International Working Group on Price Indices (Ottawa Group)
Proceedings of the Seventh Meeting

Paris, 27-29 May 2003

Editor: Thierry Lacroix
November 2003
International Working Group on Price Indices (O’Dowd Group)

Proceedings of the Seventh Meeting
## Contents

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Welcome speech by the Director General of INSEE</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>SESSION 1 - Coping with changing to complex pricing schemes</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Report of the session</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Paul Haschka, Simple Methods of Explicit QA for Services in Complex Pricing Schemes</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>Alina Gluchowska, Pricing some complex products for the CPI needs (based on the Polish practice)</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>Francoise Le Gallo and Francois Magnien, Measuring the price change of mobile phone services: an arduous task</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>SESSION 2 - Use of hedonic regression</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>Report of the session</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>Jan de Haan, Direct and indirect Time Dummy Approaches to Hedonic Price Measurement</td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>Erwin Diewert, Hedonic Regressions: A Review of Some Unresolved Issues</td>
<td></td>
<td>71</td>
</tr>
<tr>
<td>Heymerik van der Grient and Jan de Haan, An Almost Ideal Hedonic Price Index for Televisions - Room document</td>
<td></td>
<td>111</td>
</tr>
<tr>
<td>Masato Okamoto, Comparison of Hedonic Indices Compiled using Different Types of Weights - Room document</td>
<td></td>
<td>123</td>
</tr>
<tr>
<td>SESSION 3 - Elementary aggregation, superlative indexes</td>
<td></td>
<td>149</td>
</tr>
<tr>
<td>Report of the session</td>
<td></td>
<td>149</td>
</tr>
<tr>
<td>Bert M. Balk, Price Indexes for Elementary Aggregates: The Sampling Approach</td>
<td></td>
<td>151</td>
</tr>
<tr>
<td>Mick Silver and Saeed Heravi, Why Price Index Number Formulae Differ: Economic Theory and Evidence on Price Dispersion</td>
<td></td>
<td>175</td>
</tr>
<tr>
<td>Robert Cage, John Greenlees and Patrick Jackman, Introducing the Chained Consumer Price Index</td>
<td></td>
<td>213</td>
</tr>
<tr>
<td>SESSION 4 - Financial services, including insurances</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Report of the session</td>
<td>247</td>
<td></td>
</tr>
<tr>
<td>Lionel Viglino, Insurance and quality adjustment: excess and option-cost method</td>
<td>251</td>
<td></td>
</tr>
<tr>
<td>Martin Ribe, Financial Services in Swedish Price Indices</td>
<td>257</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SESSION 5 - New products, substitution between products and outlets</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report of the session</td>
<td>271</td>
</tr>
<tr>
<td>Timo Koskimäki and Mari Ylä-Jarkko, Segmented Markets and CPI Elementary Classifications</td>
<td>273</td>
</tr>
<tr>
<td>Rósmundur Guðnason, How do we measure inflation? Some measurement problems</td>
<td>289</td>
</tr>
<tr>
<td>Alexandre Makaronidis, On the treatment of newly significant goods and services in the HICP</td>
<td>321</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SESSION 6 - E-commerce</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report of the session</td>
<td>329</td>
</tr>
<tr>
<td>David Fenwick, Anna Brueton and Adrian Ball, Internet retail channels in Price Indices: the challenges involved in including non-traditional retailers in the UK Retail Prices Index</td>
<td>333</td>
</tr>
<tr>
<td>Yoel Finkel and Merav Yiftach, E-commerce in the Israeli CPI</td>
<td>349</td>
</tr>
<tr>
<td>Robin Lowe, E-commerce and consumer behaviour</td>
<td>363</td>
</tr>
<tr>
<td>Masato Okamoto and Eiji Nitta, The Present Situation of E-Commerce and a Pilot Survey on E-Commerce</td>
<td>375</td>
</tr>
<tr>
<td>Andrea Greman and Keith Woolford, E-commerce and Data Capture Opportunities for Price Indexes</td>
<td>381</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SESSION 7 - ILO activities on CPIs</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report of the session</td>
<td>395</td>
</tr>
<tr>
<td>International Labour Office, Proposals for a draft resolution concerning Consumer price indices</td>
<td>399</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SESSION 8 - Future directions, next meeting</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report of the session</td>
<td>419</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPENDIXES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Agenda of the meeting</td>
<td>423</td>
</tr>
<tr>
<td>2. List of participants</td>
<td>427</td>
</tr>
</tbody>
</table>
Introduction

This volume contains revised and lightly edited versions of papers presented at the Seventh Meeting of the International Working Group on Price Indices (the Ottawa Group). This material is complemented by reports of the various sessions arranged for discussions. The meeting was hosted by the Institut National de la Statistique et des Études Économiques (INSEE) of France in Paris from 27 to 29 May 2003.

Terms of reference of the Ottawa Group

The Ottawa Group is a city group set up in 1994 with the approval of the United Nations (Statistical Commission). It provides a forum for specialists and practitioners who work for, or are advisors to, national statistical agencies or international organisations to exchange their experiences and thoughts on crucial problems of measuring price change. The Group's strength is based on its professional authority, independence, and usefulness to national and international statistical agencies.

Without avoiding theoretical issues, the focus of the Group is on applied research, particularly though not exclusively, in the area of consumer price indices. The Group examines advantages and disadvantages of various concepts, methods and procedures in the context of realistic operational environments, supported by concrete examples whenever possible.

Only specialists actively involved in the application of the principles relating to the topics under discussion are invited to the Group's meetings as active participants. The proceedings from the sessions are edited and diffused. They contain the presented papers and the most important elements of discussions. They also include the Group's recommendations when a preponderant opinion clearly emerges from the discussions, or, if this is not the case, a summary of the discordant opinions with balanced commentary.

The Group may also assemble and publish compendia of materials related to specific topics of price statistics, composed of papers and of summaries of discussions from several meetings. Each of these publications could constitute chapters of a Handbook describing best practices in the given area of price statistics.

Organisation of the Group

The Group has a Steering Committee, which ensures both the continuity and evolution of the Group's activities. The Committee comprises representatives of the agencies that host the Group's recent or upcoming meetings together with others agreed to by the membership from time to time. The representative of the Australian Bureau of Statistics is currently serving as executive Secretary.

The Committee outlines long-term activity plans for the Group and proposes topics for the meetings. Meetings are organised in principle every 12-18 months, with topics established in advance for at least the next two meetings. The Committee extends calls to selected agencies for the submission of papers on the agreed topics and is responsible for the selection of papers to be presented. Their authors are invited to the meeting, possibly with other specialists whose
contribution to the discussion on a specific topic is considered useful. The number of active participants is kept limited for the sake of efficient discussion, however the host agency may invite a reasonable number of additional participants.

Although the meetings may be divided into sessions, these are not organised in a parallel manner. Each session is devoted to one, clearly defined topic with a designated moderator who is also responsible for producing a summary of the discussions and recording any recommendations.

The host agency provides facilities for the meeting and arranges for the hard-copy publication of proceedings. The participants bear the cost of travel, accommodation and subsistence during the sessions.

The copies of proceedings and information about the Group are also available on the Internet. The agency providing secretariat services (currently the Australian Bureau of Statistics) is responsible for maintaining a Web-site open to the public (www.ottawagroup.org). The Steering Committee may also decide to edit and periodically release compendia on selected topics, whenever it is warranted by the status of available materials.

### Previous meetings of the Group

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Date</th>
<th>Hosting Agency</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>November 1994</td>
<td>Statistics Canada</td>
<td>Ottawa</td>
</tr>
<tr>
<td>Second</td>
<td>November 1995</td>
<td>Statistics Sweden</td>
<td>Stockholm</td>
</tr>
<tr>
<td>Third</td>
<td>April 1997</td>
<td>Statistics Netherlands</td>
<td>Voorburg</td>
</tr>
<tr>
<td>Fifth</td>
<td>August 1999</td>
<td>Statistics Iceland</td>
<td>Reykjavik</td>
</tr>
<tr>
<td>Sixth</td>
<td>April 2001</td>
<td>Australian Bureau of Statistics</td>
<td>Canberra</td>
</tr>
</tbody>
</table>

### The seventh meeting

The papers were grouped into eight categories for purposes of discussion. The sessions and chairs are shown in chronological order below (see detailed program of the meeting in appendix 1).
### Session

**Coping with changing to complex pricing schemes**
- George Beelen

**Use of hedonic regression**
- Mick Silver

**Elementary aggregation, superlative indexes**
- Keith Woolford

**Financial services, including insurances**
- Thierry Lacroix

**New products, substitution between products and outlets**
- Bert Balk

**E-commerce**
- Rósmundur Guðnason

**ILO activities on CPIs**
- John Greenlees

**Future directions, next meeting**
- Timo Koskimäki, Keith Woolford

In total, some 19 papers were presented for discussion with a further 3 papers submitted as room documents. All papers are included in this volume grouped according to topic and in order of presentation.

Chairs were charged with the task of summarising discussions for each session and these summaries precede the papers. Where appropriate, they also include the Group’s recommendations for statistical agencies.

The last two sessions were specific: session 7 was devoted to ILO activities on CPIs and gave the Group the opportunity of providing comments on the draft revised CPI resolution to be approved at the ICLS meeting in November 2003; the future of the Group was discussed in session 8, that aimed in particular to agree about the topics of the next meeting.

The steering committee will report to the United Nations Statistical Commission and send them a copy of these papers and proceedings.

The next meeting of the Group will be hosted by Statistics Finland and held in Helsinki from 23 to 25 August 2004. A list of potential topics for this meeting, including following seven themes, was prepared by the Steering Committee for consideration of the participants:

- Price indices for services
- Relationships between PPI’s and CPI’s
- Housing
- Health
- Sampling
- Quality assurance of price indices
- Price index data processing — automation of the production process

This list of topics was generally welcomed. Some clarifications and additional topics were proposed during the discussion (see report of session 8 for more detail).

Participants also discussed the general role of the Ottawa Group and the ways in which the group contributes to the international statistical community. Generally, the work of the
Ottawa Group was seen as important and good practices presented in the meetings of the Group have been adopted as standards by statistical agencies and institutions. However, there was a general agreement that the Ottawa Group should, if possible, provide more formal feedback on its activities. It was proposed and accepted by the meeting that, once the new international manuals on price indices have been published, future discussions of the Group should take into account any relevant sections of the manuals. The meeting also wished to express its appreciation of the new ILO Manual on Consumer Price Indices and its thanks to all that have contributed to the work. Finally the meeting stressed the importance of ensuring the new ILO manual and the forthcoming new ILO resolution on Consumer Price Indices be consistent.

The Steering Committee has accepted the offer of the UK Office for National Statistics to host the meeting after Helsinki.

Finally, I would like to thank all the people who got involved in the organisation of this event and contributed to a pleasant, interesting and efficient meeting:

- my colleagues of the Steering Committee - Bert Balk (Statistics Netherlands), George Beelen (Statistics Canada), John Greenlees (BLS), Timo Koskimäki (Statistics Finland), Keith Woolford (ABS) - and Dominique Guédès (INSEE, Consumer Prices Division) for their assistance in planning for this meeting and their shrewd advice;
- the chairs - Steering Committee members, Rósmundur Guðnason (Statistics Iceland) and Mick Silver (Cardiff University) - who kindly conducted the debates and effectively prepared reports of the different sessions;
- all participants for the quality of presented papers and their active contributions to discussions;
- Aline Savignac, Lucile Chevret and Guy Fache (INSEE, Statistical Co-ordination and International Relations Directorate), for their invaluable help, especially regarding logistics, secretarial work, management of the website and translation.

I hope that the contents of this volume will be of some use for CPI statisticians while encouraging further research on price indices.

Thierry Lacroix
Statistical Co-ordination and International Relations Directorate
INSEE, Paris
Welcome speech by the Director General of INSEE

It is my pleasure to welcome, in the name of INSEE, the seventh meeting of the International Group on price indices, more commonly known as Ottawa Group. INSEE is indeed very attached to the smooth running of this group, all the meetings of which it has taken part since its establishment in 1994.

I beg you firstly to excuse us for being late in performing this event. For internal reasons, it has not been possible for us to hold the initially planned date of November 2002. But this further extension of time has all the same some advantages:

- it will have permitted to the authors to improve their contribution;
- you will be in a better position to take advantage of the charms of Paris, in one of the most pleasant periods of the year.

Let me now say a few words on my idea of the role of the city groups and more particularly of the one of the Ottawa Group. A city group is a flexible structure, of collegial type, which allows the best international experts of a domain - here the consumer price indices - to exchange their experiences and to inform on their research work. To be efficient, the discussions must in as much as possible be based on written contributions. To have papers also allows to disseminate them outside of the Group and therefore to reach a wider public.

If one held to these objectives, the works of the Ottawa Group would not be very different from those presented in conferences, seminars or other international symposia which, let us acknowledge, are numerous in the domain of consumer price indices. Now, according to me, a little more can be expected from a city group in its objectives and in its operation modes. With no exclusion of the purely theoretical thoughts, the works of a city group have an operational objective, which is to improve the methodological quality of the CPIs such as they are actually compiled. To gather researchers and persons in charge of operational units in statistical agencies appears to me in this respect as being a good mean to avoid two drawbacks:

- the one of academicism with thoughts which would be disconnected from reality or the costs of implementation of which would be exorbitant;
- the one of mediocrity with proposals lacking of scientific foundation or limited in their ambition because taking into excessive account the current conditions of production.

The balance between researchers and practitioners appears to me well respected within your group, which is likely to favor the emergence of relevant and realistic solutions. The presence of representatives of the main international organizations and the geographical diversity of the participants are also suitable for the plurality of the approaches and points of view.

Another specificity of a city group, the one of being strength of proposal. In the case of the Ottawa Group, this can be understood in two different ways:
directly, each of the working sessions having to result in recommendations or in identifying future axes of research when more thorough examinations are required. When divergent points of view appear, they should also be reflected in the conclusions and, above all, the reasons of these divergences should be made explicit;

more indirectly, the works performed in the framework of the Ottawa Group should, in as much as possible, be able to be considered as examples of "best practices" and it is indeed the role of the Steering Committee of the Group to select the papers in that direction. It is also important that the works of the Group receive a wide dissemination. Publishing of the proceedings of each of the Group meetings, notably through Internet, is the best mean to make the thoughts of the Group more popular and thereby more efficient. The Group must also be heard by international organizations, notably UN and ILO, through reporting its works, through informing about its proposals and giving its opinion on the projects conducted by these organizations.

The shifts taken at the time of the last two meetings of the Group, in Iceland and still more in Australia, follow the good direction hoped for. It is suitable to carry on with them. I take advantage of that to congratulate Keith Woolford, of the Australian Bureau of Statistics, to have been so kind as to take hold of the fate of the Group, the Secretaryship of which he ensures. All the participants to the Canberra meeting could appreciate the professionalism, the efficiency and the quality of the welcome that our Australian colleagues showed. If it was needed, these are signs that the future of the Group rests in good hands!

Let us now come to the topics which will be dealt with during these three days.

Two of them are about issues specific to the measure of prices in services. The services occupy in most of the developed countries a prominent place in household consumption. Now one must acknowledge that, till now, the research in the domain has remained rather limited. It is only at the time of its last meeting that the Group began to take interest in this domain, through dedicating a session to telecommunications services and another one to financial services. The initial agenda of Paris meeting planned to devote the major part of the works to the services domain in a twofold form:

- an analysis of the problems of complex pricing schemes, notably when a bunch of products, most often services, are sold in the form of package or when there is a link between the consumer's characteristics or his (or her) consuming behavior and the price of the products bought;

- a more thorough study of some sectors raising specific difficulties such as insurances, financial, health or social protection services.

The contributions related to these topics have been less numerous than expected: this must not be interpreted as a lack of interest for these questions but rather as the expression of difficulties to approach a research field where problems are complex, keep developing and where the data are often insufficient and difficult to obtain.

The statistical agencies represented within the Group have on the other hand been inspired by the theme of e-commerce. Several contributions describe thus the efforts undertaken to assess the importance of e-commerce in households consumption, study its specificity regarding
other distribution formats and develop strategies to integrate the products thus sold in the current production of CPI.

Besides, several papers deal with more “classical” subjects such as the formulae of index with different levels of data aggregation, the treatment of substitution effects or else the use of hedonic regressions in the treatment of quality effects. Though not coming under the preselected themes, these papers have been retained by the Steering Committee of the Group for two reasons:

- their interest regarding the general objectives of the Group;
- the proximity or the complementarity of the researches undertaken which has allowed to set up consistent sessions fed with sufficient material.

At last I would not be thorough if I did not mention the specific session devoted to the activities of ILO on CPI. This dissemination and updating system of the new international manual on CPIs will be unveiled, then ILO will present the provisional project of resolution on the CPIs which will be submitted at a next meeting of the Conference of European Statisticians. The main objective is to get feedback from the Ottawa Group about this project, notably on some particular points.

It is the use for the statistical agency which hosts every 18 months the meeting of the Group to present some recent national progresses in the field of CPI. In the case of INSEE, these progresses can be regrouped in two sets:

- a modernization of our collection, exchange and check of information system on sampled products (description of the product, price, nature of the quoted price). We have been working for five years on a project based on the use of portable data capture (with handheld computers) for the collection, the total computerization of information exchanges between collectors and regional and national levels of management and the setting up of high-performance checking tools at all levels. This project comes into force during this quarter. Besides the improvement of data quality and the cut in production time, this project leads to a marked cut of the current production cost of CPI. Except for collection, the whole number of people working for the CPI as an information system should decrease 40%. The managerial dimension of this project is besides very important since the role and organization of regional levels in the compilation of information system have been deeply reviewed;

- a permanent improvement of our statistical methods in order to provide the users a range of indices of the best quality as possible. For that respect the efforts of France come more and more within a European viewpoint. The harmonization of price indices in European Union, strengthened by the setting up of monetary union in 1999, requires to combine relevance and comparability. The balance between these two terms is difficult considering the tense budgetary context that are facing most of the statistical agencies in the world. It seems that, till now, INSEE did not badly succeed in this exercise if one judges after two recent operations of international benchmarking. Two recent assessments by experts of the French CPI, one performed by Eurostat, the other by IMF have actually led to very positive comments on our price index.
As a conclusion, I would like to thank all those involved in the organization of this event, the members of INSEE staff first of all, then the Steering Committee of the Group and lastly, all those who, through their work and their thought, are going to contribute to the different sessions.

I hope that you will derive a benefit from the numerous exchanges these three days sessions will provide and that you will also find some time to take advantage of your stay in Paris.

Jean-Michel CHARPIN

Director General of INSEE
Session 1 - Coping with changes to complex pricing schemes

Chair: George Beelen, Statistics Canada

Summary of session

CPIs attempt to measure price change of products of consistent quality. But products change frequently, creating the need for sample updating and quality adjustment. The papers in this session involved dealing with changes to products where complex pricing strategies raise several issues: tracking over time the price change in products of equal quality and, when fees are charged through complex plans, defining the overall price itself and taking account the high level of substitution between product offers.

The paper by Haschka deals with explicit quality adjustment methods in cases of changes in the characteristics of service products. Data on replacement rates and quality adjustment procedures used show that the impact of product change is potentially highly significant, with about 2.9% monthly (or over 30% annual) average replacement rate. Price changes often occur at the same time as changes in product characteristics, so it is often inappropriate to assume that observed price differences are entirely due to quality change.

The approach to explicit quality adjustment is based on the idea of “purpose” and practicality, using readily obtainable information and relatively simple, low-cost methods that can be performed consistently and quickly enough to maintain monthly production targets. The goal is to “be approximately right rather than perfectly wrong” in making quality adjustments.

In the case of package tour travel services, the full option price for minor changes in product characteristics (rather than only a proportion of the option price) is treated as quality change on the grounds that the range of available services is large enough that consumers purchasing a package tour value all its features. Explicit quality adjustment is not attempted for major changes in region or location of package tour. The overlap method is used for these cases, which implicitly assumes that any observed price differences are due to quality change.

The factors considered in determining direct quality adjustments are illustrated for changes in package tours, hotel services, driving schools caused by a change in legal requirements for getting a driving license, parking fees, hospital fees, telephone charges and banking fees.

The paper by Gluchowska deals with another type of change in product characteristics, that of inducements, bonuses or extras. Over 50 examples of inducements and their treatment for CPI purposes are illustrated. Inducements can be treated as full or partial quality change or they can be ignored if considered temporary and insignificant. Methods of directly valuing inducements include:

- pricing extras if the extra is sold separately, either in the same outlet or another outlet;
- estimating a value of the extra based on the time used to produce the extra (especially for services);
- price of the main product (without the extra) may be imputed based on prices of other products in the same group;

- obtaining information from outlets or producers to estimate the price of the main product without the extra.

Inducements or extras judged to be significant relative to the main product, such as, for example, a CD included with a magazine, the price increase was treated as the value of the extra, so no price change was shown for the magazine itself.

The paper by Le Gallo and Magnien illustrates the particular challenges of measuring price change for very complex services, that of mobile telephony services. These services are characterized by highly complex and frequently changing pricing schemes which make it difficult to know fees charged to consumers in connection with the range of their micro-consumptions. Moreover consumers frequently switch to more optimal products. A constant utility price index is approximated by a consumer profiles approach. The method uses an approach of tracking the minimum expenditure required to satisfy usage patterns of selected consumer “profiles”. Typical CPI methodology is to sample and track price change in products of consistent quality. For mobile telephony, quality adjustments are a big issue as they may be twofold: when products change and, within the same product, when the consumption patterns of the users evolve over time. Instead, in this new approach, a sample of consumer usage patterns is chosen and prices are tracked by choosing from the range of service plans available the plan that minimizes costs for each consumer usage profile. The method thus avoids both problems of quality adjustment...

The paper explains the method and the challenges involved in applying it to mobile telephony for both service packages and prepaid cards. The assumption of instant migration to optimal plans is then relaxed by introducing “friction” in the rate at which consumers move towards their optimal plan but the with-frictions approach has two important drawbacks: the great complexity of the model and its lack of robustness.

The discussion portion of this session supported a variety of approaches to making direct quality adjustments for product change. It is good practice for statistical agencies to track rates of product replacement and the quality adjustment methods used by product. Clarity of the basis for quality adjustment is needed to avoid excessive subjectivity in applying direct quality adjustments. For some important product categories, hedonic models may be justified. It was suggested that a cost-of-living utility framework might provide a useful basis for making direct quality adjustments.

The relatively high rates of product replacement shown (often 30% to almost 100% per annum depending on the product) provide a reminder that the concept of a “fixed basket” does not really exist at the level that prices are actually collected. The basket is “fixed” in the sense of holding weights fixed at some (usually) fairly detailed level, but not at the individual product level for which prices are collected.

The treatment of the introduction of parking fees in a downtown area was questioned. It was acknowledged that to a certain extent the introduction of parking fees does increase quality by rationing demand. But if no price change is shown, the use of the index as a deflator in the National Accounts would result in increased production in GDP due to the introduction of a
price on downtown parking. It seems unreasonable to ascribe the introduction of a parking fee entirely to quality change.

For driving school training, the purpose could be interpreted as the obtaining of the driving license and thus the increase due to added training requirements for a license could be seen as price increase. However, the change did involve additional driving practice which argued for at least partial quality change.

It was observed that inducements are predominantly marketing strategies, as are random price movements. It was suggested that when not of very high value relative to the main product, most may be safely ignored for CPI purposes since extras are often obviously temporary. This would restrict the set of extras for which quality adjustment decisions must be made. Another and perhaps more careful solution for statistical agencies would be to use simple methods to deal with minor inducements (implicit methods, or very simple explicit ones like quantity augmenting).

In some countries, consumers can easily switch suppliers for prepaid plans for telephony services. In those circumstances, it would be appropriate to group suppliers of prepaid plans and assume that consumers switch instantly to the supplier providing the minimal cost option for their consumption profile.

Consumer profiles must be revised periodically. The next generation of cell telephony services will include added services not currently accounted for in the profiles. The measurement of price change is further complicated by bundling of telephony services with other services such as television or internet services.

**Recommendations for statistical agencies**

1. Track and publish summaries and analysis of replacement rates, quality adjustment methods and frequency of inducements used by product, as a tool to promote development and application of a common set of principles and methods.

2. Further research is merited into developing a cost-of-living utility framework to be used as a basis for making direct quality adjustments. When the option cost method looks appropriate to estimate quality change (e.g. for inducements), attention should be paid to select the value of the coefficient applied to the full option.

3. Inducements that are known (or judged likely) to be temporary can be ignored or dealt with simple methods unless a significant proportion of the value of the main product. Guidelines should be developed to define more rigorously what constitutes significance for such inducements or extras.

4. For complex pricing situations subject to frequent changes in product characteristics or pricing plans such as mobile telephony, the use of a consumer profiles approach has great promise and merits further research.

5. To better reflect reality, research is needed on the best means of factoring “friction” into the consumer profiles approach.
Simple Methods of Explicit QA for Services in Complex Pricing Schemes

Paul Haschka

Statistik Austria

The author wishes to thank Mrs. H. Schimak and Mrs. R. Wucherer for their management of the local and central price collection and for their contributions and comments to this paper. Further important comments were made by my colleague Mrs A. Beisteiner.

Abstract: The paper addresses cases where more or less complex pricing schemes for services are normally applied. A number of case studies for Austria for different areas are presented. According to EU Regulations it was tried to use explicit methods for quality adjustment, i.e. explicitly estimating the amount of the difference in quality. It was also tried to use relatively simple methods which can be applied without comprehensive additional surveys and without the use of econometric or hedonic methods.

The cases studied relate to:

• QA for cars – in order to explain the principle
• Package tours – changes of the hotel services and of the duration of the tour
• Hotels – in comparison to package tours
• Driving schools – changes in the legal situation
• Hospital fees vs. Parking fees – raises from a zero price to a positive price
• Telephone charges – changes in the structure of the market suppliers and in the consumption habits of the consumers
• Banking fees – changes in the packages and in the consumption habits of the consumers

The paper might be seen as a contribution to increase the number and the range of explicit QA procedures in an area where up to now only few studies have been made available.

1. Introduction

Statistik Austria is a Central Office of medium size. In addition it was not possible to recruit many staff for academic research. Therefore we focused on relatively simple QA methodology which use much of the available information and comply with EU legislation. It was important to develop QA methods which are not too expensive and which can be applied relatively quickly in order to maintain the monthly publication. Discussion of hedonic

1 The paper is also part of a research project (Lot 1) of Statistik Austria. It will be further elaborated and be part of the reports to Eurostat.

2 The paper reflects the views of the author only and not necessarily those of the federal institution Statistik Österreich.
methods often mention the potential high costs (e.g. Schultze, C. and Mackie, C. (2001) and Koskimäki and Vartia (2001)). The aim was also not only to be cost efficient but also to be significantly above minimum standards for QA, although the perhaps highest and most perfect standard might not be reached. The principle to be applied might be abbreviated as: **Preferably be approximately right than perfectly wrong.**

The Austrian system for the treatment of prices includes the quality adjustment procedures for the Austrian CPI as well as for the HICP. For methodological, technical, organisational and practical reasons the price collection and all validation refers to the same set of prices for CPI and HICP. Therefore all improvements that have been made for QA apply for both indices. Such improvements result from the implementation of new EU-Regulations and from improvements for national uses. The latter are only a tiny minority because of the innovative speed derived from the EU-Regulations.

In EU-Regulation Nr. 1749/96 Article 5 minimum standards for procedures of quality adjustment for the HICP are laid down as follows (from HICP Compendium, Eurostat 2002a):

1. HICPs for which appropriate quality adjustments are made shall be deemed to be comparable. Where quality changes occur, Member States shall construct price indices by making appropriate quality adjustments based on explicit estimates of the value of the quality change. In the absence of national estimates, Member States shall use estimates based on information provided by the Commission (Eurostat) where these are available and relevant.

2. Where no estimates are available, price changes shall be estimated as the difference between the price of the selected substitute and that of the item it has replaced. In no case should a quality change be estimated as the whole of the difference in price between the two items, unless this can be justified as an appropriate estimate. Where replacements have to be made after goods or services have been offered at reduced prices, those replacements should be selected according to their similarity of utility to the consumer and not according to similarity of price.

Except for cases with really very small changes it was tried to follow the procedure mentioned in paragraph 1. Paragraph 2 procedures (quality change as the whole of the price change with justification) is only applied for really big changes.

2. **Overview: quality adjustment in selected COICOP classes in the Austrian HICP in 2002**

Table 1 below is derived from the Austrian CPI/HICP. The distinction is made between operational procedures rather than between major and minor changes in product specification.

3. **The principle of supported judgemental QA (example: cars)**

3.1 **Minor changes (options, small improvements of the machine)**

The principle can be illustrated with the well known option price method for cars: In case of a former option becoming standard now in a certain kind of car models several countries in the EU ask for the price of the former option. The approach is then that for some of the consumers the improvement is so big that they are in a position to pay more for it (namely the former price of the option) while for others the improvement was not worth the price. But still it should be noticed that the additional extra is a priori seen as an improvement or at least as equal to the predecessor.
Table 1

<table>
<thead>
<tr>
<th>Price collection and QA statistics and kind of QA procedure</th>
<th>07.1.1 Cars</th>
<th>09.6.0 Package tours</th>
<th>11.2.0 Hotels</th>
<th>CPI/HICP Overall Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>12N Number of price observations per year</td>
<td>1501</td>
<td>1068</td>
<td>4788</td>
<td>481.220</td>
</tr>
<tr>
<td>12# Number of changes per year</td>
<td>81</td>
<td>35</td>
<td>23</td>
<td>14.162</td>
</tr>
<tr>
<td>N Number of price series per month</td>
<td>125</td>
<td>89</td>
<td>399</td>
<td>40.102</td>
</tr>
<tr>
<td># Average number of changes per month</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>1.180</td>
</tr>
<tr>
<td>% Percent changes = replacement rate</td>
<td>5.4</td>
<td>3.3</td>
<td>0.5</td>
<td>2.9</td>
</tr>
</tbody>
</table>

of which (% of all changes):

<table>
<thead>
<tr>
<th></th>
<th>100</th>
<th>100</th>
<th>100</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q4 Full price change is deemed as quality change</td>
<td>12</td>
<td>6</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>Qx Part of the price change is due to quality change</td>
<td>54</td>
<td>49</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>QZ change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q0 Price and quality move in different directions</td>
<td>4</td>
<td>11</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>S Essentially equivalent, no quality change, full price change</td>
<td>26</td>
<td>14</td>
<td>17</td>
<td>54</td>
</tr>
<tr>
<td>W4 Major outlet change, price difference is equal to quality change</td>
<td>0</td>
<td>6</td>
<td>39</td>
<td>6</td>
</tr>
<tr>
<td>W0 Small outlet change, essentially equivalent</td>
<td>4</td>
<td>14</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Most of the EU-countries who practice option pricing use the share of 50% of the former option price for quality adjustment (Final report from Task Force QAS, Eurostat 2002c). This takes into account that not all of the consumers appreciate the option by being prepared to pay a higher price for it.

A similar calculation is performed for changes in the power of the machine which is also assumed to be seen as important for half of the consumers and negligible for the other half.

Moreover an informal tentative survey has been performed among Austrian car dealers about the share of their customers being willing to pay higher price for additional options. The popularity of options ranged typically between 20% and 70%, extreme shares like 5% or 95% did not occur. This again is an argument to approximate quality of the option by 50% of its former price.

For changes in fuel consumption the assumption is slightly different: All (100%) consumers will perceive a lower fuel consumption as better quality which can be translated into real money. In such cases the difference in operation costs for three years is calculated based on average car use and average fuel costs. The full result is taken as quality improvement because it is assumed to relate to all consumers.

If the result is an improvement but the price change is a decrease – price and quality move in different directions – the price change is increased by the amount of the quality change.
3.2 **Major changes (mainly size)**

In a new approach Eurostat tried to elaborate the conceptual frame of the HICP as being an index of fixed basic purposes (Eurostat 2003). The purpose (related to use) of the car in Austria is seen as having three dimensions of roughly the same importance. These three dimensions are:

- **Size of the vehicle** (perhaps the most important, but difficult to measure e.g. by length, height, weight, luggage space.....)
- **Power for transportation** (variables like motor power, maximum speed, torque etc., changes can be treated as mentioned above)
- **Luxury** (represented by the extras, changes can be treated with as above).

Major changes have an impact on one or more of the size variables. In that case the numeric difference of the variable is calculated and one third of it (according to the three dimensions) is compared with the price difference. If these are of the same size or if the quality change is bigger, then link-to-show-no-price-change is applied, otherwise the share of the quality change is rounded to (normally) 75% or (in some cases) 50% or 25% of the price change.

If the result is an improvement but the price change is a decrease – price and quality move in different directions – only the direct price change is taken into account because the calculation of the value of the quality difference is not as precise as above.

4. **Package tours – changes of the hotel services and of the duration of the tour**

In this area changes of the following type occurred:

- Changes from Bed&Breakfast to Half board and vice versa
- Changes from three meals per day to ‘all inclusive’ and vice versa
- Additional extra (a concert was included)
- Changes in the duration of the trip (one day difference for bus travels)
- Changes of the hotel and of the location

The first four types of changes were deemed to be smaller, i.e. explicit QA was tried. In the fifth case the overlap linking procedure was used. The underlying assumption was that the first four types of changes reflected typical consumer trends. Also the QA was easier because the type of the hotel and the location remained fixed. The last type of replacement, when it was not possible to find a similar hotel in the same region (without actually going there).

4.1 **Minor changes (extras in the service)**

Where an additional meal was provided the quality difference was estimated with respect to previous price differences or with price differences of similar providers. A source which is very often available for package tours is the information about an additional day or week to stay.

In case where a ticket for a concert or the like is included in the price of the package tour this price was estimated with the help of the provider.
Where small changes in the duration of the bus trip occurred the length of the trip was used in the QA procedure. The observed cases did not involve the bus to stay or drive longer the only difference were the duration and the hotels and meals provided.

In all cases above it was assumed that the range of available services is so large that for those consumers who purchased a package tour the full option price was treated as quality change.

4.2 Major changes

Changes in the region and location of the hotel (e.g. the other side of the island) occurred in combination with change of the producer. In such situations explicit QA was not applied but rather implicit methods (overlap, as described in Eurostat 2001).

The seasonal impact of the price movement is quite strong in the area of package tours, so the overlap took into account the respective month in the last year. As an improvement a 12 month average or another average might be used.

5. Hotels – only few changes in the services

The price collection of hotels refers to locations all over Austria according to the importance of tourism. It is rather focused on rural areas in the Alps.

As can be seen in the table above, for hotels changes occurred much less frequent than in the average of the CPI/HICP and also less frequent than e.g. for package tours. This might be an indication for perhaps either a steady market or an impact of the price collection. In the first case the Austrian hotels would change slower than in the other major holiday regions, in the second case this could result in a problem of Representativity in the longer term. It also has to be mentioned that the difference in the price collection (catalogues for package tours, direct price collection by mail for Austrian hotels) might have an impact on the replacement rate.

In the few cases mentioned QA was applied for improvements of the room (e.g. TV) or for changes in the kind of service provided (with/without breakfast). However, the QA procedure used is the overlap procedure, but the length of the overlap period is normally 6 months or 12 months. This rather long period compared with other areas of the CPI might improve the robustness of the overlap procedure.

As only few changes occurred it did not seem necessary to develop sophisticated QA procedures (and perhaps vice versa: as no sophisticated QA procedures existed, the replacement rate was tried to be kept rather low). But the simple number of cases – far below average – can indicate the necessity of a review of the sample or the review of replacement procedures. In any case the simple recording of the replacement rate gives an indication of potential differences which are perhaps not intended.

6. Driving school – changes in the legal situation

6.1 The situation

According to the HICP principle that the product definition should be according to the purpose of consumption the price collection for driving schools is related to the legal requirements to getting a driving license. Up to December 2002 that included:
Theoretical lessons (including papers for that and including first aid)  
20 practical driving lessons  
preparatory exam and exam in theory and practice, exam fees are normally collected by the driving school

From beginning of 2003 this was changed into:

Theoretical lessons (including papers for that and including first aid)  
18 practical driving lessons  
preparatory exam and exam in theory and practice, exam fees  
2 follow-up lessons 3 and 9 months after the exam, respectively  
2 lessons of group discussion after the follow-up lessons  
additional driving security training 6 months after the exam

The follow-up lessons can be paid within the package and will in almost all of the cases be provided by the same driving school. The driving security training is at the moment paid separately and not provided by the driving school but by the driving association clubs.

6.2 The principle of the replacement and QA procedure

The replacement takes into account that a majority (all) consumers will purchase the new variety of the service (because it is a law) and therefore the change is implemented in the first month the new service is available. It takes also into account that the two items differ in functionality and in the extent to which the purpose is fulfilled. This is done in the following way:

- The reduction of the required practical lessons is a quality decrease, but
- This decrease is fully compensated by the two follow-up lessons despite they take place at a later stage
- The two lessons of group discussions are worth significantly less than the driving lessons because they take place in a group and do not require a car, as perhaps half of the consumers prefer the change only 25% of the price is taken into account
- The additional driving security training is treated as quality improvement according to the 50% option principle

6.3 The practical implementation

The price change from the (virtual) old price to the new price was calculated in the following way:

package of the previous month (20 driving lessons deemed to include the 2 follow-up lessons)  
+ ½ price of the two follow-up group discussion lessons  
+ ½ price of the driving security training  
= virtual old price

new price = package including all follow-up and security training
The above mentioned methods and principles were developed in January 2003 within one month. The price collectors got little initial information and the method was fully developed at the end of the month. The implementation of the QA method requires two additional price observations. For both information was available from the providers. For the follow-up lessons a range was specified (70.- to 90.- €). In case of the absence of an observed price for the former option the price validation team should estimate the price with respect to the general price level of the driving school. For the security training the price happened to be uniform for all Austria and the price collectors and the validation team were provided with that information (135,40€).

The implementation resulted in an 11% price increase of driving schools, overlap linking or link-to-show-no-price-change or bridged overlap would have resulted in an almost zero price change.

7. Parking fees vs. Hospital fees (out-patient treatment) – raising from a zero price to a positive price

In both cases the price of a good which was free of charge before had a positive price because of new legal acts. The guiding principle was based on EU-Regulation 1687/98 Article 1 (3) which states (HICP Compendium, Eurostat 2002a)

(a)(3) Prices used in the HICP are the purchase prices paid by households to purchase individual goods and services in monetary transactions. Where goods and services have been available to consumers free of charge, and subsequently an actual price is charged, then the change from a zero price to the actual price, and vice versa, should be taken into account in the HICP.

This might actually be a new situation for CPI statisticians as it does not occur very often. The reason for this paragraph was the idea that price changes from e.g. 0,01 € to 15.-€ should be treated in the same way and with a similar result as changes from 0.-€ to 15.-€. Such changes are particularly important in the area of publicly administered prices.

According to the tariff Regulation 2646/98 Art. 5 only the consumption pattern before the tariff change is relevant and not the resulting pattern after the change. However this regulation leaves open the possibility for QA procedures. If the change in the tariff system goes along with a change in the quality of the service provided then QA procedures should be applied.

7.1 Parking fees

In the city centres usually the public space is short and expensive. Therefore an increasing number of cities introduce parking fees in an increasing part of their area. On the one hand the parking time is limited to 1-2 hours with typical prices of 1.-€ per hour(during daytime). On the other hand the people who live in the area can buy an annual resident parking permission (at approx. 150.-€ per year) which allows them to use public parking space.

Both kinds of parking fees are of course included in the HICP, so the price development of the existing parking tickets is observed. But in case of an enlargement of the area where parking is charged or of the new introduction of annual resident parking tickets the CPI/HICP did not show this as a price increase but as a quality increase. The new fees were brought into the index by linking.
The reason for this treatment was that the more efficient management of public space increased in fact the availability of parking space for non-residents and also the general availability of space for residents. The view is supported that often a majority of the population is in favour of such measures.

7.2 Hospital fees for out-patient treatment

A somewhat different situation occurred when the government introduced a new fee for out-patient treatment in hospitals — a service which is very common in Austria. The new fee was 11.-€ in a standard situation where the local doctor sends a patient to the hospital. The service at the hospital itself was not changed after the introduction of the fee.

Therefore the change in the charging system did not involve any quality adjustment. The introduction of a new fee was treated as price increase.

The calculation procedure for the CPI/HICP used of the consumption structure of out-patient treatment before the introduction of the new fee. A weighted average price was calculated for those consumers who visited the hospital being sent there by a local doctor (11.-€), coming there on their own (more expensive), coming there very often (maximum fee is 72.-€ per year) or with a child (free of charge). The resulting price change was a monthly change of +3% in the division 06. Health in May 2001 and an impact on the overall HICP of 0.12 percentage points on the overall inflation rate. As a matter of fact an eventual shift in the consumption pattern could not be measures until now.

Retrospectively, it has to be said that the administration was not successful in collecting fees from all patients and the gains from this measure are well below expectation. However the CPI/HICP did not (and could not) take these administrative shortcomings into account but regarded them as individual discounts. Consistently, when at the moment it is planned to abolish this fee the inflation rate can be expected to drop by 0.12 percentage points.

8. Telephone charges — changes in the structure of the market suppliers and in the consumption habits of the consumers

For telephone charges a national average price is calculated, for which weights are used according to the kind of tariff and the consumption habits. (New suppliers are included for mobile phones if their market share exceeds 10% and will be included for fixed network telephone as soon as their market share is more than 10%).

8.1 Suppliers

All telephone suppliers provide high quality services, every point in Austria can be reached all have help desks and all are controlled by a central governmental institution. Once they have reached a market share of 10% their services are considered as essentially equivalent to each other. Therefore the market share is taken as weight for the average price of the respective supplier. The weighted national average price is calculated without further quality adjustment.

8.2 Changes in the tariff structure

The price is calculated as weighted average for one hour telephone use. Changes in the tariff structure occur from time to time, e.g. special prices are charged if people wish to make more calls on the weekend or wish to select friends where calls are available at a reduced charge.
In the CPI/HCIP it is assumed that the people have selected that kind of tariff that fits them best from the given total list. All changes in that list of available tariffs of all suppliers are treated as essentially equivalent. If a new and cheaper option comes on the market this will show up as a price decrease, if a cheap option vanishes this will be reflected as a price increase.

8.3 *Changes in the consumption pattern*

The price is calculated as weighted average for the average consumption pattern (of timing: peak and off-peak calls) for each supplier and for each kind of tariff on the base of one hour telephone use. Changes in the consumption pattern of timing are deemed to have no impact on the measured inflation. Changes from a peak call to an off-peak call are therefore treated as having fully different quality (price difference=quality difference).

The reasoning for this last point seems perhaps weak at the moment. One point might be that for the issues mentioned before (changes in suppliers and tariff structure) no functional or objective difference could be detected. For the above mentioned timing consumption pattern the restriction of the time for the call and availability of telephone partners might be considered as quality issue. However it remains open whether the full price difference should be counted as quality difference.

9. Banking fees – changes in packages and in consumption habits of the consumers

For banking fees the price collection observes the prices of several standard packages of services in combination with account management. When the cheques were not available any more no QA was applied because the population had already ceased to use cheques long before. This was therefore deemed as a minor change, the replacement product was essentially equivalent, the fact that the cheques were not available had an impact of 0.-€.

Contrary to that the new possibility of internet banking is seen differently. According to Eurostat 2003 (and the earlier Eurostat 2002b) internet banking is seen as a new basic purpose. It is made from home and a PC equipment and internet access is needed. Therefore internet banking is seen as a major change to traditional banking and the introduction will not change the index level.

It might also be seen as a change of the basic purpose itself but there is no common general definition up to now.

10. Conclusions

For goods several approaches for explicit QA methods exist and numerous studies contain proposals for QA procedures including hedonics. For services the range is much smaller.

The paper describes a simplified method for QA for cars (which can of course also be used for other consumer durables) and provides proposals to apply the same principles for some areas of services and to extend the range of QA procedures also in these areas.

If explicit QA procedures are seen as generally superior to implicit ones the underlying principles might be accepted and applied in an even broader context. The implementation of
these principles in official price statistics will perhaps be guided by conventions and agreements.

It might also be interesting to test hedonic models on services like hotels or package tours and to compare QA procedures based on option prices with the results from regression calculations.

References


Eurostat (2003), The HICP as an Index defined on fixed user purposes, HCPI 03/446, Meeting of the Working Party HICP, March 2003.


Pricing some complex products for the CPI needs (based on the Polish practice)

Alina Gluchowska

Central Statistical Office, Poland

The author thanks Ms. Renata Minich and Ms. Joanna Piotrowicz for their contributions.

Abstract: Market competition induces retailers to be creative for the sake of increasing turnover. Price collectors have to cope with identifying products connected with another good or service. Retailers offer various inducements and in particular during recent years there has been seen in Poland a growing number of market changes which were very complex for pricing. The paper considers a practical approach to the treatment of products which are offered to consumers as packages or with any supplements or additions. Values of such additions vary. Sometimes it is only a small cosmetics sample added to another cosmetic of normal size, but there are also goods and services offered as supplements with substantial value, such as annual car insurance or a set of winter tyres included with a car. While the price of the “main” product remains unchanged, prices of inducements may have different price developments and changes in quality if sold as separate products. The paper classifies and describes different cases faced in statistical practice and presents methods to treat them in CPI calculations.

1. Introduction

The development of the Polish economy which took place in the nineties and at the beginning of the current decade has changed the domestic market in many dimensions. Both volume and structure of sales and its forms and marketing as well are undergoing changes. The growing importance of the services sector became established, making the economy more resemble those in western European countries. However, the share of consumer expenditure on services in total consumption of households is still lower than in some EU member states (currently about 30%) though remaining on a rising trend. There was a noticeable opposite trend in food expenditure at the same time (36.3% in 1998, against 28.2% in 2002, as estimated for the CPI weights). Apart from the economic development to a more market-based economy, these changes were also partially due to differential price developments, e.g. much faster for services but slower for food. Moreover, different classifications in use before and after 1999 confuse the comparisons. The national classification used in the past assigned goods and services rather differently from COICOP, e.g. catering before 1999 was classified as food but now as services; electricity, gas and other fuels were services according to the national classification but in COICOP “non-durables/energy”.

Although in COICOP some classes contain both goods and services because it is difficult to break them down into goods and services, some statistical institutions publish CPIs for an artificially constructed aggregate of “services”, which contains several 4-digit headings of COICOP for goods and services. This is not to show any weakness of COICOP but rather to

---

3 The views presented in the paper are those of the author, not necessarily reflecting the views of the Polish CSO.
grant that it reflects reality. There were always on the market tangible products (e.g. cars) and in tangible products (e.g. postal services), but also some "hybrids" such as dental services, which are classified as services, but what patients are most interested in are goods: fillings in their teeth.

Recent times have brought however newer forms of products offered, like a service plus a good, a service plus another service, a good with a service and a good plus another good. A vast number of them are inducements which according to Eurostat’s guidelines “should be disregarded if they are not significant”. The guidelines go on to say: “The market value of the inducement may be deducted if known but it should then be added back if the offer is withdrawn”. In practice not all of them are “not significant” and many have a known value. It is a source of some interpretation doubts which cause problems in price collection and put the comparability of price indices in question, especially if changes in quality of product are involved.

2. Sampling (item specification, a good, a service, a product, a product observed, inducements)

The sample of items in the Polish CPI (and HICP) contains about 1,800 products. The sample is purposively selected. Prices of goods and services are gathered by price collectors employed by regional statistical offices, in about 28,000 outlets selected within 309 price survey regions. Changes in the general consumption pattern were taken into account in selecting the survey regions, sales outlets, representative items of goods and services. The price survey regions were settled by the CSO in cooperation with the regional statistical offices. The survey region can be a town, part of a large city, a county (gmina) or a district. The choice of outlets was made by price collectors and regional statistical offices. Price collectors are recommended to monitor prices in the same outlets for at least one year.

A centrally fixed list of representative items of goods and services remains obligatory and unchanged for all regions covered by the price surveys throughout the year. From the description's point of view the list covers two types of representative items:

- products precisely described including their specific parameters (tight specification),
- products representing narrow assortment groups like clothes, underwear, footwear, etc., which are supplied on the market for short periods requiring general specifications. Selection of particular products as representative items within those groups depends on the price collector's decision.

Having loose rather than tight specifications allows more freedom for the price collectors, but it results in less comparability of price levels across the country. This has the effect of reducing the usefulness of the average prices which are published with the index. It may also result in collectors monitoring the prices of significantly different products when they should be very similar. On the other hand, tight specifications guarantee the comparability of prices of products observed in all regions, but at the same time may reduce the number of available price quotations. It also increases the probability that products so defined will not remain long on the market.

Sources of information on consumer prices of goods and services are the following:
- price recordings conducted by price collectors in selected sales outlets,
- price lists, regulations and decisions applicable to the uniform prices binding throughout part or all of the country, introduced by organs of central or local government, or other entities responsible for economic activity.

Prices of food, non-alcoholic and alcoholic beverages and tobacco are collected twice a month (in the first and the second half of the month), fruit and vegetables three times a month, and other goods and services once a month.

There are about 0.5m individual price quotations observed each month. Among them about 1% of products are those whose descriptions differ from those recorded in the previous month, of which some are new to the market (old ones were withdrawn) and some with temporary inducements e.g. an extra good or service included.

3. Changes in specification, price and/or quality; comparability; methods applied

The sample is updated annually but some adjustments take place each month on a small scale depending on local market supply changes. Each month there are a number of prices submitted by price collectors as being problematic, for which some price estimations must be made, taking account of possible quality changes.

From last year’s experience examples have been taken of some of these problematic goods and services (see the annex) for which inducements were regarded as “significant” and the method of treating them in the CPI had to be considered. The following four cases were observed:

- Services offered with extra goods.
- Services offered with extra services.
- Goods offered with extra goods.
- Goods offered with extra services.

Prices of some products with “extras” remained unchanged. However, in some cases the combined prices was higher than the price of the two items when bought separately, and in other cases the combined price was lower than the separate prices. There were observed products without any quality changes but the quality of the inducements added to them varied.

Problems to consider refer to the treatment of “extras”. Whether they should be treated as:

- Case A: A quality change of the product previously sold without extra good or service.
- Case B: Defining a new product as the combination of the original product plus the added good or service.
- Case C: A temporary and “insignificant” inducement which may be ignored.
Case A and B needed some estimations for which a group of methods marked as 1. in the annex was applied. This occurred when the extra good or service was treated as a substantial value added to the main product and its price was deducted. Different individually determined methods of valuation were used, such as:

- the price of the extra good or service was taken from another outlet (or even the same outlet) where it was sold as a separate product (e.g. for garments and personal care products),

- the price of the extra good or service was estimated using information on the time used to produce the "extras" and the estimated hourly pay of the maker (e.g. repair of therapeutic appliances),

- the price of the main product was estimated using information on the price development of other products in the group (e.g. new motor cars),

- additional information was gathered from the outlet or the producer to estimate the price of the main product (e.g. newspapers, books and stationery).

Method 2 was applied in case C where the extra good or service was ignored and not treated as a higher quality of a product, which in practice meant that the price of the main product was directly compared with its price from the reference month. For some cases there was no price change but for some there was a price decrease or increase.

Another method, numbered in the annex as method 3, refers only to products with uniform prices within the country such as a weekly magazine. The extra good (CD) was treated as a temporary addition to the main product but the CSO considered that it increased the quality of the main product. Although the only choice available to consumers was either to buy the product including the extra item – at a higher price than before – or not to buy it, in practice it was taken into the index calculation as no price change. In this particular example the CD was added to only a single issue of the weekly magazine, and the monthly average price was taken for the further calculations.

4. Summary and issues remaining for further discussions

The methods described above of dealing with those forms of products which are new to the Polish market are often decided ad hoc and applied to the calculations. The estimation method is proposed by the regional statistical office staff but a final decision is made at the central level. Each decision however may be biased by a subjective approach of persons responsible for the index calculations.

Problems still remain for consideration, particularly:

- whether and when inducements or "extras" may be ignored (what value should be treated as "significant"?)

- if values of "extras" are assessed as deductible, should the full estimated value be deducted, or only part of it?
• can inducements be treated as a better quality of the product offered even if not all consumers are interested in acquiring them? And note that it is not generally known what consumers’ views are;

• should the price of “extras” be deducted from the price of the main product if in order to obtain the extra good or service it is necessary to spend a minimum amount of money in the particular shop?

If statisticians declare to produce their indices as computed based on the data on prices actually paid by consumers, then above questions need to be answered and more detailed recommendations should be offered to NSIs.

Observing the market development it seems that some more research needs to be done in this area of CPI calculations aimed at increasing the comparability of results.
<table>
<thead>
<tr>
<th><strong>COICOP code</strong> (or below COICOP)</th>
<th><strong>Name of the group</strong></th>
<th><strong>Description of the representative item</strong></th>
<th><strong>Extra service or good</strong></th>
<th><strong>Weight of the group in 2002 (%)</strong></th>
<th><strong>Method used</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>031211</td>
<td>Garments for men</td>
<td>Men’s trousers, cotton with synthetic fibres</td>
<td>shortening</td>
<td>0,9485</td>
<td>1</td>
</tr>
<tr>
<td>031221</td>
<td>Garments for women</td>
<td>Women’s skirt, wool with synthetic fibres</td>
<td>alteration</td>
<td>1,3219</td>
<td>1</td>
</tr>
<tr>
<td>031231</td>
<td>Garments for children (3 to 13 years)</td>
<td>Trousers (6-11 age), jeans type</td>
<td>alteration</td>
<td>0,4053</td>
<td>1</td>
</tr>
<tr>
<td>045221</td>
<td>Gas from cylinder</td>
<td>Propane-butane gas (cylinder of 11 kg liquid capacity)</td>
<td>delivery</td>
<td>0,5221</td>
<td>2</td>
</tr>
<tr>
<td>051111</td>
<td>Furniture and furnishings</td>
<td>Kitchen furnishings (about 6-8 pieces)</td>
<td>delivery, design services</td>
<td>0,8884</td>
<td>1</td>
</tr>
<tr>
<td>053111</td>
<td>Refrigerators, freezers and fridge-freezers</td>
<td>Fridge-freezer, capacity about 280 l</td>
<td>delivery</td>
<td>0,1575</td>
<td>1</td>
</tr>
<tr>
<td>053121</td>
<td>Washing machines and similar major household appliance</td>
<td>Dish washing machine, capacity 6-8 dish sets</td>
<td>delivery</td>
<td>0,1557</td>
<td>1</td>
</tr>
<tr>
<td>071111</td>
<td>New motor cars</td>
<td>Passenger car &quot;Uno 1,0 S Fire&quot;, engine capacity 999 cm3, 5-doors (standard version)</td>
<td>yearly insurance</td>
<td>0,4849</td>
<td>1</td>
</tr>
<tr>
<td>08211</td>
<td>Telephone equipment</td>
<td>Mobile phone</td>
<td>subscription and additional services</td>
<td>0,0549</td>
<td>2</td>
</tr>
<tr>
<td>111121</td>
<td>Restaurants and cafes</td>
<td>Pizza with meat, sausage, etc.</td>
<td>delivery</td>
<td>1,5290</td>
<td>2</td>
</tr>
<tr>
<td>031411</td>
<td>Cleaning, laundering and dyeing of clothing</td>
<td>Men's suit dry-cleaning</td>
<td>collecting from the house and delivery</td>
<td>0,0319</td>
<td>2</td>
</tr>
<tr>
<td>044412</td>
<td>Other services relating to the dwelling</td>
<td>Cleaning of the chimney</td>
<td>travel costs of the workman included</td>
<td>0,0248</td>
<td>1</td>
</tr>
<tr>
<td>053132</td>
<td>Installation of stoves and ovens</td>
<td>Replacement of the gas oven</td>
<td>connection</td>
<td>0,0035</td>
<td>2</td>
</tr>
<tr>
<td>056221</td>
<td>Other domestic services and home care services</td>
<td>Bed-linen laundering</td>
<td>ironing</td>
<td>0,1062</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cleaning of carpet</td>
<td>travel costs of the workman included</td>
<td>0,1062</td>
<td>1</td>
</tr>
<tr>
<td>06131</td>
<td>Repair of therapeutic appliances</td>
<td>Grinding of the spherical glass</td>
<td>fitting</td>
<td>0,0053</td>
<td>2</td>
</tr>
<tr>
<td>072311</td>
<td>Maintenance and repair of the personal transport equipment</td>
<td>Repairing of car wheel (inner tube)</td>
<td>wheel balancing</td>
<td>0,5132</td>
<td>1</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Service Description</td>
<td>Unit Price</td>
<td>Quantity</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>072411</td>
<td>Other services in respect of personal transport equipment</td>
<td>Driving course for amateurs, “B” category insurance of the driver</td>
<td>0.2283</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>091125</td>
<td>Installation of the TV equipment</td>
<td>Installation of the satellite dish decoder, subscription</td>
<td>0.0053</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>121111</td>
<td>Hairdressing salons and personal grooming establishments</td>
<td>Hair dyeing hairdressing</td>
<td>0.2743</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>032211</td>
<td>Repair of footwear</td>
<td>Repairing men's shoes cost of material included</td>
<td>0.0026</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>054141</td>
<td>Repair of the glassware, tableware and households utensils</td>
<td>Installing window shutters cost of material included</td>
<td>0.0017</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>056221</td>
<td>Services for routine household maintenance</td>
<td>Glazing a window (150-60 cm), 3-4 mm glass thickness cost of material included</td>
<td>0.1062</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Door glazing (129-54 cm), transparent glass cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Making a duplicate key &quot;Yale&quot; or &quot;Lucznik&quot; cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>061311</td>
<td>Other medical products, therapeutic appliances and equipment</td>
<td>Making dentures cost of material included</td>
<td>0.2142</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>062211</td>
<td>Dental services</td>
<td>Filling the teeth cost of material included</td>
<td>0.6282</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>072311</td>
<td>Maintenance and repair of personal transport equipment</td>
<td>Car tyres - retreading cost of material included</td>
<td>0.5132</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bicycle tyres – retreading cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>121111</td>
<td>Personal care</td>
<td>Women's cold wave cost of material included</td>
<td>0.2743</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women's hair – dyeing and brushing cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women's hair &quot;Balayage&quot;, 3 colours cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Manicure cost of material included</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>121212</td>
<td>Repair of electric appliances for personal care</td>
<td>Replacing feeding cable in electric shaver cost of material included</td>
<td>0.0018</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>123112</td>
<td>Repair of jewellery, clocks and watches</td>
<td>Replacing battery in men's digital watch cost of the battery included</td>
<td>0.0106</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>127114</td>
<td>Other services</td>
<td>Visit cards printing, per 100 pcs cost of paper included</td>
<td>0.0738</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>056111</td>
<td>Goods and services for routine household maintenance</td>
<td>Dishwashing liquid &quot;Ludwik S&quot;, per 500 g with dishcloth</td>
<td>0.4406</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Details</td>
<td>Price</td>
<td>Quantity</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>071111</td>
<td>New motor cars</td>
<td>Passenger car &quot;Matiz FRIEND S&quot;, engine capacity 798 cm³ (standard version) with set of winter tyres (Nov, Dec.)</td>
<td>0.4849</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Passenger car &quot;Renault Megane Pack Authentique&quot;, engine capacity 1400 cm³ (standard version) with set of winter tyres (Dec.)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Passenger car &quot;Lanos Daewoo&quot;, engine capacity 1500 cm³ (standard version) with set of winter tyres (Nov, Dec.)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>095212</td>
<td>Newspapers, books and stationery</td>
<td>Daily newspaper - &quot;Gazeta Wyborcza&quot; with compact disc</td>
<td>0.6265</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weekly magazine - &quot;Przyjaciolka&quot; with compact disc</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weekly magazine - &quot;Kobieta i Życie&quot; with compact disc</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weekly magazine - &quot;Pani Domu&quot; with compact disc</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weekly magazine - &quot;Gala&quot; with compact disc</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monthly magazine - &quot;Claudia&quot; with compact disc</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>121311</td>
<td>Articles and products for personal care</td>
<td>Toilet soap for children with extra bar of soap</td>
<td>0.6990</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Toilet soap &quot;Fa&quot;, per 100 g with extra bar of soap</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hair shampoo &quot;Palmolive Naturals&quot;, per 200 ml with antiperspirant (50 ml)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shower gel with antiperspirant (50 ml)</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid toilet soap, per 300 ml with balsam for firmness</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>121312</td>
<td>Articles and products for personal care</td>
<td>Lipstick with lip balsam</td>
<td>0.9326</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cleansing milk with moisturising toner</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moisturising cream with hand cream</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Active balance toner with cleansing milk</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shaving foam with after shave milk</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* in 2003
Measuring price change in mobile-telephony services: an arduous task

Françoise Le Gallo and François Magnien

INSEE, France

The authors wish to thank Thierry Lacroix for his helpful comments and cogent suggestions on this text.

The recent but prodigious growth of mobile telephony has made it the emblematic product of the new information and communication technologies (NICTs). In France, its rise has been spectacular: introduced in 1995, it now boasts a combined total of nearly 40 million individual and corporate users. Three operators share the French market: Orange (50% of the customer base at end-2002), SFR (35%), and Bouygues Telecom (15%). Individual users account for somewhat over two-thirds of their customer base.

There are two reasons for the relatively high prices of mobile-telephony services. First, they are the counterpart of the massive investments by operators for their network implementation and expansion. Second, they reflect subscriber-acquisition costs in the form of bonuses paid to service-marketing companies and subsidies for handsets.

Today, cards and packages are equally popular with the public. Cards are now by far the more common solution among individual users. This has not always been the case, as chart I shows.

1. A “constant-utility” or rather “constant-usage” index

The price/volume decomposition of mobile-telephony services is a maddeningly difficult puzzle for statisticians. Measuring price change in individual countries is tricky. The methods are diverse and lack transparency. However, a suitable approach may gradually win acceptance. It is already being applied in the German consumer price index (CPI) and, since January 2003, in the French CPI. The method has also been used—admittedly on a trial basis—in the United States, as part of a study by Hausman (1999), as well as in the United Kingdom and France by their telecoms regulatory bodies, respectively OFTEL and ART. The principle is simple: it involves tracking the minimum expenditure required to satisfy individual or corporate consumption patterns as summarized in selected “profiles.” The minimum expenditure makes allowance for the abundant supply of pricing plans from mobile-telephony operators. The principle can subsequently be tailored to factor in the time taken by consumers to adjust to the choice of optimal offerings.

It would be more appropriate to describe this as a constant-usage index rather than a constant-utility index (CUI). A true CUI for prices would require a particularly delicate exercise: the explicit or implicit estimation of a utility function (Magnien and Pougnard 2000) or demand function (Hausman 1999). That would involve an unreasonable effort within the framework of current CPI production. A “constant-usage” index is a far easier proxy to develop. With the
1. Residents of metropolitan France (mainland + Corsica), users of a SIM card or pre-paid card who have made at least one call and have not exceeded the time limit for receiving incoming calls with the card.

Sources: ART; authors’ computations; INSEE

aid of surveys, we can describe consumption patterns: number of calls, call duration, call-period distribution, number of SMSs (text messages) sent, etc. For each product offered, we can then compute the expenditure entailed by the usage defined on these criteria and determine the minimum expenditure.

To go a step further in presenting the methodology of the price index for mobile-telephony services, some formalization is needed. In each month \( m \), consumers can choose from the products \( p \)—packages or cards—put on the market by operators. Assuming that consumers are rational, fully informed, and have no constraints (these hypotheses will be discussed in the final section of this paper, where we propose an alternative model), they will opt for product \( \hat{p} \), which minimizes their expenditure \( D_{p,m} \):

\[
D_{\hat{p},m} = \min_{p} D_{p,m}
\]

(such a product, not always unique, will be called “optimal”). With \( \hat{D}_m \) as this minimal expenditure, the index using month 0 as base 100 is \( \hat{D}_m / \hat{D}_0 \).

2. A detailed segmentation into consumption profiles

It is assumed that consumers change products “instantly” as soon as a more economically attractive one appears. But reality is very different. For example, switching operators has a cost—in particular, paperwork and changing phone numbers. We therefore had to restrict the application of the proposed methodology to consumer sub-categories, identifying product ranges with a very low short-term substitutability. We accordingly separated the profiles for the three operators.

Switching from a package to a card entails costs. Therefore, we also had to classify consumers according to their contract: subscription (package) or pre-paid (card).
But a classification restricted to six consumer types or "consumption profiles" would have been very inadequate. The product that minimizes a consumer's expenditure basically depends on his or her call volume. If we amalgamate consumers whose "sizes" are too diverse in the same profile, we will summarize their consumptions by an average monthly duration. In model (1), the optimal product will diverge from the actual optimal products for the consumers in that profile. Accordingly, we have specified three consumption levels (high, medium, and low) and three call time distributions (mostly daytime, mostly evenings and weekends, mixed) for the two categories of products selected (packages and cards).

These criteria yield 54 profiles, 18 per operator. We aggregated the indices of the 54 profiles using a Laspeyres procedure so as to obtain the index for all consumers in a month $m$ measured against a base month 0:

$$I_{m}^{i0} = \sum_{T} w_{T}^{i0} \hat{D}_{T}^{m} / \hat{D}_{T}^{0}$$

In this formula, $w_{T}^{i0}$ denotes the expenditure by all consumers in profile $T$ for the use of the optimal product in month 0:

$$w_{T}^{i0} = S_{T}^{i0} \hat{D}_{T}^{0} / \sum_{T=1}^{54} S_{T}^{i0} \hat{D}_{T}^{0}$$

where $S_{T}^{i0}$ is the number of consumers in profile $T$ in period 0. We chained the resulting indices (on a monthly, quarterly or annual basis) to allow for the changing structure of the population of consumers of mobile-telephony services.

3. The survey of mobile-telephony operators on consumption profiles

To calculate the price index for mobile-telephony services, INSEE has launched an "annual survey of operators on consumption profiles." The purpose is to construct the "profiles" (or typical consumers) mentioned earlier. The first survey, in early 2002, covered consumption patterns for mobile-telephony services in 2001. The questionnaire was prepared in consultation with the three operators.

Each operator divided its customer base into consumption "profiles" defined by the successive application of three criteria: contract type (package or pre-paid), total monthly call duration (short, medium or long), and the call-time distribution (daytime, evenings and weekends, mixed).

Table 1, compiled from the survey results, shows that call volume is much larger for package users. The survey allows a finer analysis of these figures at profile level—an analysis that does not appear in the table. In particular, the number of calls and monthly call duration increase with the consumption level, irrespective of whether packages or cards are used.

After classifying their customer bases by profile, the operators described the profiles by computing mean values, listed in table 2. Averages were calculated for all available months in 2001 (or at least for the latest six) for the following: monthly duration of national calls; breakdown of monthly call duration by call period; breakdown of monthly call duration by
destination; number of calls per month; monthly duration of calls to “favorite number.” The survey also asks operators for the breakdown of customers in each profile by product.

Table 1: Number of calls and average monthly call duration\(^1\) per user\(^2\)

<table>
<thead>
<tr>
<th>Contract type</th>
<th>Number of calls</th>
<th>Actual duration of one call</th>
<th>Actual monthly call duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packages</td>
<td>87</td>
<td>2</td>
<td>175</td>
</tr>
<tr>
<td>Cards</td>
<td>21</td>
<td>0.9</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>55</td>
<td>1.8</td>
<td>100</td>
</tr>
</tbody>
</table>

1. Average data, in minutes for call duration, computed over at least the last six months of 2001 for all operators.
2. Individual users.
Source: INSEE (National Accounts Department)

Table 2: Profile descriptions

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEASUREMENT UNIT</th>
<th>NOTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total monthly duration of outgoing national calls</td>
<td>mean Minutes</td>
<td>(\Lambda)</td>
</tr>
<tr>
<td></td>
<td>distribution</td>
<td>(t_1, t_2)</td>
</tr>
<tr>
<td>Percentage of total monthly call duration consisting of calls made in daytime</td>
<td>% (\alpha(t))</td>
<td></td>
</tr>
<tr>
<td>Percentage of total monthly call duration consisting of calls to fixed telephone(^1)</td>
<td>% (\beta(d))</td>
<td></td>
</tr>
<tr>
<td>mobile phone on same network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mobile phone on another network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly number of calls</td>
<td>mean Number</td>
<td>(N)</td>
</tr>
<tr>
<td></td>
<td>distribution</td>
<td>(d_1, d_2)</td>
</tr>
<tr>
<td>Percentage of total monthly duration of national calls to “favorite” number</td>
<td>% (P)</td>
<td></td>
</tr>
<tr>
<td>Contract duration</td>
<td>% (\tau)</td>
<td></td>
</tr>
<tr>
<td>Proportion of customers subscribing to itemized billing</td>
<td>% (\tau(s))</td>
<td></td>
</tr>
<tr>
<td>Proportion of customers subscribing to call waiting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of customers subscribing to caller ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMS</td>
<td>Number</td>
<td>(n)</td>
</tr>
</tbody>
</table>

1. Includes short numbers and WAP.
Source: INSEE (National Accounts Department)
4. A detailed description of call duration and volume, with the aid of distributions

The existence of thresholds in product pricing (indivisible minutes, higher charge for calls exceeding package allowances, time limit on card use, etc.) requires a description of typical consumers using distributions rather than relying exclusively on call duration and volume.

We assume call duration obeys an exponential law with an $\mu$ mean, i.e., of density $\varphi(\Delta) = 1/\mu e^{-\Delta/\mu}$. In other words, when the consumer has been using the service for some time, the likelihood of further consumption depends only on the further duration and not on the time already elapsed. We estimated the parameter $\mu$ as the ratio of the mean monthly call duration and the mean monthly number of calls, two information items supplied by operators in the survey on consumption profiles.¹

Rather than estimating the distribution of the monthly number of calls, we estimated the distribution $\lambda$ of the total monthly call duration. This was possible because the operator survey gave us the limit durations $t_1$ and $t_2$ that allowed us to classify consumers into three quantiles according to their monthly call duration.

5. Computing expenditure

The randomness of call duration and volume makes the expenditure for each profile random too:

$$D_{p,m} = E(M_{p,m})$$

where $E$ stands for the mathematical expectation. The random expenditure $M_{p,m}$ is the sum of several terms. The first is the “basic” expenditure $A_{p,m}$, whether the product is a package (optionally adjustable) or a card. This is the expenditure strictly generated by actual telephone calls. It is highly complex to calculate, given the large number of variables involved: call periods; surcharge for minutes exceeding the packages; roll-over minutes, free of charge or not; time limit on card use; favorite number; etc. Computing this basic expenditure is so complex that we devote the entire following section of this paper to it. The second is a series of expenditures that are not generated by telephone calls in the strict sense. They include charges for SMSs and services such as itemized billing, call waiting, and caller ID.

Packages and cards offer a varying number of SMSs (10, 20 or more) for a given monthly charge. Users who send more SMSs than the chosen number pay each additional SMS at a unit price listed in the package plan. A consumer who sends an average 12 SMSs a month will get a better deal by choosing a stock of 10 SMSs and paying for two extra; a consumer who sends an average of 18 SMSs should choose a stock of 20 SMSs—even if this means “losing” two—rather than paying for eight extra. Our model specifies this rational-consumer logic.

Itemized billing, call waiting, and caller ID entail a fixed monthly charge of the form $\tau(s)\pi^p(s)$, where $\tau(s)$ is the proportion of customers in a profile who have subscribed to option $s$ and $\pi^p(s)$ the monthly cost of the option if it comes with product $p$.

¹ We were able to “test” the validity of the assumption made concerning the call-duration distribution (Magnien 2003).
The final form of the monthly-expenditure expression is thus:

\[ M^{p,m} = A^{p,m} + SMS^p + \sum_s \tau(s) \pi(s) \]

where \( s \) denotes any one of the three options chosen: itemized billing, call waiting, and caller ID. We now analyze basic expenditure \( A^{p,m} \) and its mathematical expectation \( E(A^{p,m}) \), which is used in computing the expenditure \( D^{p,m} \).

6. The complexity of call pricing is unequaled

In the field of consumer products, mobile-telephony services are characterized by a pricing system of unprecedented complexity, mainly determined by the monthly call duration \( A \).

To take the changes in pricing procedures into account when tracking price change, we need a minimum amount of information on the consumption behavior of individuals, namely (1) the proportion \( \alpha(t) \) of call minutes spent in each of the call periods \( t \) defined by the operator for a given product, and (2) the proportion \( \beta(d) \) of call minutes to the “destination” \( d \): fixed network, mobile network of the same operator or mobile network of another operator. In fact, we distinguish another destination: the “favorite” number, which, for a flat charge, can be called at a cheaper rate (Magnien 2003). With \( A^m \) as the combined call duration in month \( m \), \( C^m(i,d) \), the number of call minutes in call period \( t \) to destination \( d \) is:

\[ C^m(i,d) = \alpha(i) \beta(d) A^m \]

In summer 2002, the French press talked about a “battle of the second.” Here are the facts: while all operators had long espoused the principle of an indivisible first minute, followed by 30-second increments, SFR announced in July 2002 that it was extending per-second charges beyond the indivisible first minute. In August, Orange “responded” by offering straightforward per-second charges in its packages, albeit with offsetting provisions: for minutes in excess of the “one hour” package, the indivisible first minute would continue to apply. Also, calls to its competitors’ cell phones would be charged extra, an arrangement that Orange soon rescinded. In September, it was Bouygues Telecom’s turn to introduce per-second charging, to which SFR immediately responded by offering per-second charging as well.

How have these pricing changes been incorporated into the index? For each profile and product, we have established a billing coefficient that gives the billed call duration by simply multiplying the price by the actual call duration. Each month \( m \), the user makes \( N^m \phi(\Delta) d\Delta \) calls lasting between \( \Delta \) and \( \Delta + d\Delta \) (\( N^m \) is the total monthly number of calls and \( \phi \) their distribution by duration). The monthly call duration is therefore proportional to the monthly number of calls: \( A^m = N^m \int_{\Delta}^{\Delta+d\Delta} \phi(\Delta) d\Delta \). Operators bill \( f(\Delta) \) call minutes when the actual call duration is \( \Delta \). The monthly call duration billed by the operator is thus \( A^{p,m} = N^m \int_{\Delta}^{\Delta+d\Delta} f(\Delta) \phi(\Delta) d\Delta \). It is therefore proportional to the actual duration:

\[ A^{p,m} / A^m = \int_{\Delta}^{\Delta+d\Delta} f(\Delta) \phi(\Delta) d\Delta / \int_{\Delta}^{\Delta+d\Delta} \Delta \phi(\Delta) d\Delta \]
This ratio, which exceeds 1, is the “billing coefficient.” Its computation, under the hypothesis of a Poisson distribution of the mean call duration, is described in box 1.

**Box 1: Billing coefficient**

Let us consider a product $p$. The call duration $f^p(\Delta)$ used by an operator differs from its actual duration $\Delta$:

$$f^p(\Delta) = \begin{cases} 
\Delta^p & \text{if } \Delta \leq \Delta^p \\
\Delta^p + n\Delta^p & \text{if } (n-1)\Delta^p \leq \Delta - \Delta^p \leq n\Delta^p 
\end{cases} \quad (n \geq 1)$$

where $\Delta^p$ denotes the duration of the first call segment billed and $\Delta^p$ the duration of the following segments. The likelihood that the billed call duration will equal $\Delta^p + n\Delta^p$ is:

$$p_n = \int_{\Delta^p + (n-1)\Delta^p}^{\Delta^p + n\Delta^p} \varphi(\Delta) d\Delta = \int_{\Delta^p + (n-1)\Delta^p}^{\Delta^p + n\Delta^p} \mu^{n-1} e^{-\Delta^p/\mu} d\Delta = \left(1 - e^{-\Delta^p/\mu}\right) e^{-\frac{(n-1)\Delta^p}{\mu}} e^{-n\Delta^p/\mu}$$

for $n \geq 1$ and: $p_0 = 1 - e^{-\Delta^p/\mu}$ for $n = 0$. From this we deduce the mean billed duration:

$$\sum_{n=0}^{\infty} (\Delta^p + n\Delta^p) p_n = \Delta^p p_0 + \Delta^p (1 - p_0) + \Delta^p (1 - e^{-\Delta^p/\mu}) e^{-\Delta^p/\mu} \sum_{n=1}^{\infty} n e^{-n\Delta^p/\mu}$$

or, since the sum of the series is equal to $e^{-\Delta^p/\mu}/(1 - e^{-\Delta^p/\mu})^2$:

$$\Delta^p + \Delta^p \frac{e^{-\Delta^p/\mu}}{1 - e^{-\Delta^p/\mu}}$$

This gives us the equation for the billing coefficient:

$$\frac{\int_0^\infty f^p(\Delta) \varphi(\Delta) d\Delta}{\int_0^\infty \Delta \varphi(\Delta) d\Delta} = \frac{1}{\mu} \left[ \Delta^p + \Delta^p \frac{e^{-\Delta^p/\mu}}{1 - e^{-\Delta^p/\mu}} \right]$$

The monthly call duration billed in period $t$ to a destination $d$ depends on the product $p$ used:

$$C^{p,m}(t,d) = \alpha(t) \beta(t) \Lambda^{p,m}$$

(using equation (6)), hence:

$$C^{p,m}(t,d) = \alpha(t) \beta(d) \frac{1}{\mu} \left[ \Delta^p + \Delta^p \frac{e^{-\Delta^p/\mu}}{1 - e^{-\Delta^p/\mu}} \right] \Lambda^{p,m}$$

The following table gives an estimate of the mean billing coefficient for packages and cards:
Table 3: Average call duration, actual and billed

<table>
<thead>
<tr>
<th>Contract type</th>
<th>Number of calls</th>
<th>Actual duration of one call</th>
<th>Actual monthly call duration</th>
<th>Billed monthly call duration</th>
<th>Billing coefficient, December 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packages</td>
<td>87</td>
<td>2</td>
<td>175</td>
<td>207</td>
<td>1.19</td>
</tr>
<tr>
<td>Cards</td>
<td>21</td>
<td>0.9</td>
<td>20</td>
<td>30</td>
<td>1.52</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>55</strong></td>
<td><strong>1.8</strong></td>
<td><strong>100</strong></td>
<td><strong>121</strong></td>
<td><strong>1.22</strong></td>
</tr>
</tbody>
</table>

1. Average data, in minutes for call durations, computed over at least the last six months of 2001. Source: INSEE (National Accounts Department)

Other difficulties must be resolved to calculate basic expenditure. They are specific to the type of contract: package or card. For cards, see (Magnien 2003).

The inclusion of roll-over minutes and excess call duration—an issue specific to packages—is a complex step in the determination of the price index. In principle, the excess in a given month \(m\) depends on the consumption in that month and on any minutes rolled over from the previous month. But the roll-over from \(m-1\) depends, in turn, on the number of actual call minutes in month \(m-1\) measured against the “inclusive” minutes allowed in the package, and therefore on the roll-over of unused minutes from month \(m-2\). Using a recursive procedure, from the consumption of minutes \(A_l^i (l = m-k, \ldots, m)\) in successive months since the acquisition of the package in month \(m-k\), we can determine the roll-over for month \(m\) and thus the (random) number of call minutes exceeding the package. We assume that the unused minutes for a given month were only available in the following month. There are two reasons for this: first, in practice, operators restrict roll-over arrangements; second, we wanted to simplify the calculations. The formula for excess consumption becomes

\[
DEP_{p,m} = \Theta^p (A_m^i, A_{m-1}^i) \tag{3}.
\]

The expectation of excess consumption is what we take into account when computing the expenditure that the consumer will seek to minimize. The expectation is written:

\[
\mathbb{E}(DEP_{p,m}) = \int_0^{+\infty} \Theta^p (x, y) \lambda(x) \lambda(y) dx dy
\]

where \(\lambda\) is the distribution density of the monthly call duration estimated previously from the information gathered in the operator survey.

In the index computation, call-related expenditure for a package is therefore written:

\[
\mathbb{E}(A_{p,m}) = F_p + \tau \Delta F_p + R_p + P_p + \sum \beta(d) \mathbb{E}(DEP_{p,m})
\]

where \(F_p\) is the price of the monthly subscription, \(\Delta F_p\) the change in that price due to a non-standard contract period, and \(\tau\) the proportion of customers (specific to each profile) with a contract of that duration, \(R_p\) and \(P_p\) the monthly subscription to “roll-over minutes” and “favorite number” options, \(DEP_{p,m}\) the monthly excess call duration (the only random

---

2 See (Magnien 2003) for details on the procedure.

3 See (Magnien 2003) for details on the calculations.
component of package-related expenditure (4), $\beta(d)$ the proportion of call minutes to destination $d$, and $\delta(d)$ the price of an excess minute for calls to that destination.

Operators offer highly original pricing plans that go far beyond adding options to a standard product. We have incorporated these plans into the index calculation. Examples include "adjustable" packages and low-cost "zero-use" monthly packages to receive incoming calls only.

7. Prices stopped falling three years ago

Using the methodology described earlier, we computed the price index for mobile-telephony services from January 1999 to December 2002. The main lesson of this exercise is the overall stability of prices of mobile-telephony services in the past three years, after the sharp decline in 1999 and, no doubt, in previous years (chart 2). 4

Chart 2: Prices of packages and cards have moved in tandem

![Chart 2: Prices of packages and cards have moved in tandem](chart2.png)

Source: INSEE

This break from the previous pattern is likely due to operators' heavy investments in developing the present GSM networks and the prospects for deployment of the third-generation UMTS (Universal Mobile Telecommunications System) networks.

The stabilization of prices for mobile-telephony services since 2000 is corroborated by that of the average prices of call minutes. These prices were computed from ART data by national accountants until 2002, absent a tracking of prices of mobile-telephony services in the CPI. While average prices fell 20% between 1999 and 2000, the "constant usage" price index eased 7% in annual-average terms. Prices then effectively stabilized according to both indicators.

The steepest fall in average prices between 1999 and 2000 (also observed between 1998 and 1999 using the price index based on the ART study cited earlier) is recorded by the constant-usage index. Operators offer price cuts that are often substantial but targeted, for example, at calls to the same network or calls made in specific time periods. Consumers inevitably take advantage of these price reductions by changing their calling patterns in ways we can imagine, thereby lowering average prices. But these reductions will not be incorporated into the index since it is based on a constant usage of mobile-telephony services: the index computation will be based on a fixed distribution of calls by destination or by call periods, as provided by the operator survey.

Cards and packages do not exhibit significantly different price changes over the long run (chart 2).

The prices of mobile-telephony services measured by ART in 1998 (ART 1999) already showed a sharper downtrend for the “high-volume” consumers (i.e., individuals). The finding still held true in 1999 (chart 3a). Prices stabilized in 2000; in 2001, their movements were far more beneficial to “extreme” users (low- and high-volume). However, while operators seem to promote a diversity in consumption patterns, they also seem intent on lessening the difference between contract types. Charts 3b and 3c show greater swings in card prices for low-volume consumers and package prices for high-volume consumers: cards are more widespread among low-volume users, packages among high-volume users.

Chart 3a: Price change has favored low-volume and high-volume users

![Chart 3a: Price change has favored low-volume and high-volume users](source: INSEE)
Chart 3b: Since 2002, card prices have been more attractive for high-volume consumers...

Source: INSEE

Chart 3c: ...and package prices for low-volume consumers

Source: INSEE

8. Consumers have imperfect information

Consumers’ real expenditure exceeds their minimal expenditure calculated with (1). We estimated the excess from the operator survey, which gives the consumption breakdown by product in December 2001. At that date, for 70% of package subscribers, the cost excess was lower than 30%. We do not think it is appropriate to give a more detailed distribution of consumers by excess-cost level, as the latter is tricky to evaluate and probably overestimates
the proportion of subscribers with high cost excess. This proportion is indeed artificially increased when the computation of data relies on only one month—as here—, due to month to month fluctuations in subscribers consumption. Moreover, the computed excess cost differs from the actual excess cost, since the calculation replaces the actual requirements of consumers in a given profile (monthly call duration, distribution by call period, by destination, etc.) by their mean values. The fewer the profiles, the wider the scatter of customers in each profile around their mean; as a result, the greater the bias in the calculation of a consumer's expenditure on a product and hence of the consumer's minimal expenditure. However, by including many profiles, our study reduces the bias on the excess-cost estimation.

The supply of mobile-telephony services and, above all, their pricing are so complex that consumer choices—even if rational—can be based only on a partial knowledge of product advantages and drawbacks. Consumers are also constrained by their past choices: a product purchase involves a commitment on the part of the consumer, for example, the duration of a package contract or the total cost of a pre-paid card. The financial benefits of the product change must also outweigh its non-monetary drawbacks (paperwork, new number, etc.).

9. A model with "frictions"

The preceding analysis suggests an alternative approach to tracking prices of mobile-telephony services. It consists in assuming that consumers optimize their choices but with limited, imperfect information on product supply. The construction of a constant-usage price index with "frictions" requires a re-examination of equation (1), used until now to describe the dynamics of consumer mobility between products. Each month, all consumers migrated toward the optimal product (the one that minimizes expenditure in a friction-free setting); in the "with frictions" model, only some of the consumers whose expenditure exceeds the minimal expenditure will switch to the optimal product. These dynamics are formalized in box 2.

To implement this approach, we need additional information: the user distribution by product in the base month and in each profile. This distribution is the starting point for the new dynamics of consumer mobility. It is known for December 2001 thanks to the operator survey on consumption profiles.

The aggregation of the profile indices resembles the perfect-information procedure: in relationships (2) and (3), we simply replace minimal expenditure $D^m$ with mean expenditure

\[
\sum_p f_T^{p,m} D_T^{p,m} 
\]

where

\[
 f_T^{p,m} = S_T^{p,m} / \sum_p S_T^{p,m} 
\]

is the proportion of profile-$T$ consumers in month $m$ for product $p$.$^5$

$^5$ $S_T^{p,m}$ is the number of $T$-type consumers consuming product $p$ in month $m$ as determined by the dynamics described in box 2.
Box 2: Dynamics of consumer mobility with imperfect information

The consumer-mobility dynamics provide an explicit formula, in each profile, for the transition between its successive distributions \((S^{p,m-1})_p\) and \((S^{p,m})_p\) between the products \(p\) available in each period\(^6\). In an imperfect-information model, the formalization of these dynamics requires the introduction of a “mobility coefficient” \(\pi^{p,m}\), equal to the proportion of consumers using product \(p\) in month \(m-1\) who migrate in month \(m\) to an optimal product on the market\(^7\).

The imperfect-information dynamics are as follows:

- if a product \(p\) is not optimal in month \(m\) \((D^{p,m} > \hat{D}^m)\) then:
  \[S^{p,m} = (1 - \pi^{p,m})S^{p,m-1}\]
- if a product \(p\) is optimal in month \(m\) \((D^{p,m} \leq \hat{D}^m)\), then:
  \[S^{p,m} = S^{p,m-1} \quad \text{if } p \text{ is still on the market in month } m\]
  \[S^{p,m} = \sum_{p', D^{p',m} > \hat{D}^m} \pi^{p',m}S^{p',m-1} \quad \text{if } p \text{ is no longer on the market in month } m\]

These dynamics were initialized with the “actual” distribution for December 2001. If we admit that parameter \(\pi^{p,m}\) does not depend on product \(p\), then we can easily show that:

\[
(1 - \pi^{p,m}) \sum_{p, D^{p,m} > \hat{D}^m} f^{p,m-1} = 1 - \rho^m \quad (*)
\]

where \(\rho^m\) is the fraction of consumers optimally positioned in each month \(m\) and \(f^{p,m-1}\) is the proportion \(S^{p,m-1}/\sum_{p \in m} S^{p,m}\) of consumers of product \(p\) in month \(m-1\). The sum:

\[
\sum_{p, D^{p,m} > \hat{D}^m} f^{p,m-1}
\]

is therefore the proportion of consumers who, in month \(m\), should switch to the new product. The (*) relationship is a natural one: if we take the number of consumers who should switch and subtract the number of consumers who do not (left-hand member), we obtain the number of “poorly positioned” consumers (right-hand member).

Computing the with-frictions index for months prior to December 2001 was a problem because we did not have the consumer distributions by product to initiate the dynamics. These were therefore “reversed” as explained in box 3.

Thus, there exists in \(m-1\) a state older than the state in month \(m\) if and only if the proportion of optimal customers in \(m\) exceeds the mobility coefficient. Note that the relationship becomes trivial when there are no frictions, as \(\rho^m\) and \(\pi^m\) both equal unity. With frictions, the relationship does not always obtain, which has led us to make a slight change in the index-calculation model (Magnien 2003).

10. Results depend heavily on the mobility coefficient

We conducted the calculation of the index in 2002 and its backward extrapolation to January 1999 using different values for the mobility coefficient, which we assume to be identical for

---

\(^6\) \(S^{p,m}_T\) (or, more simply, \(S^{p,m}\)) denotes the number of consumers in profile \(T\) in month \(m\) for product \(p\).

\(^7\) When several optimal products are available on the market (an uncommon situation, but one that does occur), we assume that the migrating consumers will be evenly distributed among them.
Box 3: Consumer optimality, mobility coefficients, and backward extrapolation

Assuming the mobility coefficient is product-independent, the inversion of the mobility dynamics (box 2) within the profiles gives:

\[ S_{p,m-1} = S_{p,m} \left/ \left(1 - \pi_m^p \right) \right. \]

if \( p \) is non-optimal in \( m \)

\[ S_{p,m-1} = S_{p,m} - \frac{\pi_m^p}{1 - \pi_m^p} \sum_{p' \text{ non-optimal}} S_{p',m} \]

if \( p \) is optimal and on the market in \( m \)

\[ S_{p,m-1} = S_{p,m} \]

Thus, there exists in month \( m-1 \) a state \( S_{p,m-1} \) preceding the state \( S_{p,m} \) of month \( m \) if and only if the following condition is satisfied:

\[ \pi_m^p \sum_{p' \text{ non-optimal}} S_{p',m} \geq 0 \]

for all optimal products on the market in \( m \), i.e.:

\[ \rho_m^o \geq \pi_m^o \quad (*) \]

where \( \rho_m^o \) is the proportion of optimal consumers in a profile and \( \pi_m \) the mobility coefficient within that profile.

A higher mobility seems to entail a sharper fall in the index. This finding is not all that self-evident. Indeed, we observe that the minimal-expenditure index (i.e., with "absolute" mobility) ultimately falls less than the with-frictions indices! This seemingly paradoxical result actually has a simple explanation. Despite the high rate of change in product offerings and prices, a product within a given profile often remains cheaper than the others (and its price unchanged) for what can be a fairly long period. The frictionless profile index now depends only on that product and therefore disregards changes in the prices of the other products. It remains stable during the period. By contrast, the gradual transition of users of the other, non-optimal products toward the cheaper, optimal product causes a gradual decline in the with-frictions indices.

Estimating the mobility coefficient is crucial but extremely difficult. Operators reckon that about 25% of customers change products each year, which means total mobility in four years (mobility coefficient of 1/4, or 2% a month). However, they do not necessarily switch from a non-optimal product to an optimal one: it therefore seems difficult to infer a value for the mobility coefficient as formalized above. OFTEL has adopted a far higher coefficient of 10% a month, i.e., a mobility of less than a year. Relationship (*) in box 2 yields an order of magnitude—albeit very fragile—for the mobility coefficient, also of about four years. This
Chart 4: Price-change tracking is highly sensitive to consumer mobility

Frictionless indices with one-, two-, three-, four-year mobility
Laspeyres indices chained monthly

Source: INSEE

assumes that the proportion of “poorly positioned” consumers in December 2001 (month $m-1$) stays the same in January 2002 (month $m$). The index computed with the highest mobility (one year) is the one that falls most sharply over the entire period: the proportion of consumers using optimal products rises significantly over time. This does not seem realistic and is not corroborated by the data. One possible approach would consist in using the results of the next operator survey, which will provide the December 2002 user distribution by products in each profile. Such a solution would, however, be very difficult to implement. An iterative procedure would be applied to the mobility coefficients starting with the December 2001 distribution in order to arrive at the December 2002 distribution (as yet unknown), or at least to get as close to it as possible.

11. What index should be selected for the current production of the price index?

The complexity and fragility of the with-frictions model and its lack of robustness for the chosen value of the mobility coefficient argue in favor of selecting the frictionless model for the current (monthly) production of the CPI. This is in fact what has been decided. The CPI field of coverage has therefore been extended, since January 2003, to mobile-telephony services. The frictionless model is clear, relatively simple, and neutral toward changes in the rationality of behaviors. It also allows the inclusion of a mobility factor, in a radical manner since it is dichotomous: adjustments are instantaneous within each profile, whereas the shifts between profiles—such as shifts between operators or between packages and cards—are excluded. However, a chaining procedure, characteristic of the French CPI, does allow an adjustment of profile weights. The relatively detailed segmentation of consumers into profiles therefore enables us to modulate the inclusion of their mobility between products.

---

$^8$ Thanks to the survey, we know the proportion of “poorly positioned” customers in December 2001.
References


Session 2 - Use of hedonic regressions

Chair: Mick Silver, Cardiff University, UK

Summary of session

This session benefited from two detailed and well-worked papers on the use of hedonic regressions: Jan de Haan on the time dummy approach to hedonic indices and Erwin Diewert on a review of some unresolved issues in estimating hedonic regressions.

Okamoto tabled an empirical paper on weights in hedonic regressions.

Silver tabled a paper on the use of weights in hedonic regression.

The last room document from van der Grient and de Haan compared the time dummy approach and the matched model Törnqvist price index on televisions data.

It is worth distinguishing between two uses of hedonic regressions: their estimation so that imputations can be made for the prices of unmatched, non-comparable replacements and the use of hedonic indices, such as the time dummy variable method (DT) or hedonic imputed indices (HI). The DT is the focus of de Haan’s paper while Diewert considers issues arising in both application.

The time dummy variable method has much to commend it, especially insofar as its use of all the data in the two periods being sampled. This compares with hedonic imputations under the matched models method for just (non-comparable) replacement items. The paper by de Haan is to be welcomed in its development of tools for product areas where there is a rapid turnover in items. A major development of the paper is the appreciation that since unmatched and matched items may have different (quality-adjusted) prices; this can be explicitly modelled as part of the estimation process. De Haan considers systematic differences between quality-adjusted matched and unmatched prices. It is very important and analytically useful to disentangle such effects, and to bring them into the estimation procedure. This he does.

De Haan advises a double imputation price index as a weighted geometric average of the matched-product geometric mean index and the time dummy index. This method is intuitively appealing since it explicitly restricts hedonic modelling to unmatched products while leaving the price relatives of matched products unaffected. The cleverness of his double imputation approach is use actual comparisons for matched and a ‘tailored’ set of hedonic coefficients using the ‘new’ and ‘old’ coefficients for unmatched. He further incorporates a sampling framework to allow sample estimates of target indices and draws particular attention to a Törnqvist target index.

There is no doubt a contribution in this paper, especially with regard to the need for a more careful specification required for hedonic estimates for unmatched replacements and the impact of sample design on calculated indices. There remains the issue of whether imputations for the whole sample are preferable to imputations for only the unmatched sample using the double imputation method.
De Haan recognises that the time dummy in practice could approximate the hedonic quality adjustment index quite well, but we cannot be sure of this. A possible cause for bias in double imputation comes from the restrictive specification of the hedonic modelling. Indeed the specification for the unmatched assumes the slope coefficients are constant for matched and unmatched. Moreover the instrument used to correct the unmatched items is an estimator born out of the whole sample. Ideally it should be born out of just the unmatched new and just the unmatched old. This would equate with separate regressions with separate slope and intercept variables for the unmatched old and unmatched new. But degrees of freedom problems would probably preclude this. So the unmatched old and new slope coefficients are constrained to be the same and only the intercepts differ. This is a restrictive assumption and its problems must be weighed against the problems of other imputation methods involving both matched and unmatched items for the whole sample using predicted values.

This is however, a possible path to explore and in no way detracts from the contribution of the paper.

Diewert asks five questions:

1. Should separate hedonic regressions be run for each of the comparison periods or should we use the dummy variable adjacent year regression technique?

2. Should regression coefficients be sign restricted or not?

3. Should the dependent variable be transformed or not?

4. Should the hedonic regressions be weighted or unweighted? If they should be weighted, should quantity or expenditure weights be used?

5. How should outliers in the regressions be treated? Can influence analysis be used?

The formal working of these issues is an important contribution, these matters having previously been raised in either pure statistical or discursive terms.

The discussion focussed on two issues, (1) and (4).

(1) Should separate hedonic regressions be run for each of the comparison periods or should we use the dummy variable adjacent year regression technique?

The running of separate hedonic regressions for each of the comparison periods - base and current periods - to give rise to hedonic imputed indices (HI) based on such regressions is one approach.
The dummy time (DT) variable adjacent year regression technique is another.

\[
\ln p_i = \beta_0 + \beta_1 D_i + \sum_{k=2}^{K} \beta_k z_{ki} + e_i
\]

Diewert argues that a disadvantage of HI is that two estimates result. But this is simply the spread arising from the change between the base and current period characteristic being compared. We do not express concern that a Fisher index arises from two different estimates - Laspeyres and Paasche. We would prefer the spread to be minimal but when it is not be we do not abandon Fisher. The two estimates or separate bounds provide interesting information.

Diewert also argues that DT in constraining the coefficients to be the same in the periods compared has more degrees of freedom than HI. But HI are more flexible in that the implicit functional form, a geo mean of two estimates, allows variation in slope coefficients unlike DT. So it may conserve degrees of freedom, but at a cost. Diewert finally argues that DT is less subject to multicollinearity. However, the use of predicted values for hedonic indices, rather than individual coefficients, negates some of the disadvantages of multicollinearity.

The second issue discussed was weighting.

(4) Should the hedonic regressions be weighted or unweighted? If they should be weighted, should quantity or expenditure weights be used?

Diewert argued well for the use of weights, using a WLS estimator. Some weights are better than none. The question was whether quantity or expenditure based ones were best.

Okamoto in a tabled paper undertook an extensive study for TVs. He found different weights can matter. Disparities were less so if (higher priced) wide screen (WS) sets excluded.

Diewert argued, on the basis of representativity, that quantity weights gave too little weight to high priced items for the decomposition of value changes. He argued that value weights were preferred, and for dummy variable time, value shares for homoskedastic residuals.

Silver in a tabled paper showed that OLS and WLS weights need not be representative due to influence of observations. He suggested deletions of low weighted observations with undue influence. Concern was expressed over the deletion of information.

Some participants argued that while there was a case for representative weights for hedonic DT indices, they need not be required for hedonic estimates used to correct for quality non-comparable replacements. The case was made that either the sampling of items implicitly weighted the regression. And furthermore, if all observations are in hedonic competitive equilibrium the equal weighting would not be biased. This remains an empirical matter.

Other points raised by Diewert were accepted.
Recommendations for statistical agencies

1. When using imputations for non-comparable items by predicting their prices from a hedonic regression there is a case for including dummy variables in the regression on whether the item is 'unmatched new' as well as 'unmatched old'. The double imputation framework outlined in de Haan can then be used to more appropriately make such adjustments.

2. When using DT indices some weighting is better than no weighting and value share weights seem slightly preferable to quantity weights on grounds of representativity.

3. When using hedonic regressions for estimating prices for non-comparable replacements unweighted estimates may be suitable if an appropriate sampling scheme is used or the market is in competitive hedonic equilibrium.

4. It seems preferable to use the log of the price for the dependent variable rather than the model price itself.

5. Sign restricting the estimating coefficients on the basis of theory/a priori expectations is recommended as a careful practice to help provide interpretable results.

6. Some value was found in the use of influence analysis for representativity, but caution was expressed over any undue deletion of observations with high influence unless their weights were low.
Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement

Jan de Haan

Statistics Netherlands

Acknowledgements are due to Heymerik van der Grient, Peter Hein van Mulligen, May Hua Oei and Eddy Opperoes (Statistics Netherlands), Jörgen Dalén (Dalén Consultancy), Mick Silver (Cardiff University), and participants of the Ottawa Group meeting for helpful comments.

Abstract: Quality-adjusted price indexes are frequently obtained by estimating how much of the price difference between a disappearing item and its replacement is due to a quality difference. Hedonic regression has become a popular quality-adjustment method among statistical agencies. The use of the time dummy method is still very limited, though. This paper has two aims. First, it shows how this method fits into the matched-model methodology of agencies applying a geometric mean index formula at the elementary aggregation level. Second, the paper argues that the ordinary or 'direct' time dummy approach cannot cope with systematic price effects of new and disappearing products. Several 'indirect' alternatives are discussed in which the time dummy coefficient serves as a common adjustment factor and in which systematic effects of unmatched products are explicitly taken into account. Special attention is paid to the role of the sampling design, in particular to product sampling proportional to expenditure.

Keywords: consumer price index; hedonic regression; quality adjustment; sampling.

1. Introduction

The changing mix and quality of products pose difficult problems in constructing the consumer price index (CPI). A statistical agency has to make judgments about how much of the price difference between a disappearing product and its replacement is due to a quality difference. One of the methods to arrive at quality-adjusted price changes is hedonic regression. Most economists agree that this method offers the most promising approach to account for changing product quality. A panel of experts that was asked by the U.S. Bureau of Labor Statistics (BLS) “to investigate conceptual, measurement, and other statistical issues in the development of cost-of-living indexes” shared this view (Schultze and Mackie, 2001). Yet the panel recommends the BLS not to immediately expand the use of hedonics because

1 The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

2 The panel will be referred to as the Schultze panel. Silver (2002a) reviews its book entitled At What Price?. Hausman (2003) provides a critical assessment of the panel’s views and recommendations.
there are substantial unresolved econometric, data and other measurement issues that need further attention.

An approach to hedonic modelling which has mainly been applied in the academic literature is the time dummy approach. In this method, data from multiple periods are used to estimate the coefficients of a function relating the logarithm of price to a set of product characteristics and a set of dummy variables for the periods covered. The antilogarithm (exponential) of the time dummy coefficient directly produces a quality-adjusted price index. Recommendation 4-4 of the Schultz panel states that "BLS should not allocate resources to the direct time dummy method (unless work on other hedonic methods generates empirical evidence that characteristic parameter stability exists for some products)". By parameter stability the panel refers to stability over time. In the longer run this kind of stability is indeed not expected to hold. But in the short run, and especially for consecutive months, the constant-parameter assumption does not seem too restrictive. Time series are obtained by multiplying, or 'chaining' as it is usually called, the adjacent-period (bilateral) hedonic index numbers.

Section 2 of this paper reviews the adjacent-period time dummy method and shows that this method can be viewed as an automatic hedonic quality-adjustment method using a geometric mean index formula at the elementary aggregation level - a point stressed before by Tripllett (2001) but not always fully appreciated. Section 3 accounts for the possibility that there are systematic differences between the quality-adjusted prices of matched products (sold in adjacent periods) and unmatched products. This is done by incorporating dummy variables for new and disappearing products into the model. The antilog of the time dummy coefficient should not be interpreted as a quality-adjusted price index any longer, but a hedonic quality-adjustment index akin to the time dummy index can be defined. Section 4 discusses imputation methods that attempt to estimate unobservable prices of unmatched products from the extended model. Special attention is paid to the role of the sampling design. We present an hedonic imputation estimator for the superlative Törnqvist price index under product sampling strictly proportional to expenditure in each period. Section 5 considers some practical issues and suggests a simplified, more feasible Törnqvist-type estimator. Section 6 concludes.

2. The (Direct) Time Dummy Price Index

2.1 A review

This is not the place for a thorough review of the literature on hedonics; see e.g. Rosen (1974), Tripllett (1988), Griliches (1990), and Berndt et al. (1995). It suffices to mention that there has been some debate about the preferred functional form of hedonic models. Many researchers believe that the choice of the functional form is an empirical rather than a theoretical matter. Most empirical findings favour the logarithmic model over its linear counterpart. This is in agreement with Diewert's (2001, 2003) a priori point of view. Among other things, he argues that the residuals from a logarithmic model are less likely to be heteroskedastic. Time dummy approaches to hedonic regression use a logarithmic specification merely for reasons of convenience.

We will compare two time periods: the base period 0 and the current period 1. The (semi-) logarithmic model can be expressed as

\[
\ln p_i^t = \alpha^t + \sum_{k=1}^{K} \beta_{ik}^t x_{ik} + \varepsilon_i^t \quad (t=0,1),
\]  

(1)
where $p_i^t$ is the price of product $i$ in period $t$, $x_{ik}$ its $k$-th characteristic ($k = 1, \ldots, K$), $\beta_k^t$ the corresponding parameter and $\epsilon_i^t$ an error term with an expected value of zero. Our analysis does not change when some or even all characteristics are logarithmic. All parameters are time dependent as there is no reason to believe they must be constant over time, and model (1) should preferably be estimated on cross-section data for each time period separately. However, we might expect the parameters to be approximately constant in the short run. Thus if period 0 and period 1 are adjacent periods (i.e. months) it seems justifiable to impose a priori restrictions $\beta_k^0 = \beta_k^1 = \beta_k$ for all $k$. This implies that the restricted model

$$\ln p_i^t = \alpha + \delta D_i + \sum_{k=1}^{K} \beta_k x_{ik} + \epsilon_i^t \quad (t = 0, 1),$$

may be estimated on the pooled data of both periods, where the time dummy variable $D_i$ takes on the value of one if the $i$-th observation comes from period 1 and zero otherwise. We assume that the errors are independently and identically distributed with constant variances. If these (rather strong) assumptions do not hold, then Ordinary Least Squares (OLS) estimators might be very inefficient. Moreover, a complex structure of the variance-covariance matrix might make it difficult to test whether the parameters are indeed constant over time. Notice that we did not change the notation for the errors, notwithstanding that model (2) is a restricted version of model (1). This should not lead to confusion. The method is known as the adjacent-period time dummy approach. The estimated base period and current period prices of $i$ are $\hat{p}_i^0 = \exp(\hat{\alpha} + \sum_{k=1}^{K} \hat{\beta}_k x_{ik})$ and $\hat{p}_i^1 = \exp(\hat{\alpha} + \hat{\delta} + \sum_{k=1}^{K} \hat{\beta}_k x_{ik})$, respectively. So a time dummy price index

$$\hat{P}_{TD} = \hat{p}_i^1 / \hat{p}_i^0 = \exp(\hat{\delta}) \quad (\text{for all } i)$$

can be computed directly from the estimated model. This is probably why the time dummy method is sometimes called the ‘direct’ hedonic approach.

Suppose we have a product sample $S^0$ in period 0 and a sample $S^1$ in period 1. It is assumed that hedonic model (2) is estimated by OLS regression on the pooled data of $S^0 \cup S^1$. The regression residuals are defined as $u_i^t = \ln p_i^t - \ln \hat{p}_i^t = \ln(p_i^t / \hat{p}_i^t); \ t = 0, 1$. Due to the inclusion of a constant term and a time dummy, the residuals sum to zero in both periods, or:

$$\sum_{i \in S^0} \ln \left( \frac{p_i^0}{\hat{p}_i^0} \right) = \sum_{i \in S^1} \ln \left( \frac{p_i^1}{\hat{p}_i^1} \right) = 0. \quad (4)$$

Taking antilogarithms yields

$$\prod_{i \in S^0} \left( \frac{p_i^0}{\hat{p}_i^0} \right) = \prod_{i \in S^1} \left( \frac{p_i^1}{\hat{p}_i^1} \right). \quad (5)$$

If the sample does not change ($S^1 = S^0 = S$, with size $n$), it follows from (5) that

\citefoot{1968}{Goldberger}
\[
\prod_{i \in S} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right)^{\frac{1}{n}} = \hat{P}_{TD} = \prod_{i \in S} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{\frac{1}{n}}.
\] (6)

Thus, “the price index number formula implied by a dummy variable (logarithmic) regression run on matched models is a ratio of equally-weighted geometric means” (Triplet 2001). Obviously, the model specification does not matter in this specific case. The Schultz panel also mentions Triplet's work: “Triplet (2001b; 6-7) notes that the dummy variable method, when specified in a double-log or semilog functional form, produces a price index based on the geometric mean formula. Since statistical agencies have begun moving toward using the geometric mean formula to construct elementary item indexes (for other reasons), time dummy approaches have become more consistent with the prevailing methodology.” (Schultz and Mackie, 2001, p. 4-19). It is important to realise that the expected value of any elementary price index, and thus of (6), depends on the sampling design. For a discussion of elementary price indexes from a sampling perspective, see Balk (2003).

Often the product sample does change, and this is when quality adjustment comes into play. Products observed in both time periods will be referred to as matched products; \( S^1 \cap S^0 \) is the matched sample with size \( n_M \). We assume that \( S^1 \cap S^0 \neq \emptyset \). The sub-sample of products observed during the base period that has disappeared is denoted by \( S^{0D} \), and the current period sub-sample of new products is denoted by \( S^{1N} \). For reasons of simplicity and because this reflects statistical agencies’ usual practices, the sample size \( n \) will be kept constant. Hence, the size of the unmatched part of the sample is \( n - n_M \). The following relation can be derived from equation (5):

\[
\prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) = \left( \prod_{i \in S \cap S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right) \right) \left( \prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) \right).
\] (7)

Substituting \( \hat{p}_i^{1/2} / p_i^{1/2} = \exp(\delta) = \hat{P}_{TD} \) for all \( i \in S^1 \cap S^0 \) in the left-hand side of (7) and some rearranging gives

\[
\hat{P}_{TD} = \left[ P_{M} \right]_{f_M} = \left[ \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{\frac{1}{n-M}} \right] \left[ \prod_{i \in S^1 \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right)^{\frac{1}{n-M}} \right],
\] (8)

where \( P_{M} = \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{n-M} \) denotes the geometric mean or Jevons price index for the matched sample; \( f_M = n_M / n \) is the fraction of matched products.

Let us take a closer look at equation (8). The second factor between square brackets can be viewed as a quality-adjusted geometric mean price index for the unmatched part. To show this, suppose that a statistical agency selects a replacement product \( j \in S^{1N} \) for a disappearing product. The following relation can be derived from equation (5):

\[
\prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) = \left( \prod_{i \in S \cap S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right) \right) \left( \prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) \right).
\] (7)

Substituting \( \hat{p}_i^{1/2} / p_i^{1/2} = \exp(\delta) = \hat{P}_{TD} \) for all \( i \in S^1 \cap S^0 \) in the left-hand side of (7) and some rearranging gives

\[
\hat{P}_{TD} = \left[ P_{M} \right]_{f_M} = \left[ \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{\frac{1}{n-M}} \right] \left[ \prod_{i \in S^1 \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right)^{\frac{1}{n-M}} \right],
\] (8)

where \( P_{M} = \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{n-M} \) denotes the geometric mean or Jevons price index for the matched sample; \( f_M = n_M / n \) is the fraction of matched products.

Let us take a closer look at equation (8). The second factor between square brackets can be viewed as a quality-adjusted geometric mean price index for the unmatched part. To show this, suppose that a statistical agency selects a replacement product \( j \in S^{1N} \) for a disappearing product.

\[
\prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) = \left( \prod_{i \in S \cap S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right) \right) \left( \prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) \right).
\] (7)

Substituting \( \hat{p}_i^{1/2} / p_i^{1/2} = \exp(\delta) = \hat{P}_{TD} \) for all \( i \in S^1 \cap S^0 \) in the left-hand side of (7) and some rearranging gives

\[
\hat{P}_{TD} = \left[ P_{M} \right]_{f_M} = \left[ \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{\frac{1}{n-M}} \right] \left[ \prod_{i \in S^1 \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right)^{\frac{1}{n-M}} \right],
\] (8)

where \( P_{M} = \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{n-M} \) denotes the geometric mean or Jevons price index for the matched sample; \( f_M = n_M / n \) is the fraction of matched products.

Let us take a closer look at equation (8). The second factor between square brackets can be viewed as a quality-adjusted geometric mean price index for the unmatched part. To show this, suppose that a statistical agency selects a replacement product \( j \in S^{1N} \) for a disappearing product.

\[
\prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) = \left( \prod_{i \in S \cap S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right) \right) \left( \prod_{i \in S \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right) \right).
\] (7)

Substituting \( \hat{p}_i^{1/2} / p_i^{1/2} = \exp(\delta) = \hat{P}_{TD} \) for all \( i \in S^1 \cap S^0 \) in the left-hand side of (7) and some rearranging gives

\[
\hat{P}_{TD} = \left[ P_{M} \right]_{f_M} = \left[ \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{\frac{1}{n-M}} \right] \left[ \prod_{i \in S^1 \cap S^0} \left( \frac{\hat{p}_i^{1/2}}{p_i^{1/2}} \right)^{\frac{1}{n-M}} \right],
\] (8)

where \( P_{M} = \prod_{i \in S^0} \left( \frac{p_i^{1/2}}{\hat{p}_i^{1/2}} \right)^{n-M} \) denotes the geometric mean or Jevons price index for the matched sample; \( f_M = n_M / n \) is the fraction of matched products.
product \( i \in S_0 \) in order to maintain the initial sample size. The agency observes a difference in quality and adjusts the current period price of replacement \( j \) as follows:

\[
p_j^{\text{adj}} = p_j \exp \left[ \sum_{k=1}^{K} \hat{\beta}_k (x_{jk} - x_{jk}) \right] = \hat{p}_i \left[ \frac{p_j}{\hat{p}_i} / \exp(\delta) \right]. \tag{9}
\]

Comparing the second expression on the right-hand side of (9) with (8) reveals that the time dummy method automatically produces a geometric mean price index based on hedonic quality adjustment. This has been mentioned before by e.g. Triplett (2001) and Silver and Heravi (2002). Notice that \( p_j^{\text{adj}} = p_j \) if \( x_{jk} = x_{jk} \) for all \( k \). That is, when the disappearing product and its replacement have identical characteristics, no adjustment will be made. It does not matter which product from the set of new products is attached to a certain disappearing product.

### 2.2 Sampling issues

Nothing has been said so far about the sampling design, i.e. about how products are selected. Yet this is a crucial issue because it determines the population price index the time dummy index effectively aims at (the target index). Our starting point is a sample \( S_0 \) drawn proportional to base period expenditure. This design resembles the initial sampling procedures of some statistical agencies, notably that of the U.S. BLS. Suppose for the moment that there are no disappearing products, neither in the sample nor in the universe (population). The sample geometric mean price index \( \prod_{i \in S_0} (p_i / p_0)^{1/n} = \hat{P}_D \), given by equation (6), then is an approximately unbiased estimator of

\[
P_{\text{GL}} = \prod_{i \in S_0} \left( \frac{p_i}{p_0} \right)^{s_i^0}, \tag{10}
\]

where \( s_i^0 \) is the base period expenditure share of product \( i \) and \( U_0 \) the base period universe of products belonging to the product group in question. The bias is positive and will approach zero when the sample size grows. Formula (10) is sometimes called the geometric Laspeyres price index.

Silver (2002b) criticises unweighted (OLS) regression. He concludes that ".... a WLS estimator is preferred to an OLS one", and that "the use of expenditure share weights for hedonic time dummy regression estimates is preferred to relative quantities". This is certainly true for sample data obtained from simple random sampling but not for data obtained from sampling proportional to expenditure. In the latter case, an expenditure-weighted least squares estimator places too much weight on high-expenditure products. WLS may be useful for estimating hedonic imputation indexes (which are discussed in section 4) if the errors are heteroskedastic, even under product sampling proportional to expenditure.

Next, we look at the case where some products have disappeared and are replaced by other products. We might view the adjusted current period price of replacement \( j \), given by quality-adjustment procedure (9), as an estimator \( \hat{p}_i^* \) of the unobservable price \( p_i^* \) of the disappearing product \( i \). Hence, under initial sampling proportional to base period expenditure, the time dummy index might loosely be interpreted as a quality-adjusted estimator of the
geometric Laspeyres price index (10) when some products disappeared. A problem arises because the statistical agencies' replacement procedures typically do not rely on probability sampling. This not only makes variance estimation difficult, it affects the expected value of the time dummy price index in an unknown way.

3. Systematic Effects of Unmatched Products

3.1 Extending the model

Silver and Heravi (2002), Triplett (2001), and various other authors have argued that unmatched (i.e. new and disappearing) products may have unusual prices given their characteristics. Stated otherwise, the 'law of one quality-adjusted price' might not hold. We are particularly interested in those situations where the quality-adjusted prices of unmatched and matched products differ in a systematic way. For example, the observed (current period) prices of new products might be higher than we would expect from an hedonic model—consumers face hidden price increases, it is sometimes said. In this particular case new products most likely exhibit positive residuals. Quality-adjustment procedure (9), and hence the time dummy index (8), seemingly takes this possibility into account. This can easily be shown by rewriting (8) as

$$\hat{P}_{TD} = P_{M} \left[ \exp(\bar{u}_N^1 - \bar{u}_D^0) \right]^{1/\mu}, \quad (8')$$

where $$\bar{u}_N^1 = \sum_{i \in N \setminus M} u_i^1 / (n - n_M)$$ and $$\bar{u}_D^0 = \sum_{i \in D \setminus M} u_i^0 / (n - n_M)$$ are the average residuals for the new and disappearing products, respectively. Thus if $$\bar{u}_N^1 > 0$$ (and $$\bar{u}_D^0 \approx 0$$), then the time dummy index $$\hat{P}_{TD}$$ will be greater than the matched-product geometric mean index $$P_{M}$$ there.

As a matter of fact, least squares regression will usually provide biased parameter estimators if the expected value of $$e_i$$ is non-zero for some $$i$$. Or to put it another way: when the law of one quality-adjusted price systematically fails, the model has been misspecified. So if the quality-adjusted prices of the unmatched products systematically differ from those of the matched ones, then model (2) should be adapted accordingly. We will retain our basic assumption saying that the characteristics parameters remain constant and add dummy variables (allowing the intercept term to shift) for new as well as for disappearing products. The extended model for $$t = 0, 1$$ becomes:

$$\ln p_i^t = \alpha + \delta D_i + \lambda_D^0 D_i^0 + \lambda_D^1 D_i^1 + \lambda_N^0 D_i^0 + \lambda_N^1 D_i^1 + \sum_{k=1}^{K} \beta_k x_{ik} + e_i^t, \quad (11)$$

A situation of equal quality-adjusted prices for all products, apart from a random error term, most likely holds in competitive markets. It is therefore not very surprising that empirical studies using the adjacent-period time dummy approach often find (chained) indexes quite similar to matched-product geometric mean indexes in fairly competitive markets. For such a study on Dutch CPI data for computers, see Van Mulligen (2002).
where the dummy $D_{it}^i$ takes on the value of one if product $i$ disappeared in period $t+1$ and zero otherwise, and where the dummy $D_{it}^N$ takes on the value of one if $i$ is new in period $t$ and zero otherwise. It is assumed that products can be observed in at least two consecutive periods. Products that are new in period 0, for instance, are assumed not to disappear in period 1. Scanner data indeed suggest that products are rarely sold during one month only, so this assumption is not really a restrictive one.

This paper looks at bilateral price change; prices in the current period (1) are compared with prices in the preceding period (0). The specification of (11) may seem a bit curious at first sight because (in some sense) it looks beyond those periods. The model not only takes into account systematic effects of the unmatched products - that is, products that are either new in period 1 or disappear after period 0 - but also systematic effects of matched products that are new in period 0 and those that disappear in period 2. This has been done for reasons of symmetry. For example, if in period 1 the quality-adjusted prices of new products systematically exceed those of the other products, there is reason to expect that something similar exists for new products in period 0. The extended model should account for such effects as well. Note that during period 1 no one knows what products will disappear in period 2. Equation (11) therefore cannot be estimated in real time, but here we act as if we did have this knowledge. We will return to this issue in section 5.

The OLS parameter estimators of (11) will be indicated by a tilde. Compared with the 'true' model (11), the simple model (2) suffers from omitted variables, yielding biased OLS estimators and biased predicted values. It is important to recognise that equation (11) models the whole population. This means, for example, that $D_{it}^N$ must equal 1 only if product $i$ has not been sold in period 0. Because statistical agencies lack this kind of information, $D_{it}^N$ will be set to 1 in practice if $i$ is new in the sample. So there is a sample-selection problem, and $\tilde{\lambda}_N^0$, $\tilde{\lambda}_N^1$, $\tilde{\lambda}_D^0$ and $\tilde{\lambda}_D^1$ may suffer from sample-selection or 'outside-the-sample' bias.

3.2 A hedonic quality adjustment price index

As we have seen, the time dummy index cannot cope with systematic effects of new and disappearing products. It is possible, however, to incorporate these effects into an index based on hedonic quality adjustment. We apply a quality adjustment procedure similar to (9) but estimated from the extended model (11): $p_{ij}^{adj} = p_j^i \exp[\sum_{k=1}^{K} \tilde{\beta}_k (x_{ik} - x_{jk})]$, where $j \in S_1^{1N}$ replaces $i \in S_0^{0D}$. Instead of the first expression on the right of (9), we now have

\[
p_{ij}^{adj} = p_j^i \left[ \frac{\exp(\tilde{\delta} + \tilde{\lambda}_N^0 - \tilde{\lambda}_D^0)}{(\tilde{p}_j^i / \tilde{p}_i^0)} \right].
\]

(12)

Using procedure (12) in a geometric mean framework leads to the following explicit hedonic quality adjustment price index:

\[\text{Using OLS regression, the residuals for the matched products sum to zero in both periods. It can easily be demonstrated that the antilogarithm of the estimated time dummy coefficient would be exactly equal to the matched-product geometric mean index if the latter effects had not been incorporated into the model.}\]
This approach can be called an 'indirect' time dummy method, where the time dummy coefficient serves as a common adjustment factor. As a result of adding dummies for new and disappearing products, the OLS residuals for \( i \in S^{0D} \) and \( i \in S^{11} \) satisfy

\[
\sum_{i \in S^{0D}} \ln \left( \frac{P_i^0}{P_i} \right) = \sum_{i \in S^{11}} \ln \left( \frac{P_i}{P_i} \right) = 0.
\] (14)

From taking antilogs of (14) it follows that the second factor on the right-hand side of (13) is equal to 1, yielding

\[
\bar{P}_{HQ1} = [P_{M}]^{fM} \left[ \exp(\tilde{\delta} + \tilde{\lambda}_N - \tilde{\lambda}_D) \right]^{-fM}.
\] (15)

Let us compare (15) with the following alternative expression for the time dummy price index \((8')\):

\[
\hat{P}_{TD} = [P_{M}]^{fM} \left[ \exp(\tilde{\delta} + \tilde{\mu}_N - \tilde{\mu}_D) \right]^{-fM},
\] (8’)

where \( \tilde{\mu}_N \) and \( \tilde{\mu}_D \) are the average residuals for the new and disappearing products, based on the ordinary model (2). There is a striking similarity between expressions \((8')\) and (15). The values of \( \tilde{\delta} \) and \( \tilde{\delta}' \), as well as the values of \( \tilde{\mu}_N - \tilde{\mu}_D \) and \( \tilde{\lambda}_N - \tilde{\lambda}_D \), might differ only slightly.

In practice therefore the time dummy index could approximate the hedonic quality adjustment index quite well, but we cannot be sure of this.

One difficulty remains. If initial product sampling had been performed proportional to base period expenditure, so that the geometric Laspeyres index is the target, it seems a bit odd to take systematic effects of new products into account as these products do not belong to the base period universe (but to the current period universe instead). In our opinion hedonic imputation price indexes make much more sense in this respect.

4. Hedonic Imputation Price Indexes

4.1 Sampling proportional to base period expenditure

Our starting point is again a sample \( S^0 \), which has been drawn proportional to base period expenditure. First we will assume that systematic price effects of unmatched products do not occur. We would like to impute the unobservable (fictitious) current period prices of disappearing products. In section 2.2 we noted that the time dummy price index can be interpreted in this manner. The term hedonic imputation is generally used for those methods that estimate unobservable prices directly from the hedonic model itself, though, and those methods will be investigated here. The imputation procedure should measure what the current period prices would have been — given the bundle of characteristics incorporated in the
products — had the products still been sold, that is, had they been matched products instead. Using hedonic model (2), those prices are \( \hat{p}_i^w = \exp(\alpha + \bar{\delta} + \sum_{k=1}^{K} \bar{\beta}_k x_{ik}) \) for \( i \in S^{OD} \).

A disadvantage might be substantial variability of the ‘quality-adjusted’ price index for the unmatched products due to a comparison of estimated (current period) prices with actual (base period) prices. There are several ways to deal with this problem. Silver (2002b) proposes to delete so-called influential outliers. However, unless their quality is in serious doubt, throwing away observations is not a very attractive idea. We prefer to enhance stability by replacing the observed base period prices by their predicted values based on model (2). We then obtain \( \hat{p}_i^w / \hat{p}_i^0 = \exp(\tilde{\delta}) \) for \( i \in S^{OD} \). The geometric mean formula yields a \textit{double imputation price index}:

\[
\hat{P}_{DI} = \left[ \bar{P}_{MI} \right]_{\bar{f}_M} \left[ \exp(\tilde{\delta}) \right]^{-\bar{f}_M} = \left[ \bar{P}_{MI} \right]_{\bar{f}_M} \left[ \hat{P}_{TE} \right]^{-\bar{f}_M},
\]

(16)

which is also the result of an ‘indirect’ time dummy approach.

The double imputation price index is a weighted geometric average of the matched-product geometric mean index and the time dummy index. This method is intuitively appealing since it restricts hedonic modelling explicitly — rather than implicitly, as in the time dummy index (8) — to unmatched products, while leaving the price relatives of matched products unaffected. Notice that under the present sampling design \( \bar{P}_{MI} \) is an approximately unbiased estimator of the matched-product geometric Laspeyres price index; the expected value of \( f_M \) is approximately equal to the base period expenditure share of the matched products.

Next we assume that systematic price effects of unmatched products do occur. The unobservable current period prices of disappearing products now have to be imputed according to hedonic model (11). Those prices are \( \hat{p}_i^w = \exp(\bar{\alpha} + \bar{\delta} + \sum_{k=1}^{K} \bar{\beta}_k x_{ik}) \), since we want to measure what they would have been had the disappearing products been matched products instead. Using the geometric mean framework again, the imputation index reads

\[
\tilde{P}_{GL} = \left[ \bar{P}_{MI} \right]_{\bar{f}_M} \left[ \prod_{i \in S^{OD}} \left( \exp(\bar{\alpha} + \bar{\delta} + \sum_{k=1}^{K} \bar{\beta}_k x_{ik}) \right)^{1-\bar{f}_M} \right]^{-\bar{f}_M} \left( \bar{p}_i^0 \right)^{1/(n-n_M)}.
\]

(17)

Equation (14) implies \( \prod_{i \in S^{OD}} (\bar{p}_i^0)^{1/(n-n_M)} = \prod_{i \in S^{OD}} (\tilde{p}_i^0(\tilde{\alpha} + \tilde{\delta}))^{1/(n-n_M)} \). So \( \tilde{P}_{GL} \) automatically becomes a \textit{double imputation index}. Using \( \tilde{p}_i^0 = \exp(\bar{\alpha} + \bar{\delta} + \sum_{k=1}^{K} \bar{\beta}_k x_{ik}) \) for \( i \in S^{OD} \), (17) reduces to

\[
\tilde{P}_{GL} = \left[ \bar{P}_{MI} \right]_{\bar{f}_M} \left[ \exp(\tilde{\delta} - \tilde{\alpha}) \right]^{-\bar{f}_M}.
\]

(18)

When \( \lambda_D^0 \) approaches zero we expect to find \( \tilde{\delta} \approx \hat{\delta} \) and thus \( \tilde{P}_{GL} \approx \hat{P}_{DI} \), particularly for not too small samples. This result strengthens our choice made above for double instead of single
(one-sided) imputation. $\tilde{P}_{GL}$ can be viewed as an estimator of the geometric Laspeyres price index (10) under sampling proportional to base period expenditure when some products are no longer available. In (10) we define the unobservable current period prices for $i \in U^{0D}$ as $p_i = p_i^*$, where $U^{0D}$ denotes the disappearing part of $U^0$. The procedure to select replacement products is less important, provided that model (11) holds.

4.2 Estimating the Törnqvist price index

Suppose next that our starting point was a sample $S'$ drawn proportional to current period expenditure. Products that have not been sold in the base period are 'replaced backwards' by products that did sell. Again, the actual 'replacement procedure' is of minor importance. We now wish to compute an imputation price index by imputing the unobservable base period prices for new products. The resulting imputation price index can be expressed as

$$\tilde{P}_{GP} = \left[ P_S \right]^{1/\hat{s}_i} \left[ \exp(\hat{o} + \hat{\lambda}_{N}^1) \right]^{1-\hat{f}_w}. \quad (19)$$

$\tilde{P}_{GP}$ is an estimator of the 'geometric Paasche price index'

$$P_{GP} = \prod_{i \in U'} \left( \frac{p_i^1}{p_i^0} \right)^{s_i}, \quad (20)$$

where $s_i$ is the current period expenditure share of product $i$ and $U'$ the current period product universe. In (20) we define the unobservable base period prices for $i \in U^{1N}$ as $p_i^0 = p_i^{*0}$, where $U^{1N}$ denotes the new part of $U^1$. It follows that the geometric average of (18) and (19), i.e.

$$\tilde{P}_T = \left[ P_S \right]^{1/\hat{s}_i} \left[ \exp(\hat{o} + \hat{\lambda}_{N}^1 - \hat{\lambda}_{N}^0) \right]^{1-\hat{f}_w}, \quad (21)$$

is an approximately unbiased estimator of the Törnqvist price index

$$P_T = \left[ P_{GL} P_{GP} \right]^{1/2} = \prod_{i \in U^0} \left( \frac{p_i^1}{p_i^0} \right)^{s_i^0} \prod_{i \in U^1} \left( \frac{p_i^1}{p_i^0} \right)^{s_i^1}. \quad (22)$$

under independent sampling proportional to expenditure in both periods.

The Törnqvist index is defined on a variable set of products or a 'dynamic universe' (Dalén, 2001). This index can be seen as a generalisation of the ordinary, static universe Törnqvist index, which is adjusted for quality changes in the sense that all unobservable prices are imputed. De Haan (2002) describes a similar generalisation of the Fisher price index. Estimator (21) makes no explicit use of $\hat{\lambda}_D$ and $\hat{\lambda}_N$. One should not conclude that the corresponding dummy variables $D_{1D}$ and $D_{1N}$ can be dropped from model (11), since $\delta$ would then no longer be estimated unbiasedly; see also footnote 4. Comparing equations (16) with

6 De Haan and Oppendoes (2002) applied hedonic imputation methods in a Fisher index framework, and found that single imputation sometimes led to implausible results.
and (21) makes clear that the hedonic quality adjustment index (16) is an upward (downward) biased estimator of the Törnqvist index if $\lambda_N^1 > \lambda_D^0$ ($\lambda_N^1 < \lambda_D^0$).

The interpretation of the 'backward imputation' approach used in (19) and (21) can be problematic because production of the new products may not have been possible with the technology of the earlier period. Triplett (2001) remarks: "The price index quality adjustment problem sometimes makes it necessary to estimate a price in period $t$ for a new computer that first became available in period $t+1$. One cannot ignore the fact that the new computer was not actually available in period $t$, and the possible reasons why it was not available. The special assumptions necessary to validate "backcasting" a price for a machine that was not in fact produced should be kept in mind." But we cannot escape from 'backcasting' when defining and estimating the Törnqvist price index (22), or any other type of symmetric index for that matter, in a dynamic-universe context.

It should be noted that proportional to expenditure sampling independently in both time periods might be an inefficient sampling strategy, resulting in an unnecessarily small matched sample. Some $i \in S^{1N}$, while not belonging to the matched sample, could actually belong to the matched universe $U^1 \cap U^0$. Despite the fact that their base period and current period prices can be observed, at least in principle, estimator (21) uses the predicted values.

So far it has been assumed that a difference in quality can only exist when products have different sets of characteristics, and systematic effects of new and disappearing have been treated as real price changes. For durable goods, such as televisions and computers, this assumption is a reasonable one. For certain other product groups (for example novels), on the other hand, consumers may appreciate new products over old ones with otherwise the same characteristics. Model (11) would still hold, but in this case 'newness' and 'oldness' become quality aspects themselves. One probably wants to control for those aspects as well and incorporate the dummy variables for new and disappearing products into the quality-adjustment procedure. It is straightforward to show that all imputation indexes discussed above now equal $\exp(\delta)$. Unlike in the case of the time dummy index, the expected values of $\exp(\delta)$ and the matched-product geometric mean index differ, precisely because model (11) holds instead of model (2). We will not pursue this special case any further here, and next turn to some practical considerations when systematic effects of unmatched products are deemed real.

5. Practical Considerations

The analysis in section 4 suggests we should make a choice between the imputation indexes (18), (19) and (21). As noted before, the sampling design is the crucial issue at stake. Let us start with (18), which can be viewed as an estimator of the geometric Laspeyres price index under sampling proportional to base period expenditure.

Under systematic sampling proportional to size some high-expenditure products may have a probability of 1 to be in the sample. The products belonging to this so-called self-selecting part must be weighted according to expenditure in order to obtain approximately unbiased estimators. See for example De Haan et al. (1999). We will not address this complication here.
practical relevance whatsoever. This estimator has merely been shown in order to arrive at (21).

Under sampling proportional to expenditure in both periods (21) is to be preferred from a theoretical perspective because it estimates the superlative (adjacent-period) Törnqvist price index. In its pure form this sampling design would be infeasible for statistical agencies. They might nevertheless try to mimic it. At present, many agencies select a replacement product similar to the 'old' one. While this method is understandable from a pragmatic side because it seemingly reduces the need for quality adjustment, it should not be recommended. To keep the sample up-to-date, a better procedure would be to select products that (are expected to) have a significant market share. A more offensive strategy would be to replace a part of the sample each month instead of waiting until forced replacements occur. Doing this such that high-expenditure products will have a higher inclusion probability than low-expenditure products mimics product sampling proportional to expenditure to some extent, so that estimator (21) gets a meaningful interpretation. The number of unmatched products decreases, of course, which makes hedonic regression particularly helpful.

Still the fraction of unmatched products could be small. In fact an attrition rate of 0.1 would be unusually large, not only for CPI samples but also for the universe – except perhaps in case of high-tech goods like computers. With a small unmatched sample, the standard errors of $\lambda_p$ and $\lambda_N^p$ might be very large. The variance of the price index will then be substantial. Using Taylor linearization a formula for the approximated variance of (21), conditional on the design matrix of $x$-values (and thus on $f_N$), is easily derived from the variance-covariance matrix supplied by most statistical packages. One must realise that those packages typically assume simple random sampling, which gives rise to bias in the (co)variance estimators. It would be much more difficult to estimate the unconditional variance.

Things get even more complicated because in practice one does not yet know what products will disappear in period 2. This makes real-time estimation of model (11) – particularly of $\lambda_p$ – impossible. However, using the sample-replacement rule suggested above there seems to be no need to include dummies for disappearing products since products will have left the sample well before they actually disappear from the market. This rule might nevertheless create some 'outside-the-sample bias'. An additional (and, admittedly, rather strong) assumption we shall make is $\lambda_p^1 = \lambda_N^p = \lambda_N$. We therefore propose the following simplified version of (11):

$$\ln p_i = \alpha + \delta D_i + \lambda_N D_{in} + \sum_{k=1}^{K} \beta_k x_{ik} + e_i^{r} \quad (r=0,1),$$

(23)

where the dummy $D_{in}$ takes on the value of one if $i$ is new in either period 0 or 1 and zero otherwise. Denoting the parameter estimators of the dummy variables in (23) by $\delta$ and $\lambda_N^p$, estimator (21) now reduces to

---

8 For research purposes Statistics Netherlands has bought scannerdata from GfK for some durable goods. A first inspection of the data for televisions and washing machines revealed that most products did not show substantial price changes just before they left the market. This supports the case for not including dummy variables for disappearing products. Silver and Heravi (2002), on the other hand, using similar data on washing machines, did find evidence of lower quality-adjusted prices for disappearing products.
The best way to proceed after having estimated model (23) seems to test whether \( \lambda_{N} \) differs from zero. A cautious approach is called for. Testing at low significance levels almost inevitably leads to rejecting the null hypothesis \( H_{0} : \lambda_{N} = 0 \), and the usual 5% level might be too low. If the null hypothesis cannot be rejected, model (2) should be estimated in a second stage and the (double) imputation index (16) should be computed.

The question arises whether estimators like (24) or (16) can successfully be applied in practice using price data that are collected for the compilation of the CPI. Suppose the statistical agency needs at least \( n \) price observations to estimate a geometric mean price index with sufficient accuracy. A necessary requirement for using model (23) on the pooled data of adjacent periods is \( n \geq (K + 3)/2 \). This requirement will usually be met. What is needed as well is sufficient variation in characteristics and prices. This requires the use of wide product descriptions for selecting specific items in the outlets, contrary to Statistics Netherlands' practices. For example, Statistics Netherlands collects about 30 prices monthly for each of 9 specific television sets \( (n = 270) \). Although the overall sample size is not a bottle-neck for implementing time dummy methods, current data gathering procedures should be redesigned towards using wide product descriptions and, above all, collecting data on the characteristics of all sampled products. Statistical agencies may find the latter requirement especially burdensome. It can be argued, on the other hand, that every sophisticated quality-adjustment method needs data on product characteristics, so they should have been collected anyway. Such data are becoming increasingly available via the internet. Other data sources might be the manufacturers, the importers, or even the retailers where prices are collected.

Expression (24) suggests we do not have to bother so much about the stability of the characteristics’ coefficients. Multicollinearity by itself raises no problems. Although the variance of the hedonic index will be higher than it would have been otherwise, the choice of the set of characteristics included in the model should not depend on statistical but on economic reasoning. Those characteristics must be included that are related to the product’s performance. According to Tripplett (2001) it requires “careful thought, sometimes subtle analysis, and most importantly, knowing one’s product”. Including irrelevant variables does not affect the unbiasedness of the OLS estimators for the relevant variables. So increasing \( R^{2} \) by including non-performance variables does not affect the unbiasedness of \( \hat{\delta} \) and \( \hat{\lambda}_{N} \), and hence does not introduce quality-adjustment bias. But there is a problem of inefficiency. Including irrelevant variables generally raises the standard errors of the coefficients on the relevant ones. In particular, if in model (23) the dummy variable \( D_{NW} \) is highly correlated with the unjustly included variables, \( \delta \) and \( \lambda_{N} \) would be estimated unnecessarily inefficient.

6. Conclusions

It is well established in the index number literature that the direct time dummy approach produces a geometric mean price index. The prices of newly introduced products are synthetically matched to the prices of replaced products via hedonic quality adjustment. One of our conclusions states that when there are no systematic price effects of new and disappearing products, the time dummy index should not deviate in a systematic way from the matched-product geometric mean index – provided that the characteristics parameters are
constant (which is our basic assumption). The time dummy approach can still be justified on the grounds that it might lower the standard error of the index by increasing the number of observations, albeit at the expense of possible bias caused by an imperfectly specified hedonic model.

The impact of modelling can be reduced by using an hedonic imputation approach. This approach estimates unobservable prices of unmatched product directly from the hedonic model itself, and does not try to make a synthetic match (quality adjustment) between a disappearing product and a replacement. Hedonic imputation methods allow us to take the sampling design explicitly into account and to estimate target (population) indexes like the geometric Laspeyres index. Moreover, hedonic imputation indexes can account for systematic effects of unmatched products, whereas the time dummy index cannot. To this end, the hedonic model should be extended with dummy variables for new and disappearing products. An imputation index has been proposed, which approximates the population Törnqvist index under sampling proportional to expenditure in both periods. Using a sample replacement rule whereby new products enter the sample timely and replace obsolete products well before they actually disappear from the market has an obvious advantage compared with traditional approaches in which the sample is kept fixed until forced replacements occur. Nevertheless, some ‘outside-the-sample’ bias can arise if products had unusual prices just before they disappear.

The analysis outlined in this paper depends on the appropriateness of the logarithmic specification of the hedonic model and on the assumption of constant parameters. The constant-parameter assumption has been criticised frequently, recently by the Schultze panel. For adjacent periods, though, this assumption does not seem too unrealistic. In any case it is much less restrictive than using coefficients relating to months or even years ago, which is what statistical agencies often do when applying hedonic regression. The advantage of pooling data from adjacent months is a substantial gain in efficiency, and this property could make time dummy methods feasible on actual CPI data without a need to enlarge the sample.

The Schultze panel raises another point of concern worth mentioning. “Finally, most uses of hedonics . . . predict the price that would have prevailed in period t for a variety or model not actually offered for sale in that period. While this seems sensible, it is problematic at the theoretical level: under imperfect competition, if an additional variety or model had actually been offered for sale, the prices of the other products might also have changed” (Schultze and Mackie, 2001, p. 4-40). The panel is right, obviously. Yet the criticism is somewhat misdirected because it pertains to quality adjustment methods in general, not just to hedonics. Predicting unobservable prices of unmatched products (while leaving the prices of matched products unaffected) is essentially what quality adjustment is about, either with or without using hedonics. Taking the panel’s opinion literally would make it virtually impossible to compile quality-adjusted price indexes.
References


Hausman, J. (2003), Sources of Bias and Solutions to Bias in the CPI, *Journal of Economic Perspectives*, 17, no. 1, 23-44.


Hedonic Regressions: A Review of Some Unresolved Issues

Erwin Diewert

University of British Columbia, Vancouver, Canada

The author is indebted to Ernst Berndt and Alice Nakamura for helpful comments.

1. Introduction

Three recent publications have revived interest in the topic of hedonic regressions. The first publication is Pakes (2001) who proposed a somewhat controversial view of the topic. The second publication is Chapter 4 in Schultze and Mackie (2002), where a rather cautious approach to the use of hedonic regressions was advocated due to the fact that many issues had not yet been completely resolved. A third paper by Heravi and Silver (2002) also raised questions about the usefulness of hedonic regressions since this paper presented several alternative hedonic regression methodologies and obtained different empirical results using the alternative models.

Some of the more important issues that need to be resolved before hedonic regressions can be routinely applied by statistical agencies include:

- Should the dependent variable be transformed or not?
- Should separate hedonic regressions be run for each of the comparison periods or should we use the dummy variable adjacent year regression technique initially suggested by Court (1939; 109-11) and used by Berndt, Griliches and Rappaport (1995; 260) and many others?
- Should regression coefficients be sign restricted or not?
- Should the hedonic regressions be weighted or unweighted? If they should be weighted, should quantity or expenditure weights be used?
- How should outliers in the regressions be treated? Can influence analysis be used?

The present paper takes a systematic look at the above questions. Single period hedonic regression issues are addressed in sections 2 to 5 while two year time dummy variable regression issues are addressed in sections 6 and 7. Some of the more technical material relating to section 7 is in an Appendix, which examines the properties of bilateral weighted

1 See Hulten (2002) for a nice review of the issues raised in Pakes paper.
2 The observation that different variants of hedonic regression techniques can generate quite different answers empirically dates back to Triplett and McDonald (1977; 150) at least.
3 Diewert (2002b) recently looked at these weighting issues in the context of a simplified adjacent year hedonic regression model where the only characteristics were dummy variables.
hedonic regressions. Section 8 discusses the treatment of outliers and influential observations and section 9 addresses the issue of whether the signs of hedonic regression coefficients should be restricted. Section 10 concludes.

2. To Log or Not to Log

We suppose that price data have been collected on K models or varieties of a commodity over T+1 periods. Thus \( p_k^t \) is the price of model k in period t for \( t = 0,1,...,T \) and \( k \in S(t) \) where \( S(t) \) is the set of models that are actually sold in period t. For \( k \in S(t) \), denote the number of these type k models sold during period t by \( q_k^t \). We suppose also that information is available on N relevant characteristics of each model. The amount of characteristic n that model k possesses in period t is denoted as \( z_{kn}^t \) for \( t = 0,1,...,T \), \( n = 1,...,N \) and \( k \in S(t) \). Define the N dimensional vector of characteristics for model k in period t as \( z_{k}^t = [z_{k1}^t,z_{k2}^t,...,z_{kN}^t] \) for \( t = 0,1,...,T \) and \( k \in S(t) \). We shall consider only linear hedonic regressions in this review. Hence, the unweighted linear hedonic regression for period t has the following form:

\[
(1) f(p_k^t) = \beta_0 + \sum_{n=1}^{N} f_n(z_{kn}^t)\beta_n^t + e_k^t, \quad t = 0,1,...,T; \quad k \in S(t)
\]

where \( e_k^t \) is an independently distributed error term with mean 0 and variance \( \sigma^2 \). \( f(x) \) is either the identity function \( f(x) = x \) or the natural logarithm function \( f(x) = \ln x \) and the functions of one variable \( f_n \) are either the identity function, the logarithm function or a dummy variable which takes on the value 1 if the characteristic n is present in model k or 0 otherwise. We are restricting the \( f \) and \( f_n \) in this way since the identity, log and dummy variable functions are by far the most commonly used transformation functions used in hedonic regressions.

Recall that the period t characteristics vector for model k was defined as \( z_k^t = [z_{k1}^t,z_{k2}^t,...,z_{kN}^t] \). We define also the period t vector of the \( \beta \)'s as \( \beta^t = [\beta_0^t,\beta_1^t,...,\beta_N^t] \). Using these definitions, we simplify the notation on the right hand side of (1) by defining:

\[
(2) h(t(z_k^t,\beta^t)) = \beta_0^t + \sum_{n=1}^{N} f_n(z_{kn}^t)\beta_n^t \quad t = 0,1,...,T; \quad k \in S(t).
\]

The question we now want to address is: should the dependent variable \( f(p_k^t) \) on the left hand side of (1) be \( p_k^t \) or \( \ln p_k^t \); i.e., should \( f \) be the identity function or the log function? We also would like to know if the choice of identity or log for the function \( f \) should affect our choice of identity or log for the \( f_n \) that correspond to the continuous (i.e., non dummy variable) characteristics.

Suppose that we choose \( f \) to be the identity function. Suppose further that there is only one continuous characteristic so that \( N = 1 \). In this situation, the hedonic regression is essentially a regression of price on package size and so if we want to have as a special case, that price per

---

4 Models sold in different outlets can be regarded as separate varieties or not, depending on the context.
5 If a particular model k is sold at various prices during period t, then we interpret \( q_k^t \) as the total quantity of model k that is sold in period t and \( p_k^t \) as the corresponding average price or unit value.
6 Note that the linear regression model defined by (1) can only provide a first order approximation to a general hedonic function. Diewet (2001) made a case for considering second order approximations but in this paper, we will follow current practice and consider only linear approximations.
7 Griliches (1971a; 58) noted that an advantage of the log formulation is that \( \beta_n^t \) would provide an estimate of the percentage change in price due to a one unit change in \( z_{kn}^t \) provided that \( f_n \) was the identity function. Court (1939; 111) implicitly noted this advantage of the log formulation.
unit of useful characteristic is a constant, then we should set $f_i(z_i) = z_i$. Under these conditions, the model defined by (1) and $f(p) = p$ will be consistent with the constant per unit price hypothesis if $\beta_0^t = 0$. In the case of $N$ continuous characteristics, a generalization of the constant per unit characteristic price hypothesis is the hypothesis of constant returns to scale in the vector of characteristics, so that if all characteristics are doubled, then the resulting model price is doubled. If our period $t$ model is defined by (1) and $f(p) = p$, then $h^t$ must satisfy the following property:

\begin{equation}
\beta_0^t + \sum_{n=1}^N f_n(\lambda z_{kn})\beta_n^t = \lambda[\beta_0^t + \sum_{n=1}^N f_n(z_{kn})\beta_n^t]
\end{equation}

for all $\lambda > 0$.

In order to satisfy (3), we must choose $\beta_0^t = 0$ and the $f_n$ to be identity functions. Thus if $f$ is chosen to be the identity function, then it is natural to choose the $f_n$ that correspond to continuous characteristics to be identity functions as well.

Now suppose that we choose $f$ to be the log function. Suppose again that there is only one continuous characteristic so that $N = 1$. In this situation, again the hedonic regression is essentially a regression of price on package size and so if we want to have as a special case, that price per unit of useful characteristic is a constant, then we need to set $f_1(z_1) = \ln z_1$ and $\beta_1^t = 1$. Under these conditions, the model defined by (1) and $f(p) = \ln p$ will be consistent with the constant per unit price hypothesis. In the case of $N$ continuous characteristics, a generalization of the constant per unit price hypothesis is the hypothesis of constant returns to scale in the vector of characteristics. If our period $t$ model is defined by (1) and $f(p) = \ln p$, then $h^t$ must satisfy the following property:

\begin{equation}
\beta_0^t + \sum_{n=1}^N f_n(\lambda z_{kn})\beta_n^t = \ln \lambda + \beta_0^t + \sum_{n=1}^N f_n(z_{kn})\beta_n^t
\end{equation}

for all $\lambda > 0$.

In order to satisfy (4), we must choose the $f_n(z_n)$ to be log functions and the $\beta_n^t$ must satisfy the following linear restriction:

\begin{equation}
\sum_{n=1}^N \beta_n^t = 1.
\end{equation}

Thus if $f$ is chosen to be the log function, then it is natural to choose the $f_n$ that correspond to continuous characteristics to be log functions as well.

An extremely important property that a hedonic regression model should possess is that the model be invariant to changes in the units of measurement of the continuous characteristics. Thus suppose that we have only continuous characteristics and the period $t$ model is defined by (1) with $f$ arbitrary and the $f_n(z_n) = \ln z_n$. Suppose further that new units of measurement for the $N$ characteristics are chosen, say $Z_n$, where

\begin{equation}
Z_n = z_n/c_n,
\end{equation}

\begin{equation}
n = 1, \ldots, N
\end{equation}

We are not arguing that this constant returns to scale hypothesis must necessarily hold (usually, it will not hold); we are just arguing that it is useful for the hedonic regression model to be able to model this situation as a special case. The constant returns to scale hypothesis is required in some hedonic models; e.g., see Muellbauer (1974; 988) and Pollak's (1983) “L Characteristics” model, which is also used by Triplett (1983).

If we change the units of measurement for the continuous characteristics, then the linear hedonic regression model will be unaffected by this change in the units; i.e., the change in the units for the $n$th characteristic can be absorbed into the regression coefficient $\beta_n$.

Note that all of the continuous characteristics must be measured in positive units in this case.
where the \( c_i \) are positive constants. The invariance property requires that we can find new regression coefficients, \( \beta_i^* \), such that the following equation can be satisfied identically:

\[
\beta_0^* + \sum_{n=1}^{N} (\ln z_n) \beta_n^* = \beta_0^* + \sum_{n=1}^{N} (\ln z_n) \beta_n^*
\]

\[
= \beta_0^* - \sum_{n=1}^{N} (\ln c_n) \beta_n^* + \sum_{n=1}^{N} (\ln z_n) \beta_n^*.
\]

Hence to satisfy (7) identically, we need only set \( \beta_n^* = \beta_n^* \) for \( n = 1, \ldots, N \) and set \( \beta_0^* = \beta_0^* - \sum_{n=1}^{N} (\ln c_n) \beta_n^* \). Thus in particular, the hedonic regression model where \( f \) and the \( f_n \) are all log functions will satisfy the important invariance to changes in the units of measurement of the continuous characteristics property, provided that the regression has a constant term in it.  

We now address the following question: should the dependent variable \( f(p^k_t) \) on the left hand side of (1) be \( p^k_t \) or \( \ln p^k_t \)?

If \( f \) is the identity function, then using definitions (2), equations (1) can be rewritten as follows:

\[
(8) \quad p^k_t = h(z^k_t, p^0_t) + \varepsilon^k_t;
\]

where \( \varepsilon^k_t \) is an independently distributed error term with mean 0 and variance \( \sigma^2 \). On the other hand, if \( f \) is the logarithm function, then equations (1) are equivalent to the following equations:

\[
(9) \quad p^k_t = \exp[h(z^k_t, p^0_t)] \exp[\varepsilon^k_t];
\]

where \( \eta^k_t \) is an independently distributed error term with mean 1 and constant variance. Which is more plausible: the model specified by (8) or the model specified by (9)? We argue that it is more likely that the errors in (9) are homoskedastic compared to the errors in (8) since models with very large characteristic vectors \( z^k_t \) will have high prices \( p^k_t \) and are very likely to have relatively large error terms. On the other hand, models with very small amounts of characteristics will have small prices and small means and the deviation of a model price from its mean will be necessarily small. In other words, it is more plausible to assume that the ratio of model price to its mean price is randomly distributed with mean 1 and constant variance than to assume that the difference between model price and its mean is randomly distributed with mean 0 and constant variance. Hence, from an a priori point of view, we would favor the logarithmic regression model (9) (or (1) with \( f(p) = \ln p \)) over its linear counterpart (8).

The regression models considered in this section were unweighted models and could be estimated without a knowledge of the amounts sold for each model in each period. In the following section, we assume that model quantity information \( q^k_t \) is available and we consider how this extra information could be used.

\[11\] Note that the above argument is independent of the functional form for \( f \); i.e., if the \( f_n \) for the continuous characteristics are log functions, then for any \( f \), the hedonic regression must include a constant term to be invariant to changes in the units of these continuous characteristics.
3. Quantity Weights versus Expenditure Weights

Usually, discussions of how to use quantity or expenditure weights in a hedonic regression are centered around discussions on how to reduce the heteroskedasticity of error terms. In this section, we attempt a somewhat different approach based on the idea that the regression model should be representative. In other words, if model k sold $q_k t$ times in period t, then perhaps model k should be repeated in the period t hedonic regression $q_k t$ times so that the period t regression is representative of the sales that actually occurred during the period.\(^\text{12}\)

To illustrate this idea, suppose that in period t, only three models were sold and there is only one continuous characteristic. Let the period t price of the three models be $p_1^t$, $p_2^t$ and $p_3^t$ and suppose that the three models have the amounts $z_{1i}^t$, $z_{2i}^t$ and $z_{3i}^t$ of the single characteristic respectively. Then the period t unweighted regression model (1) has only the following 3 observations and 2 unknown parameters, $\beta_0^t$ and $\beta_1^t$:

\[
\begin{align*}
(f(p_1^t)) &= \beta_0^t + f_1(z_{1i}^t)\beta_1^t + \epsilon_1^t; \\
(f(p_2^t)) &= \beta_0^t + f_1(z_{2i}^t)\beta_1^t + \epsilon_2^t; \\
(f(p_3^t)) &= \beta_0^t + f_1(z_{3i}^t)\beta_1^t + \epsilon_3^t.
\end{align*}
\]

Note that each of the 3 observations gets an equal weight in the period t hedonic regression model defined by (10). However, if say models 1 and 2 are vastly more popular than model 3, then it does not seem to be appropriate that model 3 gets the same importance as models 1 and 2.

Suppose that the integers $q_1^t$, $q_2^t$ and $q_3^t$ are the amounts sold in period t of models 1, 2 and 3 respectively. Then one way of constructing a hedonic regression that weights models according to their economic importance is to repeat each model observation according to the number of times it sold in the period. This leads to the following more representative hedonic regression model, where the error terms have been omitted:

\[
\begin{align*}
1_1f(p_1^t) &= 1_1(\beta_0^t + 1_1f_1(z_{1i}^t)\beta_1^t); \\
1_2f(p_2^t) &= 1_2(\beta_0^t + 1_2f_1(z_{2i}^t)\beta_1^t); \\
1_3f(p_3^t) &= 1_3(\beta_0^t + 1_3f_1(z_{3i}^t)\beta_1^t);
\end{align*}
\]

where $1_k$ is a vector of ones of dimension $q_k^t$ for $k = 1, 2, 3$.

Now consider the following quantity transformation of the original unweighted hedonic regression model (10):

\[
\begin{align*}
(q_1^t)^{1/2} f(p_1^t) &= (q_1^t)^{1/2} \beta_0^t + (q_1^t)^{1/2} f_1(z_{1i}^t)\beta_1^t + \epsilon_1^t; \\
(q_2^t)^{1/2} f(p_2^t) &= (q_2^t)^{1/2} \beta_0^t + (q_2^t)^{1/2} f_1(z_{2i}^t)\beta_1^t + \epsilon_2^t; \\
(q_3^t)^{1/2} f(p_3^t) &= (q_3^t)^{1/2} \beta_0^t + (q_3^t)^{1/2} f_1(z_{3i}^t)\beta_1^t + \epsilon_3^t.
\end{align*}
\]

\(^{12}\) Thus our representative approach follows along the lines of Theil's (1967; 136-138) stochastic approach to index number theory, which is also pursued by Rao (2002). The use of weights that reflect the economic importance of models was recommended by Griliches (1971b; 8): "But even here, we should use a weighted regression approach, since we are interested in an estimate of a weighted average of the pure price change, rather than just an unweighted average over all possible models, no matter how peculiar or rare." However, he did not make any explicit weighting suggestions.
Comparing (10) and (12), it can be seen that the observations in (12) are equal to the corresponding observations in (10), except that the dependent and independent variables in observation of (10) have been multiplied by the square root of the quantity sold of model k in period t for k = 1,2,3 in order to obtain the observations in (12). A sampling framework for (12) is available if we assume that the transformed residuals $\varepsilon_k^{*}$ are independently normally distributed with mean zero and constant variance.

Let $b_0^1$ and $b_1^1$ denote the least squares estimators for the parameters $\beta_0^1$ and $\beta_1^1$ in (11) and let $b_0^{*1}$ and $b_1^{*1}$ denote the least squares estimators for the parameters $\beta_0^1$ and $\beta_1^1$ in (12). Then it is straightforward to show that these two sets of least squares estimators are the same $^{13}$; i.e., we have:

\[(13) \ [b_0^{1}, b_1^{1}] = [b_0^{*1}, b_1^{*1}].\]

Thus a shortcut method for obtaining the least squares estimators for the unknown parameters, $\beta_0^1$ and $\beta_1^1$, which occur in the "representative" model (11) is to obtain the least squares estimators for the transformed model (12). This equivalence between the two models provides a justification for using the weighted model (12) in place of the original model (10). The advantage in using the transformed model (12) over the "representative" model (11) is that we can develop a sampling framework for (12) but not for (11), since the (omitted) error terms in (11) cannot be assumed to be distributed independently of each other. However, in view of the equivalence between the least squares estimators for models (11) and (12), we can now be comfortable that the regression model (12) weights observations according to their quantitative importance in period t. Hence, we definitely recommend the use of the weighted hedonic regression model (12) over its unweighted counterpart (10).

However, rather than weighting models by their quantity sold in each period, it is possible to weight each model according to the value of its sales in each period. Thus define the value of sales of model k in period t to be:

\[(14) v_k^t = p_k^t q_k^t; \quad t = 0,1,\ldots,T; \quad k \in S(t).\]

Now consider again the simple unweighted hedonic regression model defined by (10) above and round off the sales of each of the 3 models to the nearest dollar (or penny). Let $1_k$ be a vector of ones of dimension $v_k^t$ for k = 1,2,3. Repeating each model in (10) according to the value of its sales in period t leads to the following more representative period t hedonic regression model (where the errors have been omitted):

\[(15) \begin{align*}
1_1 & f(p_1^t) = 1_1 \beta_0^1 + 1_1 f_1(z_{11}) \beta_1^1; \\
1_2 & f(p_2^t) = 1_2 \beta_0^1 + 1_2 f_1(z_{21}) \beta_1^1; \\
1_3 & f(p_3^t) = 1_3 \beta_0^1 + 1_3 f_1(z_{31}) \beta_1^1.
\end{align*}\]

$^{13}$ See, for example, Greene (1993; 277-279). However, the numerical equivalence of the least squares estimates obtained by repeating multiple observations or by the square root of the weight transformation was noticed long ago as the following quotation indicates: "It is evident that an observation of weight w enters into the equations exactly as if it were w separate observations each of weight unity. The best practical method of accounting for the weight is, however, to prepare the equations of condition by multiplying each equation throughout by the square root of its weight." E. T. Whittaker and G. Robinson (1940; 224).
Now consider the following value transformation of the original unweighted hedonic regression model (10):

\[
(16) \quad (v_{1t})^{1/2} f(p_{1t}) = (v_{1t})^{1/2} \beta_0^t + (v_{1t})^{1/2} f(z_{1t})\beta_1^t + \varepsilon_1^{**};
\]

\[
(16) \quad (v_{2t})^{1/2} f(p_{2t}) = (v_{2t})^{1/2} \beta_0^t + (v_{2t})^{1/2} f(z_{2t})\beta_1^t + \varepsilon_2^{**};
\]

\[
(16) \quad (v_{3t})^{1/2} f(p_{3t}) = (v_{3t})^{1/2} \beta_0^t + (v_{3t})^{1/2} f(z_{3t})\beta_1^t + \varepsilon_3^{**}.
\]

Comparing (10) and (16), it can be seen that the observations in (12) are equal to the corresponding observations in (10), except that the dependent and independent variables in observation \( k \) of (10) have been multiplied by the square root of the value sold of model \( k \) in period \( t \) for \( k = 1, 2, 3 \) in order to obtain the observations in (16). Again, a sampling framework for (16) is available if we assume that the transformed residuals \( \varepsilon_k^{**} \) are independently distributed normal random variables with mean zero and constant variance.

Again, it is straightforward to show that the least squares estimators for the parameters \( \beta_0^t \) and \( \beta_1^t \) in (15) and (16) are the same. Thus a shortcut method for obtaining the least squares estimators for the unknown parameters, \( \beta_0^t \) and \( \beta_1^t \), which occur in the value weights representative model (15) is to obtain the least squares estimators for the transformed model (16). This equivalence between the two models provides a justification for using the value weighted model (16) in place of the original model (10). As before, the advantage in using the transformed model (16) over the value weights representative model (15) is that we can develop a sampling framework for (16) but not for (15), since the (omitted) error terms in (15) cannot be assumed to be distributed independently of each other.

It seems to us that the quantity weighted and value weighted models are clear improvements over the original unweighted model (10). Our reasoning here is similar to that used by Fisher (1922; Chapter III) in developing bilateral index number theory, who argued that prices needed to be weighted according to their quantitative or value importance in the two periods being compared.\(^{14}\) In the present context, we have a weighting problem that involves only one period so that our weighting problems are actually much simpler than those considered by Fisher: we need only choose between quantity or value weights!

But which system of weighting is better in our present context: quantity or value weighting?

The problem with quantity weighting is this: it will tend to give too little weight to models that have high prices and too much weight to cheap models that have low amounts of useful characteristics. Hence it appears to us that value weighting is clearly preferable. Thus we are taking the point of view that the main purpose of the period \( t \) hedonic regression is to enable

\[^{14}\] "It has already been observed that the purpose of any index number is to strike a ‘fair average’ of the price movements—or movements of other groups of magnitudes. At first a simple average seemed fair, just because it treated all terms alike. And, in the absence of any knowledge of the relative importance of the various commodities included in the average, the simple average is fair. But it was early recognized that there are enormous differences in importance. Everyone knows that pork is more important than coffee and wheat than quinine. Thus the quest for fairness led to the introduction of weighting.” Irving Fisher (1922; 43). "But on what principle shall we weight the terms? Arthur Young’s guess and other guesses at weighting represent, consciously or un consciously, the idea that relative money values of the various commodities should determine their weights. A value is, of course, the product of a price per unit, multiplied by the number of units taken. Such values afford the only common measure for comparing the streams of commodities produced, exchanged, or consumed, and afford almost the only basis of weighting which has ever been seriously proposed.” Irving Fisher (1922; 45).
us to decompose the market value of each model sold, \( p_k t q_k t \), into the product of a period \( t \) price for a quality adjusted unit of the hedonic commodity, say \( P_t \), times a constant utility total quantity for model \( k \), \( Q_k t \). Hence observation \( k \) in period \( t \) should have the representative weight \( Q_k t \) in constant utility units that are comparable across models. But \( Q_k t \) is equal to \( p_k t q_k t / P_t \), which in turn is equal to \( v_k t / P_t \), which in turn is proportional to \( v_k t \). Thus weighting by the values \( v_k t \) seems to be the most appropriate form of weighting.

Our conclusions about single period hedonic regressions at this point can be summarized as follows:

- With respect to taking transformations of the dependent variable in a period \( t \) hedonic regression, taking of logarithms of the model prices is our preferred transformation.

- If information on the number of models sold in each period is available, then weighting each observation by the square root of the value of model sales is our preferred method of weighting.

- If the log transformation is chosen for the dependent variable, then we have a mild preference for transforming the continuous characteristics by the logarithm transformation as well. If the continuous characteristics are transformed by the logarithmic transformation, then the regression must have a constant term to ensure that the results of the regression are invariant to the choice of units for the characteristics.

- If the dependent variable is simply the model price, then we have a mild preference for not transforming the continuous characteristics as well.

With the above general considerations in mind, we now turn to a discussion of how single period hedonic regressions can be used by statistical agencies in a sampling context.

4. The Use of Single Period Hedonic Regressions in a Replacement Sampling Context

In this section, we consider the use of single period hedonic regressions in the context of statistical agency sampling procedures where a sampled model that was available in period \( s \) is not available in a later period \( t \) and is replaced with a new model that is available in period \( t \).

We assume that \( s < t \) and that model 1 is available in period \( s \) (with price \( p_1 s \) and characteristics vector \( z_1 s \)) but is not available in period \( t \). We further assume that model 1 is replaced by model 2 in period \( t \), with price \( p_2 t \) and characteristics vector \( z_2 t \). The problem is to somehow adjust the price relative \( p_2 t / p_1 s \) so that the adjusted price relative can be averaged with other price relatives of the form \( p_k t / p_1 s \) that correspond to models \( k \) that are present in both periods \( s \) and \( t \) in order to form an overall price relative for the item level, going from period \( s \) to \( t \). If the item level index is a chain type index, then \( s \) will be equal to \( t-1 \) and if the item level index is a fixed base type index, then \( s \) will be equal to the base period 0.

Recall the family of single period hedonic regressions defined in section 2 above by equations (1). If we use definitions (2) and assume that the function of one variable \( f(x) \) has an inverse function \( f^{-1} \), then we may rewrite equations (1) as follows:

\[
(17) \quad p_k t = f^{-1}[h(z_k t, \beta_t) + \varepsilon_k t]; \quad t = 0,1,...,T; \ k \in S(t).
\]
Assume that we have a vector of estimates \( b^t \) for the period \( t \) vector of parameters \( \beta^t \) and define the model \( k \) sample residuals for period \( t \), \( e_k^t \), as follows:\(^{15}\)

\[
e_k^t = f(p_k^t) - h(z_k^t, \beta^t); \quad t = 0, 1, ..., T; \ k \in S(t).
\]

Thus the sample counterparts to equations (17) are the following equations:

\[
p_k^t = f^{-1}[h(z_k^t, b^t) + e_k^t]; \quad t = 0, 1, ..., T; \ k \in S(t).
\]

Now suppose that the period \( s \) hedonic regression is available to the statistical agency. Thus equation (19) for period \( s \) and model 1 is:

\[
p_1^s = f^{-1}[h(z_1^s, b^s) + e_1^s].
\]

Recall that model 2, the replacement for model 1 in period \( t \), has the vector of characteristics \( z_2^t \). Hence, using the period \( s \) hedonic regression, a comparable price for model 2 in period \( s \) is \( f^{-1}[h(z_2^s, b^s)] \), the predicted period \( s \) price using the period \( t \) hedonic regression for a model with the vector of characteristics \( z_2^t \). Thus our first estimator for an adjusted price relative for models 1 and 2 going from period \( s \) to \( t \) is:

\[
r(1) = \frac{p_2^t}{p_1^s} f^{-1}[h(z_2^s, b^s)].
\]

However, there is a problem with the use of (21) as an adjusted price relative. The problem will become apparent if \( z_2^t = z_1^t \), so that the two models are in fact identical. In this case, we want our price relative to equal the actual price ratio:

\[
p_2^t / p_1^s = \frac{p_2^t / f^{-1}[h(z_1^s, b^s) + e_1^s]}{p_1^s / f^{-1}[h(z_1^s, b^s)]} \quad \text{using (20)}
\]

Hence if the regression residual for model 1 in period \( s \), \( e_1^s \), is not equal to zero, then \( r(1) \) defined by (21) will not be an appropriate adjusted price relative. In order to compare like with like, we must multiply \( r(1) \) by an adjustment factor equal to

\[
f^{-1}[h(z_1^s, b^s)] / p_1^s = f^{-1}[h(z_1^s, b^s)] / f^{-1}[h(z_1^s, b^s) + e_1^s].
\]

Thus our second estimator \( r(2) \) for an adjusted price relative is \( r(1) \) defined by (21) times the adjustment factor defined by (23), which adjusts the period \( s \) observed price for model 1, \( p_1^s \), onto the period \( s \) hedonic regression surface:\(^{16}\)

\[
r(2) = \left( \frac{p_2^t / f^{-1}[h(z_2^s, b^s)]}{p_1^s / f^{-1}[h(z_1^s, b^s)]} \right) \left( \frac{f^{-1}[h(z_1^s, b^s)] / p_1^s}{f^{-1}[h(z_1^s, b^s)] / p_1^s} \right).
\]

The second expression for \( r(2) \) in (24) is instructive. We can interpret \( p_2^t / f^{-1}[h(z_2^s, b^s)] \) as the period \( t \) price for model 2 expressed in constant quality utility units, using the period \( s \) hedonic regression as the quality adjustment mechanism. Similarly, we can interpret \( p_1^s / f^{-1}[h(z_1^s, b^s)] \) as the period \( s \) price for model 1 expressed in constant quality utility units,

\(^{15}\)Definitions (18) need to be modified if weighted regressions are run instead of unweighted regressions.

\(^{16}\)If \( e_1^s = 0 \), then \( r(1) \) will equal \( r(2) \).
using the period s hedonic regression as the quality adjuster. Thus the price relative defined by (24) compares the price of model 2 in period t to the price of model 1 in period s in constant utility quantity units. Hence, the period s hedonic regression may be used to express model prices in homogeneous quality adjusted units.\(^\text{17}\)

Obviously, if the statistical agency has the period t hedonic regression available to it, then the above analysis can be repeated, with some modifications. In this case, equation (19) for period t and model 2 is:

\[
(25) p_2^t = f^{-1}[h'(z_2^t, b^t) + e_2^t].
\]

Recall that model 1 has the vector of characteristics \(z_1^s\). Hence, using the period t hedonic regression, a comparable price for model 1 in period t is \(f^{-1}[h'(z_1^s, b^t)]\), the predicted period t price using the period t hedonic regression for a model with the vector of characteristics \(z_1^s\). Thus our third estimator for an adjusted price relative for models 1 and 2 going from period s to t is:

\[
(26) r(3) = f^{-1}[ht(z_2^t,b^t)]/p_1^s.
\]

However, again, there is a problem with the use of (26) as an adjusted price relative. As above, the problem becomes apparent if \(z_2^t = z_1^s\), so that the two models are in fact identical. In this case, we want our price relative to equal the actual price ratio:

\[
(27) p_2^t/p_1^s = f^{-1}[h'(z_2^t, b^t) + e_2^t]/p_1^s \quad \text{using (25)}
\]

\[
\text{if } e_2^t \neq 0.
\]

Hence if the regression residual for model 2 in period t, \(e_2^t\), is not equal to zero, then \(r(3)\) defined by (26) will not be an appropriate adjusted price relative. In order to compare like with like, we must multiply \(r(3)\) by an adjustment factor equal to

\[
(28) p_2^t/f^{-1}[h'(z_2^t, b^t)] = p_2^t/f^{-1}[h'(z_2^t, b^t) + e_2^t]/f^{-1}[h'(z_2^t, b^t)].
\]

Thus our fourth estimator \(r(4)\) for an adjusted price relative is \(r(3)\) defined by (26) times the adjustment factor defined by (28), which adjusts the period t observed price for model 2, \(p_2^t\), onto the period t hedonic regression surface:\(^\text{18}\)

\[
(29) r(4) = \left\{p_2^t/f^{-1}[h'(z_2^t, b^t)]\right\}/\left\{p_1^s/f^{-1}[h'(z_1^s, b^t)]\right\}
\]

The second expression for \(r(4)\) in (29) is again instructive. We can interpret \(p_2^t/f^{-1}[h'(z_2^t, b^t)]\) as the period t price for model 2 expressed in constant quality utility units, using the period t hedonic regression as the quality adjustment mechanism. Similarly, we can interpret \(p_1^s/f^{-1}[h'(z_1^s, b^t)]\) as the period s price for model 1 expressed in constant quality utility units, using the period t hedonic regression as the quality adjuster. Thus the price relative defined by (29) compares the price of model 2 in period t to the price of model 1 in period s in

\(^{17}\) This basic idea can be traced back to Court (1939; 108) as his hedonic suggestion number one. The idea was explicitly laid out in Griliches (1971a; 59-60) (1971b; 6) and Dhrymes (1971; 111-112). It was implemented in a statistical agency sampling context by Triplett and McDonald (1977; 144).

\(^{18}\) Of course, if \(e_2^t = 0\), then \(r(3)\) will equal \(r(4)\).
constant utility quantity units, using the period t hedonic regression to do the quality adjustment.

If the period s and t hedonic regressions are both available to the statistical agency, then it is best to make use of both of the adjusted price relatives r(2) and r(4) and generate a final adjusted price relative that is a symmetric average of the two estimates. Thus define our final preferred adjusted price relative r(5) as the geometric mean of r(2) and r(4):

\[
(30) \quad r(5) = \left[ r(2)r(4) \right]^{1/2}.
\]

We chose the geometric mean in (30) over other simple symmetric means like the arithmetic average because the use of the geometric average leads to an adjusted price relative that will satisfy the time reversal test.

Finally, suppose that period s and t hedonic regressions are not available to the statistical agency but a base period hedonic regression is available. In this case, the obvious adjusted replacement price ratio is:

\[
(31) \quad r(6) = \frac{p_2^t/f_1^t[h^0(z_2^t,b_0)]}{p_1^s/f_1^s[h^0(z_1^s,b_0)]}.
\]

Thus the price relative defined by (31) compares the price of model 2 in period t to the price of model 1 in period s in constant utility quantity units, using the period 0 hedonic regression to do the quality adjustment.

Obviously, the adjusted price relative r(5) would generally be preferable to the price relative defined by r(6), since the period 0 hedonic regression may be quite out of date if period 0 is distant from periods s and t. Similar considerations suggest that more reliable results will be obtained if the chain principle is used in forming the adjusted price relatives defined by (5); i.e., the gap between the equally valid r(2) and r(4) is likely to be minimized if period s is chosen to be period t-1.

In the following section, we shall assume that the statistical agency has estimated single period hedonic regressions as in this section but in addition, we assume that information on quantities sold of each model is available. Hence, Paasche, Laspeyres and superlative indexes of the type advocated by Silver and Heravi (2001) (2002a) (2002b) and Pakes (2001) can be calculated.

---

19 Griliches (1971a; 59) noted the existence of these two equally valid estimates. Griliches (1971b; 7) also suggested taking an average of the two estimates and, as an alternative method of averaging or smoothing, he suggested using adjacent year regressions, which will be studied in sections 7 and 8 below.

20 See Diewert (1997; 138) for an argument along these lines.

21 Tastes will probably change over time and the characteristics domain of definition for models that exist in period 0 may be quite different from the domains of definition for the models that exist in periods s and t; i.e., the z region spanned by the period 0 hedonic regression may be quite out of date for the later periods.

22 Our advocacy of the chain principle and of averaging equally valid results seems to be consistent with the position advocated by Griliches (1971b; 6-7): "This approach calls for relatively recent and often changing 'price' weights. Since such statistics come to us in discrete intervals, we are also faced with the usual Laspeyres-Paasche problem. The often we can change such weights [i.e., run a new hedonic regression], the less of a problem it will be. In practice, while one may want to use the most recent cross section to derive the relevant price weights, such estimates may fluctuate too much for comfort as the result of multicollinearity and sampling fluctuations. They should be smoothed in some way, either by choosing \( w_i = (1/2)[w_i(t) + w_i(t+1)] \), or by using 'adjacent year' regressions in estimating these weights."
5. Single Period Hedonic Regressions in the Scanner Data Context

In this section, we assume that the statistical agency has both price and quantity (or value) data for the subset of the K models that are available in each period. As in the previous period, we will assume that the statistical agency has run single period hedonic regressions for periods s and t.23

The hedonic regression of period s can be used in order to calculate the following Paasche type index going from period s to t:24

\[
P_P(s,t) = \frac{\sum_{k \in S(t)} p_k^t q_k^t}{\sum_{k \in S(t) \cap S(s)} p_k^s q_k^t + \sum_{k \in S(t) \setminus S(s)} f^{-1}[h^s(z_k,b^s)]q_k^t}.
\]

The summation in the numerator of the right hand side of (32) is simply the sum of price \( p_k^t \) times quantity \( q_k^t \) over all of the models k sold during period t, which is the set of indexes k represented by \( S(t) \). The first summation in the denominator of the right hand side of (32) is the product of the period s model k price, \( p_k^s \), over all models that are present in both periods s and t while the second set of terms uses the period s estimated hedonic price of a model k that is sold in period t (which has characteristics defined by the vector \( z_k^t \)) but is not sold in period s, \( f^{-1}[h^s(z_k^t,b^s)] \), times the period t quantity sold for this model, \( q_k^t \). If we make the strong assumptions on demander’s period s preferences25 that are listed in Diewert (2001), then we can interpret \( f^{-1}[h^s(z_k^t,b^s)] \) as an approximate Hicksian (1940; 114) reservation price for model k that is sold in period t but not in period s; i.e., if price is above this limiting price, then purchasers will not want to buy any units of it in period s. Thus under appropriate assumptions on consumer’s preferences, the Paasche index defined by (32) will be an approximate lower bound to a theoretical Paasche-Konüs cost of living index; see Diewert (1993; 80).26 Thus the estimated period s hedonic regression enables us to calculate a matched model type Paasche index between periods s and t, where the prices for the models that were sold in period t but not period s are filled in using the period s hedonic regression.

In a similar manner, we can use the hedonic regression for period t to fill in the missing reservation prices for models that were sold in period s but not t and we can calculate the following Laspeyres type index going from period s to t:27

\[
P_L(s,t) = \frac{\sum_{k \in S(s) \cap S(t)} p_k^s q_k^t + \sum_{k \in S(s) \setminus S(t)} f^{-1}[h^s(z_k^s,b^s)]q_k^t}{\sum_{k \in S(s)} p_k^s q_k^s}.
\]

23 With the availability of quantity information on the models sold, value weighted hedonic regressions of the type recommended in section 4 can be run for each period.

24 This is Pakes’ (2001; 22) Paasche complete hedonic hybrid price index. Except for error terms, it is also equal to one of Silver and Heravi’s (2001) Paasche type lower bounding indexes for a true cost of living index.

25 A stronger but simpler set of assumptions than those of Diewert (2001) are that all period s demanders of the hedonic commodity evaluate the utility of a model with characteristics vector \( z \) according to the magnitude \( g^s(z) \), where \( g^s(z) \) is a separable (cardinal) utility function. Under these assumptions, the equilibrium price of a model with characteristics vector \( z \) should have the period s hedonic price function equal to \( g^s(z) \) times a constant. If \( f^{-1}[h^s(z^t,b^t)] \) can approximate this true period s hedonic price function and if the fit of the period s hedonic regression is good so that \( b^t \) is close to \( \beta^s \), then \( f^{-1}[h^s(z^t,b^t)] \) will be an approximate Hicksian reservation price for model k that is sold in period t but not in period s.

26 See Diewert (1993; 103-104) for an exposition of the use of Hicksian reservation prices for new and disappearing commodities in the context of Paasche and Laspeyres indexes.

27 Except for error terms, it is equal to one of Silver and Heravi’s (2001) Laspeyres type upper bounding indexes for a true cost of living index.
The summation in the denominator of the right hand side of (33) is simply the sum of price $p_k^s$ times quantity $q_k^s$ over all of the models $k$ sold during period $s$, which is the set of indexes $k$ represented by $S(s)$. The first summation in the numerator of the right hand side of (33) is the product of the period $t$ model $k$ price, $p_k^t$, over all models that are present in both periods $s$ and $t$ while the second set of terms uses the period $t$ estimated hedonic price of a model $k$ that is sold in period $s$ (which has characteristics defined by the vector $z_k^s$) but is not sold in period $t$, $f^{-1}[h'(z_k^s,b^t)]$, times the period $s$ quantity sold for this model, $q_k^s$. Under appropriate assumptions on consumer's preferences, the Laspeyres index defined by (33) will be an approximate upper bound to a theoretical Laspeyres-Konüs cost of living index; see Diewert (1993; 80). Thus the estimated period $t$ hedonic regression enables us to calculate a matched model type Laspeyres index between periods $s$ and $t$, where the prices for the models that were sold in period $s$ but not period $t$ are filled in using the period $t$ hedonic regression.

If both period $s$ and $t$ hedonic regressions are available to the statistical agency, then since the Paasche and Laspeyres measures of price change between periods $s$ and $t$ are equally valid, it is appropriate to take a symmetric average of these two estimators of price change as a "final" estimator of price change between the periods. As usual, we chose the geometric mean of $P_L$ and $P_P$ over other simple symmetric means like the arithmetic average because the use of the geometric average leads to an index that will satisfy the time reversal test. Hence, define the *Fisher* (1922) index between periods $s$ and $t$ as:

$$P_F(s,t) = \left[ P_L(s,t) \cdot P_P(s,t) \right]^{1/2}$$

where $P_P$ and $P_L$ are defined by (32) and (33). It is of some interest to compute $P_F$, $P_L$ and $P_F$ defined by (32)-(34) above for the case where there are only two models: model 1, which is available in period $s$ but not period $t$, and model 2, which is available in period $t$ but not period $s$; i.e., we are revisiting the sampling model that was studied in section 4 above. Under these conditions, $P_P$ defined by (32) simplifies to the following expression:

$$P_P(s,t) = p_2^s q_2^t f^{-1}[h'(z_2^s,b^t)] q_2^t = p_2^s f^{-1}[h'(z_2^s,b^t)] = r(1)$$

where $r(1)$ was defined in section 4 by (21). Similarly, $P_L$ defined by (33) simplifies to the following expression:

$$P_L(s,t) = f^{-1}[h'(z_1^s,b^t)] q_1^s/p_1^s q_1^s = f^{-1}[h'(z_1^s,b^t)]/p_1^s = r(3)$$

If all models are present in both periods, then the Laspeyres type index defined by (33) reduces to an ordinary Laspeyres index between periods $s$ and $t$ and the Paasche type index defined by (32) reduces to an ordinary Paasche index. It can be seen that the weights for each of these indexes is not representative of both periods and hence each of the indexes (32) and (33) will be subject to substitution or representativity bias; see Diewert (2002a; 45) on the concept of representativity bias. Hence, to eliminate this bias, it is necessary to take an average of the two indexes defined by (32) and (33).

See Diewert (1997; 138).

An argument due originally to Konüs (1924) can be used to prove that a theoretical cost of living index lies between the Paasche and Laspeyres indexes; see also Diewert (1993; 81). However, this argument will only go through for the case where all of the characteristics are of the continuous type.
where \( r(3) \) was defined in section 4 by (26). Recall that our preferred replacement price ratios obtained in section 4 were \( r(2) \) and \( r(4) \) rather than \( r(1) \) and \( r(3) \). Hence the results obtained in this section seem to be slightly inconsistent with the results obtained in section 4.\(^{31}\)

This slight inconsistency can be resolved if we make strong assumptions about the preferences of purchasers of the hedonic commodities. Suppose all purchasers of the hedonic commodity evaluate the relative utility of each model in period \( s \) according to the cardinal utility function \( g^s(z) \) so that the relative value to purchasers of a model with characteristics vector \( z_1 \) versus a model with characteristics vector \( z_2 \) is \( g^s(z_1)/g^s(z_2) \). Then in equilibrium, the period \( s \) relative price of the two models should also be \( g^s(z_1)/g^s(z_2) \). Thus the period \( s \) price of a model with characteristics vector \( z \) should be proportional to \( g^s(z) \). Finally, suppose that the period \( s \) econometrically estimated hedonic function, \( f^s_1[h^s(z,b)] \), can provide an adequate approximation to the theoretical hedonic function, \( p^s g^s(z) \), where \( p^s \) is a positive constant. Under these strong assumptions, the total market utility for period \( s \) that is provided by purchases of the hedonic commodities is equal to:

\[
Q^s = \sum_{k \in S(s)} p^s g^s(z_k) q_k^s = \sum_{k \in S(s)} f^s_1[h^s(z_k,b)] q_k^s
\]

where we have approximated the utility to purchasers of model \( k \) in period \( s \), \( p^s g^s(z_k) \), by the period \( s \) hedonic regression estimated value, \( f^s_1[h^s(z_k,b)] \). Thus \( Q^s \) can be interpreted as the aggregate quantity of all of the models purchased in period \( s \), where each model has been quality adjusted into constant utility units using the period \( s \) hedonic aggregator function, \( g^s(z) \). In what follows, we will neglect the approximation error between lines 1 and 2 of (37) so that we identify the period \( s \) aggregate quantity purchased of the hedonic commodity, \( Q^s(s) \), using the period \( s \) hedonic regression to do the quality adjustment, as follows:

\[
Q^s(s) = \sum_{k \in S(s)} f^s_1[h^s(z_k,b)] q_k^s.
\]

For each period \( t \), we can define the value of all models purchased as:

\[
V^s = \sum_{k \in S(t)} p_k^t q_k^t ; \quad t = 0,1,\ldots,T.
\]

For later reference, we also define the period \( t \) expenditure share of model \( k \) as follows:

\[
s_k^t = p_k^t q_k^t / \sum_{i \in S(t)} p_i^t q_i^t ; \quad t = 0,1,\ldots,T ; k \in S(t).
\]

Corresponding to the period \( s \) quantity aggregate defined by (38), we can define an aggregate period \( s \) price level, \( P^s(s) \), by dividing \( Q^s(s) \) into the period \( s \) value aggregate, \( V^s \):

\[
P^s(s) = V^s/Q^s(s) = V^s \sum_{k \in S(s)} f^s_1[h^s(z_k,b)] q_k^s
\]

using (38)

\[
= \frac{1}{\sum_{k \in S(s)} \{f^s_1[h^s(z_k,b)]/p_k^s\} p_k^s q_k^s V^s}
\]

using (40) for \( t = s \)

\[
= \frac{[\sum_{k \in S(s)} s_k^s \{p_k^s f^s_1[h^s(z_k,b)]\}]^{-1}}{V^s}.
\]

\(^{31}\) We say slightly inconsistent because usually the hedonic regression observed errors \( e_1^s \) and \( e_2^s \) will be small and hence the differences between \( r(1) \) and \( r(2) \) and \( r(3) \) and \( r(4) \) will also be small.
Thus the aggregate period s price level using the period s hedonic regression, \( P^s(s) \), is equal to a period s share weighted harmonic mean of the period s actual model prices, \( p^s_k \), relative to the corresponding predicted period s model prices using the period s hedonic regression, \( f^{-1}[h^s(z^s_k,b^s)] \). Since \( p^s_k = f^{-1}[h^s(z^s_k,b^s)] + e^s_k \) where \( e^s_k \) is the regression residual for model k in period s, and these residuals are typically close to 0 and randomly distributed around 0, it can be seen that under normal conditions, \( P^s(s) \) defined by (41) will be close to 1.

Now let us use the period s hedonic regression to form a constant utility quantity aggregate for the models sold in period t. Thus model k in period t, using the estimated hedonic valuation function of period s, will have the constant utility value \( f^{-1}[h^s(z^s_k,b^s)] \). Hence, the period t aggregate quantity purchased of the hedonic commodity, \( Q^t(s) \), using the period s hedonic regression to do the quality adjustment into constant utility units, can be defined as follows:

\[
(42) \quad Q^t(s) = \sum_{k \in S(t)} f^{-1}[h^s(z^s_k,b^s)] q^t_k.
\]

Corresponding to the period t quantity aggregate defined by (42), we can define an aggregate period t price level using the preferences of period s to do the quality adjustment, \( P^t(s) \), by dividing \( Q^t(s) \) into the period t value aggregate, \( V^t \):

\[
(43) \quad P^t(s) = \frac{V^t}{Q^t(s)} = \frac{V^t}{\sum_{k \in S(t)} f^{-1}[h^s(z^s_k,b^s)] q^t_k} = \frac{1}{\sum_{k \in S(t)} \{ f^{-1}[h^s(z^s_k,b^s)]/p^t_k \} p^t_k q^t_k / V^t}.
\]

Thus the aggregate period t price level using the period s hedonic regression, \( P^t(s) \), is equal to a period t share weighted harmonic mean of the period t actual model prices, \( p^t_k \), relative to the corresponding predicted period s model prices using the period s hedonic regression, \( f^{-1}[h^s(z^s_k,b^s)] \).

Having defined the period s price level \( P^s(s) \) by (41) and the period t price level \( P^t(s) \) by (43) using the hedonic regression of period s to do the constant utility quality adjustment, we can take the ratio of these two price levels to form a Paasche type price index going from period s to t, using the hedonic regression of period s, as follows:

\[
(44) \quad P^{st}(s) = P^t(s) / P^s(s) = \frac{\left[ \sum_{k \in S(t)} s^t_k \{ p^t_k / f^{-1}[h^s(z^s_k,b^s)] \}^{-1} \right]^{-1} / \left[ \sum_{k \in S(s)} s^s_k \{ p^s_k / f^{-1}[h^s(z^s_k,b^s)] \}^{-1} \right]^{-1}}.
\]

The above Paasche type index can be compared with our earlier Paasche type index defined by (32):

\[32\] It can be seen that the expression on the right hand side of (41) is a type of Paasche price index, where the price and quantity data of period s, \( p^s_k \) and \( q^s_k \) for \( k \in S(s) \), act as the comparison period data and the hedonic regression period s predicted prices, \( f^{-1}[h^s(z^s_k,b^s)] \) for \( k \in S(s) \), act as base period prices.

\[33\] Our algebra here assumes that unweighted hedonic regressions have been run. If a value weighted hedonic regression has been run for period s, then the equation \( p^s_k = f^{-1}[h^s(z^s_k,b^s)] + e^s_k \) must be replaced by \( p^s_k = f^{-1}[h^s(z^s_k,b^s) + \gamma h^{s,1/2} e^s_k] \) where the \( e^s_k \) are the residuals for the transformed period s hedonic regression.

\[34\] It can be seen that the expression on the right hand side of (43) is a Paasche price index, where the price and quantity data of period t, \( p^t_k \) and \( q^t_k \) for \( k \in S(t) \), act as the comparison period data and the hedonic regression period s predicted prices, \( f^{-1}[h^s(z^s_k,b^s)] \) for \( k \in S(s) \), act as base period prices.
Thus our old Paasche type index \( P_p(s,t) \) is approximately equal to the numerator of our new Paasche type index \( P^t(s) \). However, as we mentioned before, the denominator of \( P^t(s) \), \( P^t(s) \), will be approximately equal to 1, and hence, our new Paasche type index will be approximately equal to our old Paasche type index; i.e., we have

\[
\text{(46) } P^t(s) \approx P_p(s,t).
\]

Now consider our new Paasche type index for the case where there are only two models: model 1, which is available in period \( s \) but not period \( t \), and model 2, which is available in period \( t \) but not period \( s \) so that we are revisiting the sampling model that was studied in section 4 above. Under these conditions, \( P^t(s) \) defined by (44) simplifies to \( r(2) \) defined in section 4 by (24). Hence our new Paasche type index is perfectly consistent with the hedonically adjusted sampling price ratio \( r(2) \) defined earlier in section 4.

Obviously, the above analysis can be repeated except that the hedonic regression for period \( t \) is used to do the quality adjustment rather than the period \( s \) hedonic regression. Thus, we now suppose that all purchasers of the hedonic commodity evaluate the relative utility of each model in period \( t \) according to the cardinal utility function \( g^t(z) \). Then in equilibrium, the period \( t \) price of a model with characteristics vector \( z \) should be proportional to \( g^t(z) \). Suppose that the period \( t \) econometrically estimated hedonic function, \( f^t[h(z,b)] \), can provide an adequate approximation to the period \( t \) theoretical hedonic function, \( \rho^t g^t(z) \), where \( \rho^t \) is a positive constant. Under these strong assumptions, the total market utility for period \( t \) that is provided by purchases of the hedonic commodities is equal to:

\[
\text{(47) } Q^t = \sum_{k \in S(t)} \rho^t g^t(z_k)q_k^t
\]

where we have approximated the utility to purchasers of model \( k \) in period \( t \), \( \rho^t g^t(z_k) \), by the period \( t \) hedonic regression estimated value, \( f^t[h(z_k,b)] \). Thus \( Q^t \) can be interpreted as the aggregate quantity of all of the models purchased in period \( t \), where each model has been quality adjusted into constant utility units using the period \( t \) hedonic aggregator function, \( g^t(z) \). In what follows, we will again neglect the approximation error between lines 1 and 2 of (47) so that we identify the period \( t \) aggregate quantity purchased of the hedonic commodity, \( Q^t(t) \), using the period \( t \) hedonic regression to do the quality adjustment, as follows:

\[
\text{(48) } Q^t(t) = \sum_{k \in S(t)} f^t[h(z_k,b)]q_k^t.
\]
Corresponding to the period $t$ quantity aggregate defined by (48), we can define an aggregate period $t$ price level, $P^t(t)$, by dividing $Q^t(t)$ into the period $t$ value aggregate, $V^t$:

\begin{align*}
(49) & \quad P^t(t) = \frac{V^t}{Q^t(t)} \\
& = \frac{V^t}{\sum_{k \in S(t)} \int [h'(z_k^t, b^t)] q_k^t} \quad \text{using (48)} \\
& = \frac{1}{\sum_{k \in S(t)} \{ \int [h'(z_k^t, b^t)]/p_k^t \} p_k^t q_k^t V^t} \\
& = \frac{1}{\sum_{k \in S(t)} \{ \int [h'(z_k^t, b^t)]/p_k^t \} s_k^t} \\
& = \left[ \sum_{k \in S(t)} s_k^t \{ \int [h'(z_k^t, b^t)]/p_k^t \} \right]^{-1}.
\end{align*}

Thus the aggregate period $t$ price level using the period $t$ hedonic regression, $P^t(t)$, is equal to a period $t$ share weighted harmonic mean of the period $t$ actual model prices, $p_k^t$, relative to the corresponding predicted period $t$ model prices using the period $t$ hedonic regression, $f^{-1}[h'(z_k^t, b^t)]$. Since $p_k^t = f^{-1}[h'(z_k^t, b^t)] + e_k^t$ where $e_k^t$ is the regression residual for model $k$ in period $t$ and these residuals are typically close to 0 and randomly distributed around 0, it can be seen that under normal conditions, $P^t(t)$ defined by (49) will be close to 1.

Now use the period $t$ hedonic regression to form a constant utility quantity aggregate for the models sold in period $s$. Thus model $k$ in period $s$, using the estimated hedonic valuation function of period $t$, will have the constant utility value $f^{-1}[h'(z_k^s, b^s)]$. Hence, the period $s$ aggregate quantity purchased of the hedonic commodity, $Q^s(t)$, using the period $t$ hedonic regression to do the quality adjustment into constant utility units, can be defined as follows:

\begin{align*}
(50) & \quad Q^s(t) = \sum_{k \in S(s)} \int [h'(z_k^s, b^s)] q_k^s.
\end{align*}

Corresponding to the period $s$ quantity aggregate defined by (50), we can define an aggregate period $s$ price level using the preferences of period $t$ to do the quality adjustment, $P^s(t)$, by dividing $Q^s(t)$ into the period $s$ value aggregate, $V^s$:

\begin{align*}
(51) & \quad P^s(t) = \frac{V^s}{Q^s(t)} \\
& = \frac{V^s}{\sum_{k \in S(s)} \int [h'(z_k^s, b^s)] q_k^s} \quad \text{using (50)} \\
& = \frac{1}{\sum_{k \in S(s)} \{ \int [h'(z_k^s, b^s)]/p_k^s \} p_k^s q_k^s V^s} \\
& = \frac{1}{\sum_{k \in S(s)} \{ \int [h'(z_k^s, b^s)]/p_k^s \} s_k^s} \\
& = \left[ \sum_{k \in S(s)} s_k^s \{ \int [h'(z_k^s, b^s)]/p_k^s \} \right]^{-1}.
\end{align*}

Thus the aggregate period $s$ price level using the period $t$ hedonic regression, $P^s(t)$, is equal to the reciprocal of a period $s$ share weighted arithmetic mean of the predicted period $s$ model prices in period $t$ using the period $t$ hedonic regression, $f^{-1}[h'(z_k^s, b^s)]$, relative to the period $s$ actual model prices, $p_k^s$.\(^{36}\)

Having defined the period $s$ price level $P^s(t)$ by (51) and the corresponding period $t$ price level $P^t(t)$ by (49) using the hedonic regression of period $t$ to do the constant utility quality

\(^{35}\) It can be seen that the expression on the right hand side of (49) is a type of Paasche price index, where the price and quantity data of period $t$, $p_k^t$ and $q_k^t$ for $k \in S(t)$, act as the comparison period data and the hedonic regression period $t$ predicted prices, $f^{-1}[h'(z_k^t, b^t)]$ for $k \in S(t)$, act as base period prices.

\(^{36}\) It can be seen that the expression on the right hand side of (51) is the reciprocal of a kind of Laspeyres price index, where the price and quantity data of period $s$, $p_k^s$ and $q_k^s$ for $k \in S(s)$, act as the base period price and quantity data and the hedonic regression period $t$ predicted prices, $f^{-1}[h'(z_k^s, b^s)]$ for $k \in S(s)$, act as comparison period prices.
adjustment, we can take the ratio of these two price levels to form a Laspeyres type price index going from period \( s \) to \( t \), using the hedonic regression of period \( t \), as follows:

\[
P(t) = \frac{P(t)}{P(t)} = \frac{\sum_{k \in S(t)} \sum_{k \in S(s)} \left( \frac{p_k}{f^{-1}\left[h(z_k^s, b^s)\right]} \right)^{-1}}{\sum_{k \in S(s)} \left( \frac{p_k}{f^{-1}\left[h(z_k^s, b^s)\right]} \right)^{-1}}
\]

where the last line above follows from the assumption that \( \sum_{k \in S(s)} \left( \frac{p_k}{f^{-1}\left[h(z_k^s, b^s)\right]} \right)^{-1} \) will be approximately equal to 1. 37

The above Laspeyres type index can be compared with our earlier Laspeyres type index defined by (33):

\[
PL(s, t) = \frac{\sum_{k \in S(s)} \left( \frac{p_k}{f^{-1}\left[h(z_k^s, b^s)\right]} \right)^{-1}}{\sum_{k \in S(s)} \left( \frac{p_k}{f^{-1}\left[h(z_k^s, b^s)\right]} \right)^{-1}}
\]

Thus our old Laspeyres type index \( PL(s, t) \) is approximately equal to our new Laspeyres type index \( P(t) \).

Consider our new Laspeyres type index for the case where there are only two models: model 1, which is available in period \( s \) but not period \( t \), and model 2, which is available in period \( t \) but not period \( s \) so that we are revisiting the sampling model that was studied in section 4 above. Under these conditions, \( P(t) \) defined by (52) simplifies to \( r(4) \) defined in section 4 by (29). Hence our new Laspeyres type index, \( P(t) \), is perfectly consistent with the hedonically adjusted sampling price ratio \( r(4) \) defined earlier in section 4.

As usual, if hedonic regressions are available for both periods \( s \) and \( t \), then the two indexes \( P(s) \) and \( P(t) \), defined by (44) and (52) respectively, should be averaged geometrically to form a final Fisher type estimate of price change going from period \( s \) to \( t \).

We now turn our attention to bilateral hedonic regressions (i.e., hedonic regressions that involve the data of two periods instead of only one period) that also make use of a time dummy variable.

\[
\text{Footnote: } 37 \text{ The last line on the right hand side of (52) is the hedonic index that is advocated by Pakes (2001; 26). Pakes assumes that } s = t-1.
\]
6. Unweighted Bilateral Hedonic Regressions with Time as a Dummy Variable

We now consider the following hedonic regression model, which utilizes the data of periods $s$ and $t$:

\begin{align}
(54) \quad f(p_k^s) &= \beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n + \varepsilon_k^s; \quad k \in S(s); \\
(55) \quad f(p_k^t) &= \gamma_{st} + \beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n + \varepsilon_k^t; \quad k \in S(t); 
\end{align}

where $\varepsilon_k^s$ and $\varepsilon_k^t$ are independently distributed error terms with mean 0 and variance $\sigma^2$, $f(x)$ is either the identity function $f(x) = x$ or the natural logarithm function $f(x) = \ln x$ and the functions of one variable $f_n$ are either the identity function, the logarithm function or a dummy variable which takes on the value 1 if the characteristic $n$ is present in model $k$ or 0 otherwise. Note that the $\beta$ regression coefficients in (54) are constrained to be the same as the corresponding $\beta$ coefficients in (55). Note also that equations (55) have added a time dummy variable, $\gamma_{st}$, and this coefficient will summarize the overall price change in the various models going from period $s$ to $t$.

Before proceeding further, we briefly discuss some of the advantages and disadvantages of the dummy variable model defined by (54) and (55) versus running separate single period regressions of the type defined by (1) for periods $s$ and $t$ and then using these separate regressions to form two separate estimates of quality adjusted prices which would be averaged in some way in order to form an overall measure of price change between periods $s$ and $t$. The main advantage of the latter method is that it is more flexible; i.e., changes in tastes between periods can readily be accommodated. However, this method has the disadvantage that two distinct estimates of period $s$ to $t$ price change will be generated by the method (one using the regression for period $s$ and the other using the regression for period $s$) and it is somewhat arbitrary how these two estimates are to be averaged to form a single estimate of price change. The main advantages of the dummy variable method are that it conserves degrees of freedom and is less subject to multicollinearity problems and there is no ambiguity about the measure of overall price change between periods $s$ and $t$.

We have considered only the case of two periods since this is the case of most interest to statistical agencies who must provide measures of price change between two periods.

---

38 This two period time dummy variable hedonic regression (and its extension to many periods) was first considered explicitly by Court (1939; 109-111) as his hedonic suggestion number two. Court (1939; 110) chose to transform the prices by the log transformation on empirical grounds: "Prices were included in the form of their logarithms, since preliminary analysis indicated that this gave more nearly linear and higher simple correlations." Court (1939; 111) then used adjacent period time dummy hedonic regressions as links in a longer chain of comparisons extending from 1920 to 1939 for US automobiles: "The net regressions on time shown above are in effect price link relatives for cars of constant specifications. By joining these together, a continuous index is secured." If the two periods being compared are consecutive periods, Griliches (1971b; 7) coined the term "adjacent year regression" to describe this dummy variable hedonic regression model.

39 This advantage was noted by Griliches (1971b; 8): "The time dummy approach does have the advantage, if the comparability problem can be solved, of allowing us to ignore the ever present problem of multicollinearity among the various dimensions."

40 Griliches (1971b; 7) has the following very nice summary justification for the use of the time dummy variable method: "The justification for this [method] is very simple and appealing: we allow as best we can for all of the major differences in specifications by 'holding them constant' through regression techniques. That part of the average price change which is not accounted for by any of the included specifications will be reflected in the coefficient of the time dummy and represents our best estimate of the 'unexplained-by-specification-change average price change.'"
However, the bilateral model defined by (54) and (55) can encompass both the fixed base situation (where s will equal the base period 0) or the chained situation where s will equal t-1. It is also of interest to consider the two period case because in this situation, we can draw on many of the ideas that have been introduced into bilateral index number theory, which also deals with the problem of measuring price change between two periods.

We first consider the case where f is the identity transformation. Let us estimate the unknown parameters in (54) and (55) by least squares regression and denote the estimates for the $\beta_n$ by $b_n$ for $n = 0, 1, ..., N$ and the estimate for $\gamma_{st}$ by $c_{st}$. Denote the least squares residuals for equations (54) and (55) with f defined to be the identity transformation by $e_{st}$ and $e_{k}^t$ respectively. Then we have the following equations, which relate the model prices in the two periods to their predicted values and the sample residuals:

\begin{align*}
(56) \quad & p_k^s = b_0 + \sum_{i=1}^{N} f_n(z_{kn})b_n + e_k^s; \quad k \in S(s); \\
(57) \quad & p_k^t = c_{st} + b_0 + \sum_{n=1}^{N} f_n(z_{kn})b_n + e_k^t; \quad k \in S(t).
\end{align*}

Now consider a hypothetical situation where the models sold during periods s and t are exactly the same so that there are say K common models pertaining to the two periods. Suppose further that the model prices in period t are all exactly $\lambda$ times greater than the corresponding model prices in period s, where $\lambda$ is a positive constant. Under these conditions, it seems reasonable to ask that the regression predicted values for the period t models be exactly equal to $\lambda$ times the regression predicted values for the same models in period s; i.e., we want the following equations to be satisfied:

\begin{align*}
(58) \quad & c_{st} + b_0 + \sum_{n=1}^{N} f_n(z_{kn})b_n = \lambda [b_0 + \sum_{n=1}^{N} f_n(z_{kn})b_n]; \quad k = 1, ..., K.
\end{align*}

In general, if $K > N+2$ and $\lambda \neq 1$, it can be seen that equations (56) cannot be solved for any coefficients $c_{st}$, $b_0$, $b_1$, ..., $b_N$. Hence, our conclusion is that the linear time dummy hedonic regression model defined by (56) and (57) is not a very good one, since it will not give us the "right" answer in a simple situation where all model prices are proportional for the two periods.\footnote{Diewert (2001) also argued on theoretical grounds that dummy variable hedonic regression models that used untransformed prices as dependent variables did not have good properties.} Of course, this homogeneity problem with the linear dummy variable regression model can be solved if we replace equations (57) by the following equations:

\begin{align*}
(59) \quad & p_k^t = c_{st}b_0 + \sum_{n=1}^{N} f_n(z_{kn})b_n + e_k^t; \quad k \in S(t).
\end{align*}

In equations (59), the time dummy variable, $c_{st}$, now appears in a multiplicative fashion. Thus, the problem with the estimating equations (56) and (59) is that we no longer have a linear regression model; nonlinear estimation techniques would have to be used.

Since nonlinear regression models are more difficult to estimate and may suffer from reproducibility problems, we will turn our attention to the second set of bilateral hedonic regression models, where f is the log transformation. In this case, the counterparts to equations (56) and (57) are the following equations:

\begin{align*}
(60) \quad & \ln p_k^s = b_0 + \sum_{i=1}^{N} f_n(z_{kn})b_n + e_k^s; \quad k \in S(s); \\
(61) \quad & \ln p_k^t = c_{st} + b_0 + \sum_{n=1}^{N} f_n(z_{kn})b_n + e_k^t; \quad k \in S(t).
\end{align*}
Exponentiating both sides of (60) and (61) leads to the following equations that will be satisfied by the data and the least squares estimators for (60) and (61):

\[(62) \quad p_k^s = \exp[b_0 + \sum_{n=1}^{N} f_n(z_{kn}^s) b_n] \exp[e_k^s] \quad k \in S(s)\]

\[(63) \quad p_k^t = \exp[c_n] \exp[b_0 + \sum_{n=1}^{N} f_n(z_{kn}^t) b_n] \exp[e_k^t] \quad k \in S(t)\]

Again consider a hypothetical situation where the models sold during periods s and t are exactly the same so that there are K common models pertaining to the two periods. Again suppose that the model prices in period t are all exactly \(\lambda\) times greater than the corresponding model prices in period s, where \(\lambda\) is a positive constant. Again we ask that the regression predicted values for the period t models be exactly equal to \(\lambda\) times the regression predicted values for the same models in period s; i.e., we want the following equations to be satisfied:

\[(64) \quad \exp[c_n] \exp[b_0 + \sum_{n=1}^{N} f_n(z_{kn}^t) b_n] = \lambda \{ \exp[b_0 + \sum_{n=1}^{N} f_n(z_{kn}^s) b_n] \} \quad k = 1, \ldots, K.\]

It can be seen that if we choose \(c_n = \ln \lambda\), then we can satisfy equations (64). Hence we conclude (from a test approach perspective) that it is preferable to run linear bilateral dummy variable hedonic regressions using the log transformation for the dependent variable rather than leaving the model prices untransformed. Thus, we have again reinforced the case for using the log transformation on the dependent variable in hedonic regression models.

The bilateral log hedonic regression model is defined by (54) and (55) where \(f\) is the log transformation. It can be seen that in this case, the theoretical index of price change going from period \(s\) to \(t\) is \(\exp[\gamma_{st}]\) and the sample estimator of this population measure is:

\[(65) \quad P(s,t) = \exp[c_{st}]\]

where \(c_{st}\) is the least squares estimator for the shift parameter \(\gamma_{st}\). Note that we put the shift parameter in equations (55) rather than in equations (54). The choice of base period should not matter so let us consider the following bilateral log regression model which puts the shift parameter \(\gamma_{ts}\) in the period \(s\) equations rather than in the period \(t\) equations:

\[(66) \quad \ln p_k^s = \gamma_{ts} + \beta_0^s + \sum_{n=1}^{N} f_n(z_{kn}^s) \beta_n^s + \varepsilon_k^s \quad k \in S(s)\]

\[(67) \quad \ln p_k^t = \beta_0^t + \sum_{n=1}^{N} f_n(z_{kn}^t) \beta_n^t + \varepsilon_k^t \quad k \in S(t)\]

Denote the least squares estimates for \(\beta_n^s\) by \(b_n^s\) for \(n = 0, 1, \ldots, N\) and the estimate for \(\gamma_{ts}\) by \(c_{st}\). For the regression model defined by (66) and (67), it can be seen that the theoretical index of price change going from period \(t\) to \(s\) is \(\exp[\gamma_{ts}]\) and the sample estimator of this population measure is:

\[(68) \quad P(t,s) = \exp[c_{ts}]\]

The question now is: how does \(P(s,t)\) defined by (65) relate to \(P(t,s)\) defined by (68)? Ideally, we would like these two estimators of price change to satisfy the following \textit{time reversal test}:

\[(69) \quad P(t,s) = 1/P(s,t)\]

If we compare the original log linear regression model defined by (54) and (55) (with \(f\) being the log transformation) with the new model defined by (66) and (67), it can be seen that the
right hand side exogenous variables are identical except that $\gamma_{ts}$ appears in the first set of equations in (66) and (67) while $\gamma_{st}$ appears in the second set of equations in (54) and (55). The transpose of the column in the X matrix that corresponds to $\gamma_{ts}$ in (66) and (67) is equal to $[1_1^T,0_2^T]$ where $1_1$ is a column vector of ones of dimension equal to the number of models in the set $S(s)$ and $0_2$ is a column vector of zeros of dimension equal to the number of models in the set $S(t)$. The transpose of the column in the X matrix that corresponds to $\gamma_{st}$ in (54) and (55) is equal to $[0_1^T,1_2^T]$ where $0_1$ is a column vector of zeros of dimension equal to the number of models in the set $S(s)$ and $1_2$ is a column vector of ones of dimension equal to the number of models in the set $S(t)$. However, note that both models have the constant term $\beta_0$ (or $\beta_0^*$) in every equation and the transpose of the column in the X matrix that corresponds to this constant term is equal to $[1_1^T,1_2^T]$ in both models. It can be seen that the subspace spanned by the X columns corresponding to $\beta_0$ and $\gamma_{st}$ in (54) and (55) is equal to the subspace spanned by the X columns corresponding to $\beta_0^*$ and $\gamma_{ts}$ in (66) and (67) and the two sets of parameters are related by the following equations:

\[
(70) \begin{bmatrix} 0_1^T,1_2^T \end{bmatrix} \gamma_{st} + [1_1^T,0_2^T] \beta_0 = [1_1^T,0_2^T] \gamma_{ts} + [1_1^T,1_2^T] \beta_0^*.
\]

Equations (70) are equivalent to the following 2 equations in the four variables $\gamma_{st}$, $\beta_0$, $\gamma_{ts}$ and $\beta_0^*$:

\[
(71) \begin{align*}
0 & \gamma_{st} + 1 \beta_0 = 1 \gamma_{ts} + 1 \beta_0^*; \\
1 & \gamma_{st} + 1 \beta_0 = 0 \gamma_{ts} + 1 \beta_0.
\end{align*}
\]

Thus given $\gamma_{st}$ and $\beta_0$, the corresponding $\gamma_{ts}$ and $\beta_0^*$ can be obtained using equations (71) as:

\[
(72) \begin{align*}
\gamma_{ts} &= -\gamma_{st}; \\
\beta_0^* &= \gamma_{ts} + \beta_0.
\end{align*}
\]

Equations (72) also hold for the least squares estimators for the two hedonic regression models. In particular, we have:

\[
(73) c_{ts} = -c_{st}.
\]

Hence, exponentiating both sides of (73) gives us $\exp[c_{ts}] = 1/\exp[c_{st}]$ and this equation is equivalent to (69) using definitions (68) and (65). Thus we have shown that the estimator of price change $P(s,t)$ defined by (65) (which corresponds to the least squares estimators of the initial log hedonic regression model defined by (54) and (55) with $f(p) = \ln p$) is equal to the reciprocal of the estimator of price change $P(t,s)$ defined by (68) (which corresponds to the second log hedonic regression model defined by (66) and (67) so that the two bilateral dummy variable hedonic regressions satisfy the time reversal test (69).

The results in this section strongly support the use of the logarithms of model prices as the dependent variables in a bilateral hedonic regression model with a time dummy variable. In the following section, we will study the properties of weighted bilateral hedonic regression models.

### 7. Weighted Bilateral Hedonic Regressions with Time as a Dummy Variable

Given the results in the previous section, we consider only weighted bilateral hedonic regressions that use the log of model prices as the dependent variable, before weighting the
equations. We also draw on the results in section 3 and consider only value weighting. Thus we now consider the following value weighted hedonic regression model, which utilizes the data of periods $s$ and $t$:

$$
\begin{align*}
(74) \quad & \left( v_{ks} \right)^{1/2} \ln p_{ks} = \left( v_{ks} \right)^{1/2} \left[ \beta_0 + \sum_{n=1}^{N} f_n(z_{kn}) \beta_n \right] + \varepsilon_{ks}^s; \\
& \quad k \in S(s); \\
(75) \quad & \left( v_{kt} \right)^{1/2} \ln p_{kt} = \left( v_{kt} \right)^{1/2} \left[ \gamma_{st}^s + \beta_0 + \sum_{n=1}^{N} f_n(z_{kn}) \beta_n \right] + \varepsilon_{kt}^t; \\
& \quad k \in S(t);
\end{align*}
$$

where the model sales values for period $t$, $v_{kt}^t$, were defined by (14) and $\varepsilon_{ks}^s$ and $\varepsilon_{kt}^t$ are independently distributed error terms with mean 0 and variance $\sigma^2$.

The weighted model defined by (74) and (75) is the bilateral counterpart to our single equation weighted hedonic regression model that was studied in section 3 above. However, in the present bilateral context, we now encounter a problem that was absent in the single equation context. The problem is this: if there is high inflation going from period $s$ to $t$, then the period $t$ model sales values $v_{kt}^t$ can be very much bigger than the corresponding period $s$ model sales values $v_{ks}^s$ due to this general inflation. Hence, the assumption of homoskedastic residuals between equations (74) and (75) is unlikely to be satisfied. Hence, it is necessary to pick new weights that will eliminate this problem.

Our tentative initial suggested solution to the above problem caused by general inflation between the two periods is to use the model expenditure shares, $s_{ks}$ and $s_{kt}$ defined earlier by (40) as the weights in (74) and (75) in place of the model expenditures, $v_{ks}^s$ and $v_{kt}^t$. Thus we recommend the use of the following expenditure share weighted hedonic regression model, which utilizes the data of periods $s$ and $t$:

$$
\begin{align*}
(76) \quad & \left( s_{ks} \right)^{1/2} \ln p_{ks} = \left( s_{ks} \right)^{1/2} \left[ \beta_0 + \sum_{n=1}^{N} f_n(z_{kn}) \beta_n \right] + \varepsilon_{ks}^s; \\
& \quad k \in S(s); \\
(77) \quad & \left( s_{kt} \right)^{1/2} \ln p_{kt} = \left( s_{kt} \right)^{1/2} \left[ \gamma_{st}^s + \beta_0 + \sum_{n=1}^{N} f_n(z_{kn}) \beta_n \right] + \varepsilon_{kt}^t; \\
& \quad k \in S(t);
\end{align*}
$$

where $\varepsilon_{ks}^s$ and $\varepsilon_{kt}^t$ are independently distributed error terms with mean 0 and variance $\sigma^2$.

Denote the least squares estimates for $\beta_n$ by $b_n$ for $n = 0, 1, \ldots, N$ and the estimate for $\gamma_{st}$ by $c_{st}$. For the regression model defined by (76) and (77), it can be seen that the theoretical index of price change going from period $t$ to $s$ is $\exp[\gamma_{st}]$ and the sample estimator of this population measure is:

$$
(78) \quad P_{1}(s,t) = \exp[c_{st}].
$$

It can be shown that $P_{1}(s,t)$ defined by (78) in this section has the same desirable property that $P(s,t)$ defined by (65) in the previous section had: namely, if the models are identical in the two periods (and the model expenditure shares are identical for the two periods) and the model prices in period $t$ are all exactly $\lambda$ times greater than the corresponding model prices in period $s$, then $P_{1}(s,t)$ is exactly equal to $\lambda$.\(^{43}\)

\(^{42}\)Diewert (2002b) considered a model similar to (76) and (77) except that all of the explanatory variables were dummy variables and showed that weighting by the square roots of expenditure shares led to a very reasonable index number formula to measure the price change between the two periods. Thus the model defined by (76) and (77) is consistent with his results.

\(^{43}\)See Proposition 1 in the Appendix.
The restriction that the expenditure shares be identical in the two periods in the identical model case is a bit unrealistic. Moreover, in the identical models case, it would be nice if \( P_1(s,t) \) defined by (78) turned out to equal the Törnqvist price index, since this index is a preferred one from the viewpoints of both the stochastic and economic approaches to index number theory.\(^{44}\) Hence in place of the model defined by (76) and (77), when a model is present in both periods, let us use the average sales share for that model, \( (1/2)(s_k^s + s_k^t) \), as the weight for that model in both periods. In this revised weighting scheme, the old period \( s \) equations (76) are replaced by the following two sets of equations:

\[
(79) \quad (s_k^s)^{1/2} \ln p_k^s = (s_k^s)^{1/2}[\beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n] + \varepsilon_k^s; \quad k \in [S(s)\cap S(t)];
\]

\[
(80) \quad [(1/2)(s_k^s+s_k^t)]^{1/2} \ln p_k^s = [(1/2)(s_k^s+s_k^t)]^{1/2}[\beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n] + \varepsilon_k^s; \quad k \in S(s)\cap S(t).
\]

Thus if a model \( k \) is present in period \( s \) but not present in period \( t \), then we use the square root of the period \( s \) sales share for that model, \( (s_k^s)^{1/2} \), as the weight, which means this model is included in equations (79). On the other hand, if model \( k \) is present in both periods, then we use the square root of the arithmetic average of the period \( s \) and \( t \) sales shares for that model, \( [(1/2)(s_k^s+s_k^t)]^{1/2} \), as the weight, which means this model is included in equations (80).

Similarly, the old period \( s \) equations (77) are replaced by the following two sets of equations:

\[
(81) \quad (s_k^t)^{1/2} \ln p_k^t = (s_k^t)^{1/2}[\gamma_{st} + \beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n] + \varepsilon_k^t; \quad k \in [S(t)\cap S(s)];
\]

\[
(82) \quad [(1/2)(s_k^s+s_k^t)]^{1/2} \ln p_k^t = [(1/2)(s_k^s+s_k^t)]^{1/2}[\gamma_{st} + \beta_0 + \sum_{n=1}^{N} f_n(z_{kn})\beta_n] + \varepsilon_k^t; \quad k \in S(s)\cap S(t).
\]

Thus if a model \( k \) is present in period \( t \) but not present in period \( s \), then we use the square root of the period \( t \) sales share for that model, \( (s_k^t)^{1/2} \), as the weight, which means this model is included in equations (81). On the other hand, if model \( k \) is present in both periods, then we use the square root of the arithmetic average of the period \( s \) and \( t \) sales shares for that model, \( [(1/2)(s_k^s+s_k^t)]^{1/2} \), as the weight, which means this model is included in equations (82). As usual, we assume that \( \varepsilon_k^s \) and \( \varepsilon_k^t \) are independently distributed error terms with mean 0 and variance \( \sigma^2 \).

Denote the least squares estimates for \( \beta_n \) by \( \hat{b}_n \) for \( n = 0,1,\ldots,N \) and the estimate for \( \gamma_{st} \) by \( c_{st} \).

For the regression model defined by (79)-(82), it can be seen that the theoretical index of price change going from period \( t \) to \( s \) is \( \exp[k_{st}] \) and the sample estimator of this population measure is:

\[
(83) \quad P_2(s,t) = \exp[c_{st}].
\]

It can be shown that \( P_2(s,t) \) defined by (83) has the following desirable property: if the models are identical in the two periods, then \( P_2(s,t) \) is equal to the Törnqvist price index between the two periods.\(^{45}\) Hence it appears that the weighted hedonic regression model defined by (79)-(82) is a "natural" weighted hedonic regression model that provides a generalization of the Törnqvist price index to cover the case where the models are not matched. If there are no models in common for the two periods under consideration, then the model defined by (79)-(82) reduces to our earlier model defined by (76)-(77).

\(^{44}\) See Diewert (2002a).

\(^{45}\) This follows from Corollary 5.2 in the Appendix.
As in the previous section, it is somewhat arbitrary whether we put the time dummy variable in the period t equations or whether we put it in the period s equations. If we put the time dummy in the period s equations as the parameter $\gamma_{ts}$ and obtain a weighted least squares estimate $c_{ts}$ for this population parameter, the theoretical index of price change going from period t to s is $\exp[\gamma_{ts}]$ and the sample estimator of this population measure is:

$$ (84) \quad P^*(t,s) = \exp[c_{ts}] $$

As in the previous section, we would like $P^*(t,s)$ to equal the reciprocal of $P(s,t)$. It turns out that this property is true for the weighted hedonic regressions defined by (76) and (77) and (79)-(82) in this section as well as for the unweighted ones defined in the previous section; see Proposition 4 in the Appendix. Hence it does not matter whether we put the time dummy variable in period s or t: our measure of overall price change between the two periods will be invariant to this choice for the two weighted hedonic regressions considered in this section.

Using the results in the Appendix, we can also show that $P_1(s,t)$ and $P_2(s,t)$ both satisfy the identity test (A6), the homogeneity tests (A4) and (A5) and the time reversal test (A7) as we have already indicated. Thus both of these hedonic price indexes have some good a priori properties.

Which bilateral weighted hedonic index is best? From the viewpoint of representativity, $P_1(s,t)$ seems best; the models present in each period are weighted by expenditure shares that pertain to that period. However, the loss of representativity for $P_2(s,t)$ is probably not large in most applications and $P_2(s,t)$ has the advantage of being consistent with the use of a Törnqvist price index in the matched models case.46 Thus at this stage of research, we lean towards the use of $P_2(s,t)$.

We turn now to a discussion of the treatment of regression outliers.

8. Outliers and Influence Analysis

In the context of traditional sampling techniques used by statistical agencies, usually provision is made for the deletion of outliers in the samples of prices collected. This raises the issue as to whether outliers should also be deleted in the hedonic regression context.

In the unweighted context, the deletion of sample outlier observations should be permitted. Since influence analysis is just an extension of outlier analysis (an influential observation is one which greatly influences the estimated regression coefficients), the deletion of influential observations should also be permitted.47

However, in the weighted context, the situation is somewhat different for two reasons.

Assuming that we have complete market information on the prices and quantities sold of all models being considered for the two periods under consideration, then we are in the same situation as we would be if we were applying traditional bilateral index number theory. In

46 Moreover, in the matched models case, $P_2(s,t)$ has the advantage of being independent of the hedonic regression coefficients $b_0, b_1, ..., b_N$, whereas $P_1(s,t)$ is not.
47 In the unweighted bilateral hedonic regression context, we need only delete observations that influence the estimate of $\gamma_{ts}$ since the other parameters are not of great significance in this context. For an exposition of the various approaches to influence analysis, see Chapter 4 in Chatterjee and Hadi (1988).
this latter context, (assuming reliable data), traditional index number theory does not suggest dropping out prices and quantities that look a bit unusual.\textsuperscript{48} Thus the dropping of outliers or influential observations in the context of running a hedonic regression could lead to results that would not be comparable to the results of say a maximum overlap superlative index.

Our second reason for advocating a bit of caution in dropping outliers or influential observations in the weighted hedonic regression context is that in this weighted context, an individual observation does not have equal weight! Consider the case of a hugely popular model that accounts for almost all of the sales in a given period. Dropping this weighted observation would frequently lead to a big change in the weighted regression but on representativeness grounds, we would not want to drop this observation. Hence traditional outlier and influence analysis would have to be somehow adapted to deal with this situation. Until this new methodology is available, I would urge a cautious approach to the dropping of observations in the context of weighted hedonic regressions.

9. Do the Signs of Hedonic Regression Coefficients Matter?

A recent paper by Pakes (2001) has stimulated a certain amount of controversy in the literature on hedonic regressions. Hulten (2002) has provided a nice summary of the more controversival parts of Pakes' analysis and Hulten labels these controversival Propositions as Pakes I, Pakes II and Pakes III. We will follow Hulten's interpretation of Pakes below.

Hulten (2002; 23) states the Pakes Proposition I as follows: the hedonic function is equal to a producers' marginal cost function plus a market power function that depends on the elasticities of demand for characteristics.\textsuperscript{49}

Hulten's (2002; 25) Pakes Proposition II is the following corollary to Proposition I: the price of a product can go down when it acquires more of a given characteristic. In other words, the sign of a hedonic coefficient for a characteristic can go in the "wrong" direction!\textsuperscript{50}

Hulten's (2002; 25) Pakes Proposition III is that the two single period hedonic regressions pertaining to any two periods being compared may not bear any close relationship to each other (this follows from the first Proposition that implies that changing market power between the two periods might lead to quite different hedonic regressions) but only one of the two regressions is required to undertake a quality adjustment.\textsuperscript{51}

\textsuperscript{48} This is not quite true since there is some literature on measuring core inflation that suggests the deletion of outliers. However, the goal of this literature is usually to obtain better estimates of trend or future inflation. Traditional index numbers that do not drop observations are still acceptable as measures of past inflation.

\textsuperscript{49} Pakes (2001; 10) distinguishes between the competitive case and the more normal market power case as follows: "That is, in the marginal cost pricing equilibrium the hedonic function is the marginal cost function. However in the Bertrand equilibrium the hedonic function is the sum of the marginal cost function and a function that summarizes the relationship between markups and characteristics."

\textsuperscript{50} "Hedonic regressions have been used in research for some time and they are often found to have coefficients which are 'unstable' either over time or across markets, and which clash with the naive intuition that the 'marginal willingness to pay for a characteristic equaled its marginal cost of production'. I hope this discussion has made it amply clear that these models can be very misleading. The derivatives of a hedonic price function should not be interpreted as either willingness to pay derivatives or cost derivatives; rather they are formed from a complex equilibrium process." Ariel Pakes (2001; 14).

\textsuperscript{51} The hedonic indexes that Pakes (2001; 26) considers are all of the type defined by the last line of (52) above where the chain principle is used so that $s = t - 1$. 

96  International Working Group on Price Indices - Seventh Meeting
We discuss each of these Propositions below.

I agree with Pakes that the determination of model prices in any one period is determined by the interaction of demander’s preferences, the costs of model suppliers and the degree of market power of suppliers. However, I disagree with Pakes’ contention that the hedonic function must be equal to marginal cost function plus a markup function. It seems to me that Pakes is viewing the hedonic function through the eyes of producers when what is required is a consumer view. After all, the purpose of the hedonic exercise is to find how demanders (and not suppliers) of the product value alternative models in a given period. Thus for the present purpose, it is the preferences of consumers that should be decisive, and not the technology and market power of producers. The situation is similar to ordinary general equilibrium theory where equilibrium price and quantity for each commodity is determined by the interaction of consumer preferences and producer’s technology sets and market power. However, there is a big branch of applied econometrics that ignores this complex interaction and simply uses information on the prices that consumers face, the quantities that they demand and perhaps demographic information in order to estimate systems of consumer demand functions. Then these estimated demand functions are used to form estimates of consumer utility functions and these functions are often used in applied welfare economics. What producers are doing is entirely irrelevant to these exercises in applied econometrics with the exception of the prices that they are offering to sell at. In other words, we do not need information on producer marginal costs and markups in order to estimate consumer preferences: all we need are selling prices. I believe that the situation is similar in the context of estimating a hedonic price function for a given period. For welfare purposes, we need to assume that the hedonic model price function is proportional to a (separable from everything else) hedonic utility function that gives the utility that demanders will get as a function of model characteristics. We then make the heroic assumption that the actual prices of models that were sold during the period under consideration are proportional to this assumed hedonic utility function.

52 This position seems consistent with the position of Griliches (1971b; 14), who argued that it is the user value or utility of a model that is the “right” characteristic for government statisticians to attempt to measure: “Most economists would agree that they would like the ‘price’ index to be a ‘price-of-living’ or ‘utility’ indicator. Many government statisticians in charge of producing actual price indexes will reply that the cannot achieve this and that therefore they should not even try, but should concentrate instead on some more ‘objective’ index of ‘transaction’ prices and/or allow only for those ‘quality’ changes which are based on ‘production’ costs. The fact that ‘truth’ cannot be achieved doesn’t mean that one shouldn’t strive to do so, though I sympathize with the position that it is better to measure well something definite than to do a very poor job on a more interesting but also more nebulous concept. Nevertheless, I would deny the contention that ‘transaction’ units or ‘production’ costs are much more definitive concepts. In general, they too make little sense without some appeal to utility considerations.” Griliches (1971b; 14-15) went on to definitely reject the production cost viewpoint of adjusting for quality changes: “Nor are ‘production costs’ an adequate guide to quality changes without a check of their utility implications. ... Nor should we ignore ‘costless’ changes if we can measure them. If the consumer is in fact buying ‘horsepower’, and if a design change makes it possible to deliver more horsepower from the same size and ‘cost’ engine, then the price of horsepower to the consumer has fallen and he is better off!”

53 Hulten (2002; 26) also comments on the similarity of hedonic regression estimation with the estimation of conventional supply and demand functions: “Indeed, the Pakes II result has precedent in conventional price-quantity analysis. When the price of a good is regressed on its quantity, it is well known that the underlying supply and demand curves generally cannot be identified separately, and that the regression coefficients will be unstable and can easily have the ‘wrong’ sign.” This is true as far as it goes but what enables consumer demand analysis to “succeed” in this situation is that we do not estimate a single demand function but rather estimate a system of demand functions, where an exogenous identifying variable is “income”. If we were using the same price and quantity data to estimate a system of producer supply functions, there would be different exogenous variables appearing in the producer system such as capital and labor used by the producers.
It is clear that there are a number of problems with the above assumptions:

- The separability assumption is very unrealistic.

- Different demanders may have very different hedonic utility functions and so over time as characteristic costs and markups change, different classes of consumers may be induced to enter the market and thus the estimated hedonic utility function may be quite unstable over time.

- The assumption that the market is in equilibrium is also suspect, particularly in the context of new products, where it takes time for demanders to discover the new products.

- Another consequence of the equilibrium assumptions that we have made is that all models which are sold in a given period are equally desirable; i.e., they all yield equal utility per dollar spent. But in practice, some models are vastly more popular than others and our suggested approach does not directly take this fact into account.

In spite of the above problems, I believe that the consumer valuation approach is more appropriate in the context of making quality adjustments for CPI purposes than the producer valuation approach proposed by Pakes. Note that there do not appear to be any welfare implications whatsoever in making hedonic price adjustments using the framework of Pakes.

We turn now to Pakes Proposition II; namely, that the price of a product can go down when it acquires more of a given characteristic. In other words, the signs of hedonic regression coefficients do not have to be sign restricted. In evaluating this Proposition, it is again necessary to keep in mind the purpose for running the hedonic regression, which is to provide utility valuations for possibly hypothetical models that are sold in one period but not in the other period. For this welfare economics type purpose, I believe that it makes sense to impose a priori sign restrictions on the regression coefficients. Hence if we believe that demanders of a model will, on average, get a higher utility from a model that has more of an a priori desirable feature, then we should make sure that our estimated hedonic regression does not contradict these a priori beliefs. Thus we are taking the point of view here that we impose our theory on the data and do the econometric estimation which fits the data best, consistent with our prior beliefs. An alternative (but perfectly defensible point of view) is that we use the data to discover whether our a priori theories are consistent with the data. However, it

---

54 However the use of expenditure or share weighted hedonic regressions can partially overcome this defect of the theory.

55 This position seems to be consistent with the following remarks by Griliches (1971b; 8): “The time dummy approach does have the advantage, if the comparability problem can be solved, of allowing us to ignore the ever present problem of multicollinearity among the various dimensions. Using it, we may not care that in one year the coefficient of weight is high and horsepower is low while in another year these coefficients reverse themselves, as long as the two coefficients taken together hold the joint effect of weight and horsepower constant.”

56 But only for characteristics where we are fairly certain that more is better!

57 These a priori beliefs can be imposed on the regression by replacing coefficients by squares of coefficients and then using nonlinear regression estimation techniques.

58 Again there is an analogy with traditional consumer demand analysis: if we are interested in welfare analysis, we will impose the curvature conditions implied by economic theory on the econometric estimation method as is done by Diewert and Wales (1993) whereas if we are interested in testing traditional demand theory, then we would not impose these curvature conditions.
seems to me that this theory testing point of view is not what statistical agencies are interested in when they do quality adjustment: they want to make quality adjustments that are consistent with the public’s a priori view that more of a desirable characteristic should increase the price of the product (or at least not decrease it).

Finally, we discuss Pakes Proposition III; namely that only one of the two single period hedonic regressions that pertain to the two periods under consideration needs to be used in order to undertake a quality adjustment. Obviously, this Proposition is true! However, as we have argued in the earlier sections of this paper, if two estimates are available, then it always is better to use an average of the two estimates rather than just one of them. This is particularly true for Pakes’ preferred index defined by the last line of (52) since this index is likely to suffer from substitution or representativity bias. Hence it will usually be best to match up a Laspeyres type estimator like the estimator preferred by Pakes with a corresponding Paasche type estimator (if both are available).

10. Conclusion

The theory of hedonic regressions has left a great deal of leeway open to the empirical investigator with respect to the details of implementation of the models. Our strategy in this review of the issues has been to use some of the ideas that are present in the test approach to index number theory in an attempt to narrow down some of these somewhat arbitrary choices. The problem with arbitrary choices is that the end results may not be invariant to these choices and hence if hedonic regression techniques are used by statistical agencies, the resulting estimates of price change may not be reproducible. We have made a start on narrowing down some of these choices but much work remains to be done.

In order to narrow down the range of outcomes that can result from the use of hedonic regression techniques, we make the following suggestions:

- It seems preferable to use the log of the model price as the dependent variable rather than the model price itself.

- If expenditure weights are available, use them to weight the observations as suggested in this section for a dummy variable hedonic regression and as suggested in section 3. Expenditure weights are preferable to quantity weights.

- In the context of running single period hedonic regressions, it seems preferable to run separate regressions for both periods and use a symmetric average of the results from both regressions in the final measure of price change between the two periods.

- It is preferable to use hedonic regression techniques in the context of the chain principle where the prices of period $t$ and $t-1$ are compared since this will tend to minimize the spread between estimates of price change over longer periods that are obtained using alternative hedonic techniques.

59 Again, this position seems to be consistent with that of Griliches (1971b; 7).

60 This can be most clearly seen in the matched model context where the index defined by the last line of (52) will be approximately equal to the ordinary Laspeyres index between periods $s$ and $t$.

61 The work of Heravi and Silver (2002) shows that the use of different hedonic regression techniques can lead to quite different estimates of price change.
• It seems preferable to sign restrict regression coefficients in accordance with a priori theory.

One issue that we did not discuss above is whether hedonic regressions should include brand dummy variables as independent variables. The argument against doing this is that brand dummy variables should be superfluous if we have entered all of the important characteristics of the product into the regression and hence, by including brand dummy variables, we will just use up valuable degrees of freedom and increase multicollinearity. The argument for entering brand dummy variables is that they capture in an efficient manner certain characteristics of the product that would be difficult to specify otherwise. At the present stage of research in this area, I would be inclined to allow the use of brands as admissible dummy variables.

We conclude by noting that the cautious attitude towards the use of hedonic regressions expressed by Schultze and Mackie (2002) echoes the following comments made by Bean in his discussion of Court's (1939) pioneering paper on hedonic regressions:

"Mr. Court's interesting work should be carried much further, as he suggests. We should, however, not be disappointed if neither public agencies nor trade associations adopt the policy of publishing prices, values and index numbers based on the relatively tricky results that one is sure to get by applying the device of multiple correlation. The only group who would sponsor such a procedure would be the non-existent National Association of Experts in Multiple Correlation, the demand for whose services would be enormously increased." Louis H. Bean (1939; 119).

Hopefully, in the next few years, as users form a consensus on what the "best" procedures are, then the use of hedonic regressions by statistical agencies will become much more widespread and routine.

---

62 These difficult to specify characteristics might include reliability, availability of the product in the local marketplace and the degree of consumer knowledge about the product; i.e., some producers will choose to heavily advertise their products while others will not and the effect of this advertising may be to create a brand premium.

63 The hedonic regression Manual being prepared by Jack Triplett (2002) should help form this consensus.
Appendix: Properties of Bilateral Weighted Hedonic Regressions

We consider some of the mathematical properties of a slight generalization of the share weighted bilateral hedonic regression model defined by (76) and (77) in section 7. The generalization is that we do not restrict the weights to sum up to 1 in each period. Thus, we replace the period s share weights $s_k^s$ in (76) and the period t share weights $s_k^t$ in (77) by the positive weights $w_k^s$ and $w_k^t$ respectively, where these weights do not necessarily sum to 1 in each period. We assume that these weight functions are known functions of the price and quantity data pertaining to periods s and t; i.e., we have for some functions, $g_k^s$ and $g_k^t$:

(A1) $w_k^s = g_k^s(p^s,p^t,q^s,q^t)$ for $k \in S(s)$ ; $w_k^t = g_k^t(p^s,p^t,q^s,q^t)$ for $k \in S(t)$

where $p^s$ and $p^t$ are price vectors of the model prices for periods s and t respectively and $q^s$ and $q^t$ are the corresponding period s and t quantity vectors of the models sold in periods s and t. In the Propositions below, we will place further restrictions on the weighting functions $g_k^s$ and $g_k^t$ as they are needed.

The weighted least squares estimators for $\gamma_{st}$, $\beta_0$, $\beta_1$, ..., $\beta_N$ for this new model are the solutions $c_{st}$, $b_0$, $b_1$, ..., $b_N$ to the following quadratic weighted least squares minimization problem:

(A2) $\min_{b's and c} \left\{ \sum_{k \in S(s)} w_k^s [\ln p_k^s - b_0 - \sum_{n=1}^N f_n(z_{kn}^s)b_n]^2 + \sum_{k \in S(t)} w_k^t [\ln p_k^t - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^t)b_n]^2 \right\}$.

The bilateral price index $P$ that summarizes the overall change in prices going from period s to t is defined as the exponential of the $c_{st}$ solution to (A2); i.e., we have:

(A3) $P(p^s,p^t,q^s,q^t) = \exp[c_{st}^*]$.

We would like to show that the hedonic index number formula defined by (A3) has some of the properties that bilateral index number formulae defined over matched models usually have. Thus we are attempting to extend the test approach to index number theory to weighted bilateral hedonic regressions. In particular, we would like to establish the following properties for P:

(A4) Homogeneity of degree one in period t prices; i.e., $P(p^s,\lambda p^t,q^s,q^t) = \lambda P(p^s,p^t,q^s,q^t)$ for all $\lambda > 0$.

(A5) Homogeneity of degree minus one in period s prices; i.e., $P(\lambda p^s,p^t,q^s,q^t) = \lambda^{-1} P(p^s,p^t,q^s,q^t)$ for all $\lambda > 0$.

(A6) Identity; i.e., if the models in the two periods are identical and the selling prices are equal so that $p^s = p^t = p$ and, in addition, the same quantities of each model are sold in the two periods so that $q^s = q^t = q$, then the resulting price index $P(p,p,q,q) = 1$.

64 Throughout this Appendix, we assume that the X matrix that corresponds to the linear regression model defined by (A2) has full column rank so that the solution to (A2) exists and is unique.
65 The test approach to index number theory was largely developed by Walsh (1901) (1921), Fisher (1911) (1922) and Eichhorn and Voeller (1976). For more recent contributions, see Diewert (1992) (1993), Balk (1995) and von Auer (2001).
(A7) Time reversal; i.e., $P^*(p^1, p^s, q^t, q^s) = 1 / P(p^s, p^1, q^s, q^t)$.

The above property says that if we interchange the order of our data and measure the overall change in prices going backwards from period $t$ to $s$, then the resulting index $P^*(p^1, p^s, q^t, q^s)$ is equal to the reciprocal of the original index $P(p^s, p^1, q^s, q^t)$, which measured the degree of overall price change going from period $s$ to $t$. In order to formally define the price index $P^*$, let $c_{st}^*$, $b_0^*$, $b_1^*$, ..., $b_N^*$ be the solution to the following quadratic weighted least squares minimization problem, which corresponds to reversing the ordering of the two periods:

\[
(A8) \min_{b^s} \{ \sum_{k \in S(s)} w_k [\ln p_k^s - c_{st}^* - b_0 - \sum_{n=1}^N f_n(z_{kn}^s)b_n]^2 
+ \sum_{k \in S(t)} w_k [\ln p_k^t - c_{st}^* - b_0 - \sum_{n=1}^N f_n(z_{kn}^t)b_n]^2 \}.
\]

The bilateral price index $P^*$ that summarizes the overall change in prices going from period $s$ to $t$ is defined as the exponential of the $c_{st}^*$ solution to (A8); i.e., we have:

\[
(A9) P^*(p^1, p^s, q^t, q^s) = \exp[c_{st}^*].
\]

In the remainder of this Appendix, we shall find conditions which ensure that the tests (A4)-(A7) are satisfied.

**Proposition 1:** Suppose that: (i) all models are identical in the two periods so that $S(s) = S(t)$ and $z_{kn}^s = z_{kn}^t = z_{kn}$ for $n = 1, ..., N$ and $k = 1, ..., K$; (ii) the model prices in period $s$ are equal to the corresponding model prices in period $t$ so that $p_k^s = p_k^t = p_k$ for $k = 1, ..., K$; (iii) the model quantities sold in period $s$ are equal to the corresponding sales in period $t$ so that $q_{ks} = q_{kt} = q_k$ for $k = 1, ..., K$; (iv) the model weights are equal across the two periods for each model so that $w_k^s = w_k^t = w_k$ for $k = 1, ..., K$. Under these hypotheses, the identity test (A6) is satisfied.

**Proof:** Under the above hypotheses, the least squares minimization problem (A2) becomes:

\[
(A10) \min_{b^s} \{ \sum_{k \in S(s)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^s)b_n]^2 
+ \sum_{k \in S(t)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^t)b_n]^2 \}.
\]

From the general properties of minimization problems, it can be seen that the following inequality is valid:

\[
(A11) \min_{b^s} \{ \sum_{k \in S(s)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^s)b_n]^2 
+ \sum_{k \in S(t)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^t)b_n]^2 \} 
\geq \min_{b_{st}} \{ \sum_{k \in S(s)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^s)b_n]^2 
+ \sum_{k \in S(t)} w_k [\ln p_k - c_{st} - b_0 - \sum_{n=1}^N f_n(z_{kn}^t)b_n]^2 \}.
\]

Let $b_0^*$, $b_1^*$, ..., $b_N^*$ solve the first minimization problem on the right hand side of (A11). Now look at the second minimization problem on the right hand side of (A11). Obviously the parameters $c_{st}$ and $b_0$ cannot be separately identified so one of them can be set equal to zero; we choose to set $c_{st} = 0$. But after setting $c_{st} = 0$, we see that the second minimization problem is identical to the first minimization problem on the right hand side of (A11), and hence $c_{st}^* = 0$ and $b_0^*$, $b_1^*$, ..., $b_N^*$ solve the second minimization problem. However, $c_{st}^* = 0$ and $b_0^*$, $b_1^*$, ..., $b_N^*$ are feasible for the minimization problem on the left hand side of (A11) and since the objective function evaluated at this feasible solution attains a lower bound, we conclude
that \( c_{st}^* = 0 \) and \( b_0^*, b_1^*, ..., b_N^* \) solves (A10). But \( c_{st}^* = 0 \) implies \( P(p,p,q,q) = \exp[c_{st}^*] = \exp[0] = 1 \), which is the desired result (A6). Q.E.D.

**Proposition 2:** Suppose that the weight functions defined by (A1) are homogeneous of degree zero in the components of the period \( t \) price vector \( p^t \), so that for all \( \lambda > 0 \), \( g_k(z^s, \lambda p^t, q^s, q^t) = g_k(z^s, p^t, q^s, q^t) \) for \( k \in S(s) \) and \( g_k(z^t, \lambda p^t, q^s, q^t) = g_k(z^t, p^t, q^s, q^t) \) for \( k \in S(t) \). Then the hedonic price index \( P(p^s, p^t, q^s, q^t) \) defined by (A3) will satisfy the homogeneity of degree one property (A4).

Proof: Let \( c_{st}^*, b_0^*, b_1^*, ..., b_N^* \) solve the initial minimization problem (A2) before we multiply the period \( t \) price vector by \( \lambda > 0 \). Now consider a new weighted least squares minimization problem where \( p^t \) has been replaced by \( \lambda p^t \). Under our hypotheses, the weights will not be changed by this change in the period \( t \) prices and so the new minimization problem will be:

\[
\begin{align*}
&\text{(A12)} & \sum_{k \in S(s)} w_k \left[ \ln p_k^s - b_0 - \sum_{n=1}^N f_n(z_n^s) b_n \right]^2 \\
&\quad + \sum_{k \in S(t)} w_k \left[ \ln p_k^t + \ln \lambda - c_{st} - b_0 - \sum_{n=1}^N f_n(z_n^t) b_n \right]^2
\end{align*}
\]

\[
\begin{align*}
&\text{(A13)} & \sum_{k \in S(s)} w_k \left[ \ln p_k^s - b_0 - \sum_{n=1}^N f_n(z_n^s) b_n \right]^2 \\
&\quad + \sum_{k \in S(t)} w_k \left[ \ln p_k^t - c_{st} - b_0 - \sum_{n=1}^N f_n(z_n^t) b_n \right]^2
\end{align*}
\]

where the new \( c_{st}^* \) variable is defined as follows:

\[
\begin{align*}
&\text{(A14)} & c_{st}^* = c_{st} - \ln \lambda.
\end{align*}
\]

Denote the solution to (A13) as \( c_{st}^{**}, b_0^{**}, b_1^{**}, ..., b_N^{**} \). However, it can be seen that the solution to (A13) is exactly the same as the solution to the initial problem, (A2). Hence \( c_{st}^{**} = c_{st}^* \), and the \( c_{st} \) solution to (A12), which we denote by \( c_{st}^{**} \), satisfies (A14):

\[
\begin{align*}
&\text{(A15)} & c_{st}^* = c_{st}^{**} = c_{st}^{**} - \ln \lambda \quad \text{or} \\
&\text{(A16)} & c_{st}^{**} = c_{st}^* + \ln \lambda.
\end{align*}
\]

Hence

\[
\begin{align*}
&\text{(A17)} & P(p^s, \lambda p^t, q^s, q^t) = \exp[c_{st}^{**}] \\
&\quad = \exp[c_{st}^* + \ln \lambda] \quad \text{using (A16)} \\
&\quad = \lambda \exp[c_{st}^*] \quad \text{using definition (A3)} \\
&\quad = \lambda P(p^s, p^t, q^s, q^t)
\end{align*}
\]

which establishes the desired result (A4). Q.E.D.

**Proposition 3:** Suppose that the weight functions defined by (A1) are homogeneous of degree zero in the components of the period \( s \) price vector \( p^s \), so that for all \( \lambda > 0 \), \( g_k(z^s, \lambda p^s, p^t, q^s, q^t) = g_k(z^s, p^s, p^t, q^s, q^t) \) for \( k \in S(s) \) and \( g_k(z^t, \lambda p^s, p^t, q^s, q^t) = g_k(z^t, p^s, p^t, q^s, q^t) \) for \( k \in S(t) \). Then the hedonic price index \( P(p^s, p^t, q^s, q^t) \) defined by (A3) will satisfy the homogeneity of degree minus one property (A5).

---

66 We assume that the period \( t \) quantity vector \( q^t \) remains the same if the period \( t \) prices change from \( p^t \) to \( \lambda p^t \).

67 We assume that the period \( s \) quantity vector \( q^s \) remains the same if the period \( s \) prices change from \( p^s \) to \( \lambda p^s \).
Proof: Let \( c_{st}^*, b_0^*, b_1^*, ..., b_N^* \) solve the initial minimization problem (A2) before we multiply the period \( s \) price vector by \( \lambda > 0 \). Now consider a new weighted least squares minimization problem where \( p_s^i \) has been replaced by \( \lambda p_s^i \). Under our hypotheses, the weights will not be changed by this change in the period \( s \) prices and so the new minimization problem will be:

\[
A(18) \quad \text{min} \ b_0^s \text{ and } c \{ \sum_{k \in S(s)} w_k^s [\ln p_k^s + \ln \lambda - b_0^s - \sum_{n=1}^N f_n(z_{kn}) b_n] \}^2 + \sum_{k \in S(t)} w_k^t [\ln p_k^t - c_{st}^s - b_0^t - \sum_{n=1}^N f_n(z_{kn}) b_n] \}
\]

\[
A(19) = \text{min} \ b_0^s \text{ and } c \{ \sum_{k \in S(s)} w_k^s [\ln p_k^s - b_0^s - \sum_{n=1}^N f_n(z_{kn}) b_n] \}^2 + \sum_{k \in S(t)} w_k^t [\ln p_k^t - c_{st}^* - b_0^t - \sum_{n=1}^N f_n(z_{kn}) b_n] \}
\]

where the new \( b_0^s \) and \( c_{st}^s \) variables are defined as follows:

\[
A(20) \quad b_0^s = b_0^s - \ln \lambda; \quad c_{st}^s = c_{st}^s + \ln \lambda.
\]

Denote the solution to (A19) as \( c_{st}^{**}, b_0^{**}, b_1^{**}, ..., b_N^{**} \). However, it can be seen that the solution to (A19) is exactly the same as the solution to the initial problem, (A2). Hence \( c_{st}^{**} = c_{st}^* \) and \( b_0^{**} = b_0^* \). Thus the \( c_{st} \) solution to (A18), which we denote by \( c_{st}^{**} \), satisfies the following equations, where we have substituted into equations (A20):

\[
A(21) \quad b_0^* = b_0^{**} - \ln \lambda; \quad c_{st}^* = c_{st}^{**} + \ln \lambda.
\]

Using the second equation in (A21), we have:

\[
A(22) \quad c_{st}^{**} = c_{st}^* - \ln \lambda.
\]

Hence

\[
A(23) \quad P(p_s^s, p_t^s, q_s^s, q_t^s) = \exp[c_{st}^{**}] = \exp[c_{st}^* - \ln \lambda] = \lambda^{-1} \exp[c_{st}^*] = \lambda^{-1} P(p_s^s, p_t^s, q_s^s, q_t^s) \quad \text{using (A22)}
\]

which establishes the desired result (A5). Q.E.D.

Note that in both Propositions 2 and 3, it is not necessary that the weights \( w_k^s \) and \( w_k^t \) sum to one for each period \( s \) and \( t \).

**Proposition 4:** The bilateral hedonic price index which measures price change going from period \( s \) to \( t \), \( P(p_s^s, p_t^s, q_s^s, q_t^s) \) defined by (A3), and the bilateral hedonic price index which measures price change going from period \( t \) to \( s \), \( P^*(p_t^s, p_s^s, q_t^s, q_s^s) \) defined by (A9), satisfy the time reversal test (A7).

Proof: As usual, denote the solution to (A2) as \( c_{st}^*, b_0^*, b_1^*, ..., b_N^* \). The minimization problem, which corresponds to reversing the ordering of the two periods, is (A24) below and it has the solution \( c_{st}^{**}, b_0^{**}, b_1^{**}, ..., b_N^{**} \):

\[
A(24) \quad \text{min} \ b_0^s \text{ and } c \{ \sum_{k \in S(s)} w_k^s [\ln p_k^s - c_{st}^s - b_0^s - \sum_{n=1}^N f_n(z_{kn}) b_n] \}^2 + \sum_{k \in S(t)} w_k^t [\ln p_k^t - b_0^t - \sum_{n=1}^N f_n(z_{kn}) b_n] \}
\]
\[(A25) = \min b's \text{ and } c \{ \sum_{k \in S(s)} w_k^s [\ln p_k^s - b_0^s - \sum_{n=1}^N f_n^s(z_{kn})b_n^s]^2 \\
+ \sum_{k \in S(t)} w_k^t [\ln p_k^t - c_{st}^t - b_0^t - \sum_{n=1}^N f_n^t(z_{kn})b_n^t]^2 \} \]

where we have defined the new variables \(b_0'\) and \(c_{st}'\) in terms of the old variables \(b_0\) and \(c_{st}\) as follows:

\[(A26) b_0' = b_0 + c_{st} \; ; \; c_{st}' = -c_{st} \; .\]

Denote the solution to \((A25)\) as \(c_{st}^*, b_0^*, b_1^*, \ldots, b_N^*\). However, it can be seen that the solution to \((A25)\) is exactly the same as the solution to the initial problem, \((A2)\). Hence \(c_{st}^* = c_{st}^*\) and \(b_0^* = b_0\). Thus the \(c_{st}\) solution to \((A24)\), which we denoted by \(c_{st}^*\), satisfies the following equations, where we have substituted into equations \((A26)\):

\[(A27) b_0^* = b_0^* + c_{st}^* \; ; \; c_{st}^* = -c_{st}^* \; .\]

Using definition \((A9)\), we have:

\[(A28) P^*(p^s,p^t,q^s,q^t) = \exp[c_{st}^*] = \exp[-c_{st}^*] = 1/\exp[c_{st}^*] = 1/P(p^s,p^t,q^s,q^t) \; \text{using definition \((A3)\)} \]

which establishes the desired result \((A7)\). Q.E.D.

**Proposition 5:** Let \(c_{st}^*, b_0^*, b_1^*, \ldots, b_N^*\) denote the solution to the weighted least squares problem \((A2)\). Then \(c_{st}^*\), which is the logarithm of the bilateral hedonic price index \(P(p^s,p^t,q^s,q^t)\) defined by \((A3)\), satisfies the following equation:

\[(A29) \sum_{k \in S(s)} w_k^s c_{st}^* = \sum_{k \in S(t)} w_k^t \ln p_k^t - \sum_{k \in S(s)} w_k^s \ln p_k^s - \sum_{k \in S(s)} w_k^s b_0^* - \sum_{k \in S(s)} w_k^s \sum_{n=1}^N f_n^s(z_{kn})b_n^s - \sum_{k \in S(s)} w_k^s \ln p_k^s - b_0^* - \sum_{k \in S(s)} w_k^s \sum_{n=1}^N f_n^s(z_{kn})b_n^s \; .\]

Proof: The solution \(c_{st}^*, b_0^*, b_1^*, \ldots, b_N^*\) to the minimization problem \((A2)\) can be obtained by applying least squares to the following linear regression model:

\[(A30) (w_k^s)^{1/2} \ln p_k^s = (w_k^s)^{1/2} [b_0^s + \sum_{n=1}^N f_n^s(z_{kn})b_n^s] + e_k^s \; ; \; k \in S(s); (w_k^t)^{1/2} \ln p_k^t = (w_k^t)^{1/2} [c_{st}^* + b_0^t + \sum_{n=1}^N f_n^t(z_{kn})b_n^t] + e_k^t \; ; \; k \in S(t).\]

We have inserted the optimal least squares estimators, \(c_{st}^*, b_0^*, b_1^*, \ldots, b_N^*\), into equations \((A30)\) so that we can use these equations to define the least squares residuals \(e_k^s\) and \(e_k^t\) for the period \(s\) and \(t\) observations. It is well known that the column vector of these residuals is orthogonal to the columns of the \(X\) matrix, which correspond to the exogenous variables on the right hand side of equations \((A30)\). These orthogonality relations applied to the columns that correspond to the constant term \(b_0\) and the time dummy variable \(c_{st}\) give us the following 2 equations:

---

68 The two equations in \((A29)\) are generalizations of a similar formula derived by Triplett and McDonald (1977; 150) in the unweighted context. This unweighted formula was also used by Triplett (2000; 39). The technique of proof used in this Proposition was used in section 4 of Diewert (2001).
Corollary 5.1: If the models are identical during the two periods and the weights are also identical across periods for the same model, then the hedonic price index $P(p_s, p_t, q_s, q_t)$ defined by (A3) is equal to a weighted geometric mean of the model price relatives, where the weights are proportional to the common model weights, $w_k^s = w_k^t = w_k$.

Proof: Under the stated hypotheses, the last 4 sets of terms on the right hand side of (A29) sum to zero and hence the logarithm of $P(p_s, p_t, q_s, q_t)$ is equal to:

\[(A35) \ c_{st}^* = \sum_{k=1}^{K} w_k \frac{p_k^t}{p_k^s} \frac{1}{\sum_{j=1}^{K} w_j} \]

which establishes the desired result. Q.E.D.

Corollary 5.2: If the models are identical during the two periods and the weight for model $k$ is chosen to be the arithmetic average of the expenditure shares on the model for the two periods, $(1/2)s_{k^s} + (1/2)s_{k^t}$, then the hedonic price index $P(p_s, p_t, q_s, q_t)$ defined by (A3) is equal to the Törnqvist (1936) price index.

Proof: Apply (A35) with $w_k \equiv (1/2)s_{k^s} + (1/2)s_{k^t}$. Q.E.D.

We conclude this Appendix by noting that the second equation in (A29) has a nice interpretation in the light of our discussion of quality adjusted price relatives in section 4 above: it can be seen that $c_{st}^*$ is equal to a weighted sum of the logarithms of the quality adjusted prices of the models sold in period $t$ less another weighted sum of the logarithms of the quality adjusted prices of the models sold in period $s$. If the weights sum to unity in each period, then the two weighted sums become weighted averages of the logarithms of quality adjusted prices.
References


An Almost Ideal Hedonic Price Index for Televisions

Heymerik van der Grient and Jan de Haan

Statistics Netherlands

Abstract: The time dummy approach, where the hedonic model is estimated by (a certain type of) weighted least squares regression, implicitly leads to a generalised Törnqvist price index. We present a decomposition of this almost ideal index and investigate when it reduces to the matched-model Törnqvist index. The analysis is illustrated using scanner data for televisions during the period 1999-2001. The matched-model price index appears to approximate the generalised version quite well.

Keywords: consumer price index, hedonic regression, scanner data.

1. Introduction

Statistics Netherlands is conducting research into the treatment of durable goods in the Consumer Price Index (CPI); see Van der Grient (2001) for the project plan. One of the aims is to investigate how to incorporate scanner data (obtained directly from retailers) for household appliances and electronic equipment. The central question is whether hedonic quality adjustments are necessary or a matched-model approach suffices. This question can only be answered properly if we had a benchmark at our disposal. By this we mean a price index for the population – so that sampling aspects do not come into play – which has been constructed using a sound methodology, including hedonic quality adjustments. This paper addresses the compilation of such a benchmark index for televisions.

One of the methods to compile a hedonic price index is the time dummy approach. We will consider two periods, denoted by 0 and 1. The semi-log (log-linear) hedonic model reads

\[ \ln(p_i^t) = \alpha + \delta D_i^t + \sum_{k=1}^{K} \beta_k x_{ik} + \epsilon_i^t \quad (t=0,1), \]  

where \( p_i^t \) denotes the price of model (good) \( i \) in period \( t \), \( x_{ik} \) its \( k \)-th characteristic \( (k=1,\ldots,K) \), \( \beta_k \) the corresponding parameter, \( \epsilon_i^t \) an error term with an expected value of zero, and \( D_i^t \) a dummy variable that takes on the value of 1 when the observation comes from period 1 (otherwise 0). Notice that the \( \beta_k \)'s are assumed constant over time. This restriction does not pose a serious problem if 0 and 1 are adjacent, short periods. Model (1) will be estimated by least squares regression on the pooled data from both periods. The estimated or

\[1\] The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.
predicted price of \( i \) is \( \hat{p}_i^t \), the residual
\( u_i^t = \ln(p_i^t) - \ln(\hat{p}_i^t) = \ln(p_i^t / \hat{p}_i^t) \). The antilogarithm (exponential) of the time dummy coefficient \( \delta \) yields a quality-adjusted price change.

In most empirical studies Ordinary Least Squares (OLS) has been used to estimate (1). Silver (2002) criticises this approach. Because this particular hedonic technique directly estimates a price index, the observations should be weighted according to their economic importance. Thus, a certain type of Weighted Least Squares (WLS) is needed. Diewert (2002) suggests to use the average expenditure shares as weights in the WLS-procedure for a model that has been sold in both periods. If there happen to be no new or disappearing models, in which case there are only matched models, the WLS-estimator coincides with the Törnqvist index. For new and disappearing models, which are by definition available in one period only, the expenditure shares relating to that period should serve as weights. According to Diewert the resulting WLS-estimator provides a generalisation of the Törnqvist index. Since the Törnqvist index belongs to the class of superlative indexes, we take it for granted that his choice for the weights, when applied to the whole population, leads to the desired benchmark price index.

In this paper we analyse the difference between the generalised Törnqvist index and its matched-model counterpart. The second one is much easier to construct because product characteristics do not have to be collected. Section 2 derives a general expression for the WLS time dummy price index and formulates two requirements which the weights must satisfy. By doing so it will immediately become clear why Diewert’s choice is the most logical one. Next we show under what assumption the generalised Törnqvist price index can be written as the product of the matched-model Törnqvist price index and a factor containing the average residuals of new and disappearing models. Section 3 illustrates the decomposition using scanner data for televisions. Section 4 concludes.

2. The WLS time dummy price index

2.1 A decomposition

We will start by introducing some notation. \( U^t \) denotes the population of models belonging to a certain product group in period \( t \) (\( t=0,1 \)). The matched population is defined as \( U_M = U^0 \cap U^1 \); it is assumed throughout that \( U_M \neq \emptyset \). \( U_D \) is that part of \( U^0 \) that is no longer available in period 1 (the disappearing part), and \( U_N \) is that part of \( U^1 \) that did not exist in period 0 (the new part). Regression model (1) will be estimated by WLS, in which \( w_i^0 \) and \( w_i^1 \) denote the weights for \( i \in U^0 \) and \( i \in U^1 \), respectively. Since the regression residuals sum to zero in each period, the following relation holds:

\[
\prod_{i \in U_M} \left( \frac{p_i^0}{\hat{p}_i^0} \right)^{w_i^0} \prod_{i \in U_D} \left( \frac{p_i^0}{\hat{p}_i^0} \right)^{w_i^0} \prod_{i \in U_M} \left( \frac{p_i^1}{\hat{p}_i^1} \right)^{w_i^1} \prod_{i \in U_N} \left( \frac{p_i^1}{\hat{p}_i^1} \right)^{w_i^1} = 1. \quad (2)
\]

After some rearranging and substitution of \( \hat{p}_i^1 / \hat{p}_i^0 = \exp(\hat{\delta}_{WLS}) \) for \( i \in U_M \) we obtain the following general expression for the WLS time dummy price index:

\[
P_{TD} = \exp(\hat{\delta}_{WLS}) = \prod_{i \in U_M} \left( \frac{p_i^1}{\hat{p}_i^0} \right)^{w_i^0} \prod_{i \in U_D} \left( \frac{p_i^0}{\hat{p}_i^0} \right)^{w_i^0} \prod_{i \in U_M} \left( \frac{p_i^1}{\hat{p}_i^1} \right)^{w_i^1} \prod_{i \in U_N} \left( \frac{p_i^0}{\hat{p}_i^0} \right)^{w_i^1}. \quad (3)
\]
where \( w_{t}^{j} = \sum_{i \in U_{t}} w_{i}^{j} \).

A first requirement is that the resulting index should be wholly based on observed prices when there are no new or disappearing models. Quality changes do not occur in that case (although the quality mix usually does change because of changes in the quantities sold), and we want the outcome to be independent of the chosen set of characteristics. Yet the resulting price index is in a certain sense model dependent: it will automatically have a geometric structure due to the log-linear specification of the regression model, in which the weights play a crucial role. If \( U_{t} = U_{t-1} = U_{M} \), then (3) reduces to

\[
P_{TD} = \prod_{i \in U_{M}} \left( \frac{P_{i}^{t}}{P_{i}^{t-1}} \right) \left( \frac{w_{i}^{t}}{w_{i}^{t-1}} \right) = \prod_{i \in U_{M}} \left( \frac{P_{i}^{t}}{P_{i}^{t-1}} \right) \prod_{i \in U_{M}} \left[ \exp(u_{i}^{t}) \right] \left( \frac{w_{i}^{t}}{w_{i}^{t-1}} \right).
\]

The second factor of (3') contains the period 0 residuals. This factor depends on the choice of the characteristics and usually differs from 1. Hence, the first requirement will generally not be met. We therefore impose the restriction \( w_{i}^{t} = w_{i}^{t} = w_{i} \) for \( i \in U_{M} \), which assures that the time dummy index equals the matched-model index \( P_{M} = \prod_{i \in U_{M}} \left( \frac{P_{i}^{t}}{P_{i}^{t-1}} \right) w_{i}^{t} / w_{i}^{t-1} \).

Summarising: the use of time-dependent weights for the matched models, whether that be (relative) expenditures or quantities, must be avoided. No weighting at all of the observations, i.e. the use of OLS, does satisfy the requirement.

A second requirement is that the resulting index can be defended on the grounds of index number theory. This implies that the price relatives of the individual models must somehow be weighted by expenditure shares, and not for example by (relative) quantities. The use of OLS, which leads to the unweighted geometric or Jevons price index, is no longer an option. For the matched models we have expenditure shares of both period 0 and period 1, denoted by \( s_{t}^{0} \) and \( s_{t}^{1} \). For reasons of symmetry their unweighted average is a natural choice, which also meets the first requirement. This is precisely what Diewert (2002) proposes. For disappearing and new models we only have expenditure shares for period 0 and period 1, respectively. Substitution of \( w_{i}^{0} = w_{i}^{1} = (s_{i}^{0} + s_{i}^{1}) / 2 \) for \( i \in U_{M} \), \( w_{i}^{0} = s_{i}^{0} \) for \( i \in U_{P} \) and \( w_{i}^{1} = s_{i}^{1} \) for \( i \in U_{N} \) in the general WLS expression (3) yields a decomposition of Diewert’s generalised Törnqvist price index:

\[
P_{GT} = \left( P_{MGL} \right)^{\frac{s_{M}^{1}}{s_{M}^{0} + s_{M}^{1}}} \left( P_{MGP} \right)^{\frac{s_{P}^{1}}{s_{P}^{0} + s_{P}^{1}}} \left[ \exp(u_{N}) \right]^{\frac{2(1-t_{1}^{0})}{s_{M}^{0} + s_{M}^{1}}} \left[ \exp(-u_{D}) \right]^{\frac{2(1-t_{1}^{0})}{s_{P}^{0} + s_{P}^{1}}} ,
\]

where \( s_{t}^{t} \) denotes the period \( t \) share (\( t = 0,1 \)) of the matched population in the total expenditure; \( u_{N} = \sum_{i \in U_{N}} s_{i}^{1} u_{i}^{1} / \sum_{i \in U_{M}} s_{i}^{1} \) and \( u_{D} = \sum_{i \in U_{P}} s_{i}^{0} u_{i}^{0} / \sum_{i \in U_{P}} s_{i}^{0} \) are the expenditure-weighted average residuals of new and disappearing models. The other part of (4) is the weighted geometric average of two matched-model price indexes: the geometric Laspeyres index \( P_{MGL} = \prod_{i \in U_{M}} \left( \frac{P_{i}^{t}}{P_{i}^{t-1}} \right) ^{s_{M}^{1}} \) and what we will call the geometric Paasche index \( P_{MGP} = \prod_{i \in U_{M}} \left( \frac{P_{i}^{t}}{P_{i}^{t-1}} \right) ^{s_{M}^{1}} \), with \( s_{i}^{t} = s_{i}^{1} / s_{M}^{1} \) being the share of \( i \) in the period \( t \) expenditures on all matched models.
2.2 Two assumptions

The matched-model Törnqvist price index is defined as the unweighted geometric average of $P_{MGL}$ and $P_{MGP}$, i.e. $P_{MT} = (P_{MGL} P_{MGP})^{1/2}$. We need two assumptions to achieve that this index coincides with the generalised Törnqvist index:

**Assumption i:** $s_M^1 = s_M^0$.

Replacing $s_M^1$ by $s_M^0$ in decomposition (4) gives

$$P_{GT(i)} = P_{MT} \left[ \exp(\bar{u}_N - \bar{u}_D) \right]^{1/2}.$$  \hspace{1cm} (5)

**Assumption ii:** $\bar{u}_N = \bar{u}_D$.

Under this assumption $P_{GT(i)}$ reduces to $P_{MT}$. Note that it is not the variability of the quality-adjusted prices of unmatched models that matters. Rather it is systematic effects – giving rise to positive or negative average residuals – that matter. Whether ii) holds depends on the prevailing market circumstances. Under perfect competition the ‘law of one quality-adjusted price’ predicts that the average residuals $\bar{u}_D$ en $\bar{u}_N$ will be (close to) zero. However, firms can try to reduce competition, for example through market segmentation. Also, consumers may not be completely informed or they may be faced with search costs. When such market imperfections exist it is conceivable that new and disappearing models have unusual prices, i.e. prices that are relatively high or low given their characteristics, so that assumption ii) becomes questionable. One possibility is that manufacturers or retailers manage to enforce ‘hidden’ price increases during the introduction of new models, so that $\bar{u}_N > 0$. Furthermore, old models may be sold at unusually low prices (‘sales’), in which case $\bar{u}_D < 0$. If both phenomena occur simultaneously, the error caused by assumption ii) might be large. Of course, other situations can arise too.

While systematic differences between quality-adjusted prices are unlikely to occur under perfect competition, the law of one quality-adjusted price will not perfectly hold. At a given point in time there will be random differences. Indeed, one might even view these differences as evidence of heavy competition since firms compete by (temporarily) lowering their prices. Such short-term relative price decreases can cause substantial relative quantity changes. This makes the use of a superlative chain index formula like the Törnqvist particularly meaningful.

2.3 Identification of goods

An important topic, which has not been discussed so far, is how goods should be identified. Almost every consumer durable has a model (or type) number attached to it, which is usually available in scanner data sets. This number is in principle unique and can serve as an identification key. The set of matched models, for example, can be found by matching model numbers (in a certain outlet type) in adjacent periods. However, changes in model numbers do not necessarily imply real quality changes. A matched-model approach is problematic if model-number changes coincide with price changes. The effects should be visible in the average residuals of new models.

But what constitutes a genuine quality change? The hedonic hypothesis states that a good should be viewed as a specific combination of quality-determining attributes or
characteristics; a correct hedonic model contains all characteristics that determine the performance of the good in question. From an economic point of view, therefore, it would be preferable to identify a good based on its characteristics. Models (model numbers) with identical characteristics — including outlet-specific conditions — yield the same utility to the consumer and are in fact identical goods. A new good should be defined as a combination of characteristics that did not exist before. It may be a new combination of already existing characteristics or the addition of entirely new characteristics. The time dummy method permits adding (dummy variables for) new characteristics, which is an advantage compared to most other hedonic methods, in particular the hedonic imputation approach. It is most likely that the number of new and disappearing goods measured by their characteristics is smaller than the number of new and disappearing models, and the impact of hedonic modelling will diminish accordingly.

Despite its attractiveness we will not identify products based on their characteristics and restrict ourselves to identification by model numbers. This is because our study aims at finding practical solutions for implementing scanner data obtainable directly from Dutch retailers themselves, who are generally unable or unwilling to provide Statistics Netherlands with product characteristics.

3. An illustration on scanner data for televisions

3.1 The data and the hedonic model

For research purposes Statistics Netherlands has bought scanner data from market research company GfK for a limited number of durable goods. They refer to 18 two-month periods for the years 1999-2001. The data sets contain (per type of outlet) for all models many characteristics, unit values, and quantities sold. Expenditures and expenditure shares can easily be computed. One of the product groups is television sets. Van der Grient (2002) describes this group and compares the scanner data with the CPI data. We did some data cleaning and deleted models that were only sold incidentally. Still, as many as 24,773 observations (unit values per model per period per outlet type) remained in the televisions data set, covering 97.8% of total sales.

Van der Grient (2003) gives a thorough description of the selection of the hedonic model utilized in our study. The model incorporates 71 explanatory variables, most of which are dummies. These include 29 technical characteristics (e.g. size of screen and availability of teletext), 38 brand names, and 4 outlet types. The $R^2$ was 0.96 in case of OLS, and 0.97 in case of WLS. Almost all coefficients were statistically significant and their signs accorded with a priori expectations. The assumption of constant coefficients during adjacent periods has been tested and was not rejected. As a matter of fact the coefficients for the most important technical characteristics appeared to be extremely stable over the entire three-year period.

The unit values indicated that the average price of television sets increased by 18%. This increase has nothing to do with inflation, however, but is due to the appearance on the market

---

2 The unit value (average transaction price) over all relevant models sold in a particular outlet type then is the natural price concept.
3 Regression results are available from the authors. Statistics Netherlands investigated earlier the possibility of estimating a hedonic model for TVs using data from a 'price comparison website' called Consumerdesk (Van der Grient and Oei, 2001).
of models with higher-valued characteristics. As will be shown below, quality-adjusted prices decreased by some 17%.

3.2 Results

Figure 1 clearly demonstrates the importance of weighting in the regression. Three different time dummy price indexes are shown: one using OLS, the second using WLS with time-specific expenditure shares serving as weights and the third using Diewert’s WLS proposal. The OLS version approximates the generalised Törnqvist index surprisingly well, although both series seem to diverge at the end of the period studied. The use of time-specific expenditure shares as weights, on the other hand, gives rise to upward bias.

The official CPI, which is presented in figure 1 for comparison, also overstates the generalised Törnqvist index. One should bear in mind that the CPI methodology and the data used differ a lot from the time dummy indexes. The CPI is based on a sample of only 20 television models and is a Laspeyres-type index; the (constant) weights reflect the 1995 expenditure pattern. Moreover, explicit quality adjustments have not been carried out; a matched-model approach has been applied.

Figure 1: Price change of televisions; three time dummy indexes and the CPI

Table 1 decomposes the generalised Törnqvist price index according to expression (4). ‘Factor L/P’ denotes the weighted average of the matched-model Laspeyres and Paasche indexes and ‘Factor residuals’ denotes the remaining part of the right-hand side of (4). The latter appears to be negligible.

Table 2 shows for all adjacent periods the expenditure shares of the matched models in both period t-1 and period t to infer to what extent assumption $i) s_M^1 = s_M^0$ holds. The matched

---

4 All price index numbers can be found in appendix 1.
5 Statistics Netherlands does not publish a separate CPI for television sets. We computed this index ourselves from the official CPI data.
expenditure share $s_M^t$ in period $t$ is on average about 95% and slightly smaller than the share $s_M^{t-1}$ in the preceding period, which is on average about 98%. So while the shares are not exactly equal, the differences are very small.

Table 1: $P_{GT}$ and factors of decomposition (4)

<table>
<thead>
<tr>
<th>Year</th>
<th>$P_{GT}$</th>
<th>Factor L/P</th>
<th>Factor residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>199902</td>
<td>100.00</td>
<td>100.00</td>
<td>1.000</td>
</tr>
<tr>
<td>199904</td>
<td>97.26</td>
<td>96.95</td>
<td>1.003</td>
</tr>
<tr>
<td>199906</td>
<td>94.97</td>
<td>94.87</td>
<td>1.001</td>
</tr>
<tr>
<td>199908</td>
<td>93.24</td>
<td>93.24</td>
<td>1.000</td>
</tr>
<tr>
<td>199910</td>
<td>90.86</td>
<td>91.10</td>
<td>0.997</td>
</tr>
<tr>
<td>199912</td>
<td>89.62</td>
<td>89.79</td>
<td>0.998</td>
</tr>
<tr>
<td>200002</td>
<td>88.43</td>
<td>88.72</td>
<td>0.997</td>
</tr>
<tr>
<td>200004</td>
<td>87.81</td>
<td>87.63</td>
<td>1.002</td>
</tr>
<tr>
<td>200006</td>
<td>87.16</td>
<td>87.31</td>
<td>0.998</td>
</tr>
<tr>
<td>200008</td>
<td>86.44</td>
<td>86.62</td>
<td>0.998</td>
</tr>
<tr>
<td>200010</td>
<td>85.31</td>
<td>85.76</td>
<td>0.995</td>
</tr>
<tr>
<td>200012</td>
<td>85.17</td>
<td>85.40</td>
<td>0.997</td>
</tr>
<tr>
<td>200102</td>
<td>85.55</td>
<td>85.70</td>
<td>0.998</td>
</tr>
<tr>
<td>200104</td>
<td>85.72</td>
<td>85.68</td>
<td>1.000</td>
</tr>
<tr>
<td>200106</td>
<td>85.75</td>
<td>85.74</td>
<td>1.000</td>
</tr>
<tr>
<td>200108</td>
<td>84.90</td>
<td>84.75</td>
<td>1.002</td>
</tr>
<tr>
<td>200110</td>
<td>83.99</td>
<td>84.08</td>
<td>0.999</td>
</tr>
<tr>
<td>200112</td>
<td>83.25</td>
<td>83.42</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Table 2: Expenditure shares of the matched models in period $t-1$ and period $t$

<table>
<thead>
<tr>
<th>Year</th>
<th>Year</th>
<th>$s_M^{t-1}$</th>
<th>$s_M^t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>199902</td>
<td>199904</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>199904</td>
<td>199906</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>199906</td>
<td>199908</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>199908</td>
<td>199910</td>
<td>0.99</td>
<td>0.93</td>
</tr>
<tr>
<td>199910</td>
<td>199912</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>199912</td>
<td>200002</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>200002</td>
<td>200004</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>200004</td>
<td>200006</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>200006</td>
<td>200008</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>200008</td>
<td>200010</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>200010</td>
<td>200012</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>200012</td>
<td>200102</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>200102</td>
<td>200104</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>200104</td>
<td>200106</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>200106</td>
<td>200108</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>200108</td>
<td>200110</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>200110</td>
<td>200112</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 2 suggests that the bias caused by assumption i) will be small as well. This is indeed confirmed by table 3, which contains both factors from the right-hand side of decomposition (5). $P_{MT}$ exceeds the weighted average of the geometric Laspeyres and Paasche indexes ('Factor L/P' in table 1) by no more than 0.02 index points in December 2001. The impact on

---

6 The geometric Paasche index turns out to be higher than the geometric Laspeyres index (86.25 against 80.73 in December 2001).
‘Factor residuals’ is somewhat larger and in the opposite direction. On balance $P_{GT(t)}$ exhibits a downward bias of 0.28 index points with respect to the generalised Törnqvist index.

#### Table 3: $P_{GT(t)}$ and factors of decomposition (5)

<table>
<thead>
<tr>
<th>$P_{GT(t)}$</th>
<th>$P_{MT}$</th>
<th>Factor residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>199902</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>199904</td>
<td>97.01</td>
<td>96.95</td>
</tr>
<tr>
<td>199906</td>
<td>94.73</td>
<td>94.87</td>
</tr>
<tr>
<td>199908</td>
<td>92.99</td>
<td>93.25</td>
</tr>
<tr>
<td>199910</td>
<td>90.75</td>
<td>91.10</td>
</tr>
<tr>
<td>199912</td>
<td>89.44</td>
<td>89.79</td>
</tr>
<tr>
<td>200002</td>
<td>88.24</td>
<td>88.72</td>
</tr>
<tr>
<td>200004</td>
<td>87.61</td>
<td>87.63</td>
</tr>
<tr>
<td>200006</td>
<td>87.07</td>
<td>87.31</td>
</tr>
<tr>
<td>200008</td>
<td>86.34</td>
<td>86.63</td>
</tr>
<tr>
<td>200100</td>
<td>85.27</td>
<td>85.77</td>
</tr>
<tr>
<td>200102</td>
<td>85.01</td>
<td>85.42</td>
</tr>
<tr>
<td>200104</td>
<td>85.34</td>
<td>85.71</td>
</tr>
<tr>
<td>200106</td>
<td>85.48</td>
<td>85.70</td>
</tr>
<tr>
<td>200108</td>
<td>85.52</td>
<td>85.76</td>
</tr>
<tr>
<td>200110</td>
<td>84.56</td>
<td>84.77</td>
</tr>
<tr>
<td>200112</td>
<td>83.68</td>
<td>84.10</td>
</tr>
<tr>
<td>200112</td>
<td>82.97</td>
<td>83.44</td>
</tr>
</tbody>
</table>

Table 4 shows to what extent assumption ii) $\bar{u}_N = \bar{u}_D$ holds. The average residuals fluctuate around zero but the positive values predominate. In general, the average residuals of the disappearing models exceed those of the new models. Assumption ii) more than offsets the downward bias caused by assumption i). In December 2001 the matched-model Törnqvist index (see table 3) overstates the generalised version by 0.19 index points.

#### Table 4: Weighted average residuals of new and disappearing models (assumption ii)

<table>
<thead>
<tr>
<th>$t-1$</th>
<th>$t$</th>
<th>$\bar{u}_N$</th>
<th>$\bar{u}_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>199902</td>
<td>199904</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>199904</td>
<td>199906</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>199906</td>
<td>199908</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>199908</td>
<td>199910</td>
<td>-0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>199910</td>
<td>199912</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>199912</td>
<td>200002</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>200002</td>
<td>200004</td>
<td>0.12</td>
<td>-0.04</td>
</tr>
<tr>
<td>200004</td>
<td>200006</td>
<td>-0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>200006</td>
<td>200008</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>200008</td>
<td>200010</td>
<td>-0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>200010</td>
<td>200012</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>200012</td>
<td>200102</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>200102</td>
<td>200104</td>
<td>0.05</td>
<td>-0.02</td>
</tr>
<tr>
<td>200104</td>
<td>200106</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>200106</td>
<td>200108</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>200108</td>
<td>200110</td>
<td>-0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>200110</td>
<td>200112</td>
<td>-0.04</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Both assumptions are thus approximately valid. This is again illustrated in Figure 2, which depicts the period-to-period changes of the generalised Törnqvist index and its approximations $P_{GT(t)}$ en $P_{MT}$.
4. Conclusion

The matched-model Törnqvist price index for televisions approximated Diewert's (2002) generalised Törnqvist index well during the years 1999-2001. There are two reasons for this: the expenditure share of the matched models was very high, and the average residuals of the new and disappearing models were small. This means that explicit quality adjustments were not necessary so that the matched-model Törnqvist index sufficed. This result seems fairly robust. In any case the average residuals did not display a striking pattern, not even during the turning point in the business cycle around 2000. In this respect it is a pity the data did not cover the conversion of the Dutch guilder to the euro in January 2002. That would have provided us with an additional check on the robustness of the results, since it is conceivable that markets were less transparent at that time.

Whether implementation in the CPI of scanner data for televisions – let alone for all electronic equipment – without hedonic quality adjustments is acceptable, remains debatable. We will repeat our exercise for video recorders, refrigerators, washing machines, and PCs. PCs in particular are a suitable test case, because the dynamics are far greater here than for the other commodity groups. However, even if we would find similar results, there is still no guarantee that the same will hold in the future. Note that Statistics Netherlands brings in new televisions almost always by initially assigning them the aggregate price change of the other models; see Van der Grint (2002). This method is known as linking, or bridged overlap in HICP jargon. We refer to it as a matched-model procedure, albeit not exactly a good one. Explicit quality adjustments for television sets are thus not made in current Dutch practice either, and an almost automatic procedure is applied.

Although Diewert's (2002) WLS time dummy index can be defended on the grounds of index number theory, we nevertheless make some critical remarks. First, the data may contain outliers that have a substantial influence on the outcome. This concerns especially outliers which receive large weights in the WLS procedure. Silver (2002) suggests to delete such influential outliers. We did not follow his suggestion as that might have affected the
population character of the index too heavily. We did clean the database for obvious errors, though.

Second, weighting of observations is only necessary in order to interpret the antilog of the time dummy coefficient as a certain type of price index. In econometrics WLS is recommended if there is evidence of heteroskedasticity (non-constant variance of the errors). Diewert (2002) notes that the log-linear model is preferred over its linear counterpart, because heteroskedasticity is less likely in that case. WLS might now introduce heteroskedasticity. As such, this is not problematic: as long as the weights are exogenous, the estimators remain unbiased. The weights (expenditure shares) in Diewert's procedure are not exogenous, however, since they contain the explained variable (price), which makes them stochastic variables. This introduces a bias in the WLS estimator.

Third, there seems to be an inconsistency in the time dummy approach, irrespective of the weighting procedure. The hedonic regression model should hold for all goods (TV models), both for matched models and new and disappearing ones. Is it then not a bit strange if the residuals of new and disappearing models exhibited a systematic pattern? De Haan (2003) proposes to incorporate dummy variables for those models. A drawback is that the antilogarithm of the time dummy coefficient can no longer be interpreted as a price index, so that one has to use the hedonic imputation method. Since this approach is controversial we did not try it.

So there is something to be said against the WLS procedure used on econometric grounds, the more so because the time dummy approach assumes constancy of the parameters. The generalised Törnqvist index may thus not be the final answer, but the advantages surely counterbalance the problems. The qualification 'almost ideal' seems in place – and that is what the title of our paper refers to.


Grient, H.A. van der, and M.H. Oei (2001), Possibilities for Hedonic Index Numbers for Televisions Based on Inte


Appendix 1: Price index numbers for televisions (199902=100)

<table>
<thead>
<tr>
<th>Year</th>
<th>( P_{GT} )</th>
<th>( P_{GT(i)} )</th>
<th>( P_{MT} )</th>
<th>WLS time dummy *)</th>
<th>OLS time dummy</th>
<th>CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>199902</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.0</td>
</tr>
<tr>
<td>199904</td>
<td>97.26</td>
<td>97.01</td>
<td>96.95</td>
<td>97.42</td>
<td>97.57</td>
<td>97.5</td>
</tr>
<tr>
<td>199906</td>
<td>94.97</td>
<td>94.73</td>
<td>94.87</td>
<td>95.42</td>
<td>95.78</td>
<td>96.2</td>
</tr>
<tr>
<td>199908</td>
<td>93.24</td>
<td>92.99</td>
<td>93.25</td>
<td>93.84</td>
<td>94.42</td>
<td>94.5</td>
</tr>
<tr>
<td>199910</td>
<td>90.86</td>
<td>90.75</td>
<td>91.10</td>
<td>91.82</td>
<td>92.70</td>
<td>93.1</td>
</tr>
<tr>
<td>199912</td>
<td>89.62</td>
<td>89.44</td>
<td>89.79</td>
<td>91.02</td>
<td>90.87</td>
<td>91.7</td>
</tr>
<tr>
<td>200002</td>
<td>88.43</td>
<td>88.24</td>
<td>88.72</td>
<td>89.90</td>
<td>89.94</td>
<td>90.9</td>
</tr>
<tr>
<td>200004</td>
<td>87.81</td>
<td>87.61</td>
<td>87.63</td>
<td>89.53</td>
<td>88.91</td>
<td>90.8</td>
</tr>
<tr>
<td>200006</td>
<td>87.16</td>
<td>87.07</td>
<td>87.31</td>
<td>88.30</td>
<td>97.89</td>
<td>90.8</td>
</tr>
<tr>
<td>200008</td>
<td>86.44</td>
<td>86.34</td>
<td>86.63</td>
<td>87.82</td>
<td>86.90</td>
<td>90.0</td>
</tr>
<tr>
<td>200010</td>
<td>85.31</td>
<td>85.27</td>
<td>85.77</td>
<td>87.18</td>
<td>85.45</td>
<td>89.6</td>
</tr>
<tr>
<td>200012</td>
<td>85.17</td>
<td>85.01</td>
<td>85.42</td>
<td>87.51</td>
<td>85.64</td>
<td>89.0</td>
</tr>
<tr>
<td>200102</td>
<td>85.55</td>
<td>85.34</td>
<td>85.71</td>
<td>88.13</td>
<td>85.22</td>
<td>89.3</td>
</tr>
<tr>
<td>200104</td>
<td>85.72</td>
<td>85.48</td>
<td>85.70</td>
<td>87.99</td>
<td>85.87</td>
<td>89.5</td>
</tr>
<tr>
<td>200106</td>
<td>85.75</td>
<td>85.52</td>
<td>85.76</td>
<td>88.48</td>
<td>85.95</td>
<td>89.9</td>
</tr>
<tr>
<td>200108</td>
<td>84.90</td>
<td>84.56</td>
<td>84.77</td>
<td>87.75</td>
<td>85.57</td>
<td>88.3</td>
</tr>
<tr>
<td>200110</td>
<td>83.99</td>
<td>83.68</td>
<td>84.10</td>
<td>86.98</td>
<td>85.21</td>
<td>86.9</td>
</tr>
<tr>
<td>200112</td>
<td>83.25</td>
<td>82.97</td>
<td>83.44</td>
<td>86.45</td>
<td>85.45</td>
<td>86.6</td>
</tr>
</tbody>
</table>

*) Time-specific expenditure shares as weights.
Comparison of Hedonic Indices Compiled using Different Types of Weights

Masato Okamoto
National Statistics Center IAI, Japan

1. Summary / Conclusion

I did this paper on the issue what type of weights should be used for the hedonic regression, being stimulated by Diewert’s paper (2003) presented in this meeting. I recompiled chained hedonic indices using expenditure weights and the same formula with Okamoto and Sato (2001) in order to compare with those of quantity weights. In addition, the index compilation using intermediate types of weights is also studied for the Box-Cox test. Conclusions drawn from this empirical study basically corresponds to three suggestions made by Diewert (2003), however there seem points to be noticed in practice as follows:

• From the point of view of the index number theory, expenditure weights are seemingly preferable if the log of the model price is used as the dependent variable. In the case of PCs or digital still cameras, hedonic index compiled using quantity weights does not much differ from that of expenditure weights, while both indices substantially differ from each other in the case of color TVs.

These results imply that the quantity-weighted or un-weighted hedonic regression may yields price indices sufficiently close to the expenditure-weighted hedonic regression but it should be carefully considered whether price change tend to differ remarkably depending on characteristics and what causes the differences if exists. In the case of color TVs, indices derived from the quantity-weighted and expenditure-weighted hedonic regression do not differ so much if excluding wide-screen models. There is some doubt whether index is really appropriate if including wide-screen models, especially when the expenditure-weighted hedonic regression is applied. In case both indices are substantially different, it should be inquired if a particular reason has effect on the indices, not simply choose the expenditure-weighted hedonic regression. It also should be noted that probably the issue arising in the study of color TVs is not peculiar to the hedonic method in essentials.

• The Box-Cox test was applied to the hedonic regression using intermediate weights and the Box-Cox transformation of the model price as the dependent variable. The test results indicate model price should be log-transformed in all three cases – i.e. PCs, color TVs and digital still cameras.

1 The opinions expressed in this paper are those of the author and do not represent official views of either Statistics Bureau, MPHPT, or National Statistics Center IAI.
The dummy variable adjacent period regression technique (named "two-months method" in this paper) and separate hedonic regression for each of the comparison periods (named "single-month method" in this paper) yield indices close to each other if the log of model price is used as the dependent variable while they yield significantly different indices if model price is not transformed. Probably, single-month method is better when a sufficient number of model prices are available but no significant problem has been found so far about two-months method if the log of model prices is used as the dependent variable.

2. Formula used

2.1 Single-month method, semi-log regression model

(weighted) hedonic regression: \( \log p_{t-1} = \sum \beta_{t,i-1} x_{i,t-1} \), \( \log p_{t} = \sum \beta_{t,i} x_{i,t} \)

index calculation:
\[
\log \hat{I}_{t,t-1} = \sum \beta_{t,i} \bar{x}_{i,t} - \sum \beta_{t,i-1} \bar{x}_{i,t-1}, \quad \log \bar{I}_{t,t-1} = \sum \beta_{t,i} \bar{x}_{i,t} - \sum \beta_{t,i-1} \bar{x}_{i,t-1}
\]
\[
I_{t,t-1} = \sqrt{\frac{\hat{I}_{t,t-1}}{\bar{I}_{t,t-1}}}
\]
\[
S_{T,0} = \prod_{t} I_{t,t-1}
\]

where
- \( \bar{x}_{i,t-1} \): weighted average of characteristic \( i \) in month \( t-1 \)
- \( \bar{x}_{i,t} \): weighted average of characteristic \( i \) in month \( t \)

2.2 Two-month method, semi-log regression model

(weighted) hedonic regression: \( \log p_{t,t-1} = \alpha_{t,t-1} \delta_{t,t-1} + \sum \beta_{t,i/t-1} x_{i,t,t-1} \)

index calculation:
\[
I_{t,t-1} = \exp \alpha_{t,t-1}
\]
\[
D_{T,0} = \prod_{t} I_{t,t-1}
\]

where
- \( \delta_{t,t-1} \): 0 (period \( t-1 \)), 1 (period \( t \))

2.3 Single-month method, non-log regression model

(weighted) hedonic regression: \( p_{t-1} = \sum \beta_{t,i-1} x_{i,t-1} \), \( p_{t} = \sum \beta_{t,i} x_{i,t} \)

index calculation:
\[
\hat{I}_{t,t-1} = \frac{\sum \beta_{t,i} \bar{x}_{i,t}}{\sum \beta_{t,i-1} \bar{x}_{i,t-1}}, \quad \bar{I}_{t,t-1} = \frac{\sum \beta_{t,i} \bar{x}_{i,t}}{\sum \beta_{t,i-1} \bar{x}_{i,t-1}}
\]
\[
I_{t,t-1} = \sqrt{\frac{\hat{I}_{t,t-1}}{\bar{I}_{t,t-1}}}
\]
\[
S_{T,0} = \prod_{t} I_{t,t-1}
\]
2.4 Two-month method, non-log regression model

(weighted) hedonic regression: \( p_{t,\text{ort-1}} = \exp[\alpha_{t,\text{ort-1}} \delta_{t,\text{ort-1}} \left( \sum \beta_{t,\text{ort-1}} x_{t,\text{ort-1}} \right)] \)

index calculation:

\[ I_{t,\text{ort-1}} = \exp \alpha_{t,\text{ort-1}} \]

\[ D_{t,0} = \prod_{t} I_{t,\text{ort-1}} \]

As for formula 2.1 – 2.4, hedonic regressions are performed using expenditure weights \((= pq)\) or quantity weights \((= q)\). An intermediate type of weights \((= p^{1-2}\, q)\) is used for the following formula 2.5.

2.5 Single-month method, (modified) Box-Cox model

(weighted) hedonic regression: \( \frac{p_{t,\text{ort-1}}^{\lambda_{t}-1}}{\lambda_{t}} = \sum \beta_{t,\text{ort-1}} x_{t,\text{ort-1}} \) .......................... (*)

\( \frac{p_{t,\text{ort-1}}^{\lambda_{t}}}{\lambda_{t}} = \sum \beta_{t,x_{t,\text{ort-1}}} x_{t,\text{ort-1}} \) .......................... (**) 

index calculation:

\[ \tilde{I}_{t,\text{ort-1}} = \left( \sum \beta_{t,\text{ort-1}} x_{t,\text{ort-1}} \right)^{1/\lambda_{t}} \]

\[ \tilde{I}_{t,\text{ort-1}} = \left( \sum \beta_{t,\text{ort-1}} x_{t,\text{ort-1}} \right)^{1/\lambda_{t}} \]

\[ I_{t,\text{ort-1}} = \left( \frac{\tilde{I}_{t,\text{ort-1}}}{\tilde{I}_{t,\text{ort-1}}} \right)^{1/\lambda_{t}} \]

\[ B_{t,0} = \prod_{t} I_{t,\text{ort-1}} \]

where

\( x_{t,\text{ort-1}} \): average of \( x_{t,\text{ort-1}} \) with weights \( \{p_{t}^{\lambda_{t}} q_{t}\} \)

\( x_{t,\text{ort-1}} \): average of \( x_{t,\text{ort-1}} \) with weights \( \{p_{t}^{\lambda_{t-1}} q_{t}\} \)

\( x_{t,\text{ort-1}} \): average of \( x_{t,\text{ort-1}} \) with weights \( \{p_{t}^{\lambda_{t}} q_{t-1}\} \)

\( \lambda_{t-1} = \inf_{0 < \lambda \leq 1} \left[ (1 - \lambda) \log \bar{p}_{t-1} + \frac{1}{2} \log \hat{\sigma}_{t,t-1}^{2} - \frac{1}{2} \log \tilde{s}_{t,t-1} \right] \)

\( \lambda_{t} = \inf_{0 < \lambda \leq 1} \left[ (1 - \lambda) \log \bar{p}_{t} + \frac{1}{2} \log \hat{\sigma}_{t,t}^{2} - \frac{1}{2} \log \tilde{s}_{t,t} \right] \)

\( \bar{p}_{t-1}, \bar{p}_{t} \): simple geometric mean of model prices

\( \tilde{s}_{t,t-1}, \tilde{s}_{t,t} \): simple geometric mean of share in weights \( \{= p_{t}^{\lambda_{t}} q_{t}/\sum p_{t}^{\lambda_{t}} q_{t}\} \)

\( \hat{\sigma}_{t,t-1}, \hat{\sigma}_{t,t}^{2} \): estimate of residual variance of regression (*) , (**) respectively when \( \lambda \) is given
The intermediate type of weights \((=p^{1-\lambda}q)\) seems to be appropriate for the (modified) Box-Cox model mentioned above, taking the following superlative index into consideration.

\[
\left[\frac{\sum s_i^t \lambda}{\sum s_i}\right]^{1/\lambda} = \left[\frac{\sum (p_i^t - q_0^t)p_i^t}{\sum (p_i^t - q_0^t)q_0^t}\right]^{1/\lambda}
\]

3. Results

Chained hedonic indices are presented in the following Chart 1, Table 1–2 and Table A1–A3 in the appendix.

Chained hedonic indices derived from the “expenditure-weighted semi-log hedonic regression – single month method” are close to the corresponding indices derived from the “quantity-weighted non-log hedonic regression – single month method” as expected in the case of PCs and color TVs (see in Table 1 below, Chart 1 and Table A1–A2). In the case of digital still cameras, difference between the two indices seems relatively large (see in Table 1 below and Table A3). Although the exact reason is not yet found at present (maybe, some explanatory variables such as pixels and zoom should be appropriately transformed), it may be natural that the difference of the two indices become relatively large since the Box-Cox test shows the appropriate \(\lambda\) is always almost zero as shown in Table A3.

Table 1: Chained hedonic indices compiled using different types of weights

<table>
<thead>
<tr>
<th></th>
<th>expenditure weights</th>
<th>quantity weights</th>
<th>intermediate weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log single-month</td>
<td>semi-log single-month</td>
<td>Box-Cox single-month</td>
</tr>
<tr>
<td>two-months</td>
<td>(Jan. 1995 = 1.00)</td>
<td>(Jan. 2000 = 1.00)</td>
<td></td>
</tr>
<tr>
<td>PCs</td>
<td>0.07560</td>
<td>0.54641</td>
<td>0.61140</td>
</tr>
<tr>
<td>Color TVs</td>
<td>0.07213</td>
<td>0.54592</td>
<td>0.60934</td>
</tr>
<tr>
<td>Digital Cameras</td>
<td>0.07116</td>
<td>0.61913</td>
<td>0.62504</td>
</tr>
<tr>
<td>TVs</td>
<td>0.06879</td>
<td>0.61913</td>
<td>0.62543</td>
</tr>
<tr>
<td></td>
<td>0.07276</td>
<td>0.53697</td>
<td>0.63477</td>
</tr>
<tr>
<td></td>
<td>0.07233</td>
<td>0.55495</td>
<td>0.60909</td>
</tr>
</tbody>
</table>

Chained hedonic indices derived from the “quantity-weighted semi-log hedonic regression – single month” are relatively close to “expenditure-weighted semi-log hedonic regression – single month method” in the case of PCs and digital still cameras. However, the two indices substantially differ from each other in the case of color TVs (see Table 1 above, Chart 1 and Table A2). This is mainly due to remarkably different price changes of wide-screen models. If excluding wide-screen models from the hedonic regression and the index compilation, the two indices become relatively close to each other (see Table 2 below and Table A2s). Okamoto and Sato (2001) showed the share of wide-screen models in sales has shrunk, whereas prices of wide-screen models tend to fall down much faster than the basic models (see Chart 2 – 3). "Wide screen" may be a relatively long-lasting but essentially a temporary "boom".
Mysteriously, it took long time for TV sellers to notice consumers do not value “wide screen” in reality. I guess there are certain grounds for a negative view about inclusion of wide-screen models in the hedonic regression and the index compilation.

Table 2: Chained hedonic indices compiled using different types of weights (in the case of color TVs, excluding wide-screen models)

<table>
<thead>
<tr>
<th></th>
<th>expenditure weights</th>
<th>quantity weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log</td>
<td>semi-log</td>
</tr>
<tr>
<td>two-months</td>
<td>single-month</td>
<td>two-months</td>
</tr>
<tr>
<td>color TVs</td>
<td>Jan. 1995 = 1.00</td>
<td>0.67162</td>
</tr>
<tr>
<td></td>
<td>Jun. 1999</td>
<td>0.67379</td>
</tr>
</tbody>
</table>

In all three categories, the Box-Cox test shows the appropriate $\lambda$ is closer to zero rather than one. Chained hedonic indices derived from the hedonic regression using intermediate weights and the appropriate Box-Cox transformation of model prices (“(modified) Box-Cox model – single month method”) are close to the corresponding indices derived from the “expenditure-weighted semi-log hedonic regression – single month method” (see Table 1, Chart 1 and Table A1 – A3). Thus, semi-log regression model is probably preferable. Further study is required to obtain the appropriate transformation of some explanatory variables.

Chart 1: Chained hedonic indices compiled using different types of weights (in the case of color TVs, 1995=1.00, single month method)

Similar to the quantity-weighted semi-log hedonic regression, the “expenditure-weighted semi-log hedonic regression – two-months method” yields indices close to the corresponding
indices obtained from the “expenditure-weighted semi-log hedonic regression – single-month method” in all three categories, while the “quantity-weighted non-log hedonic regression – two-months method” yields indices significantly different from the corresponding indices obtained from the “quantity-weighted non-log hedonic regression – single-month method” (see Table 1 and Table A1 – A3). No specific problem has been found so far about two-month method if the log of model price is used as the dependent variable.

Chart 2: Price indices for the specific types of color TVs estimated from the hedonic regressions

Note. The above chart is drawn using the relevant figures in Table 8. of Okamoto and Sato (2001)

Chart 3-1: Wide screen TVs, share in the total sales by size
Chart 3-2: Wide screen TVs, share in the total number of units sold

Note. Charts 3-1, 3-2 corresponds to Chart.9-1., 9-2. of Okamoto and Sato (2001) respectively.

References


<table>
<thead>
<tr>
<th></th>
<th>expenditure weight</th>
<th>quantity weight</th>
<th>intermediate weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log single-month</td>
<td>non-log single-month</td>
<td>semi-log single-month</td>
</tr>
<tr>
<td></td>
<td>two-months</td>
<td>single-month</td>
<td>two-months</td>
</tr>
<tr>
<td>95 1</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>2</td>
<td>0.91303</td>
<td>0.91282</td>
<td>0.92409</td>
</tr>
<tr>
<td>3</td>
<td>0.87544</td>
<td>0.87497</td>
<td>0.88327</td>
</tr>
<tr>
<td>4</td>
<td>0.85986</td>
<td>0.85953</td>
<td>0.85947</td>
</tr>
<tr>
<td>5</td>
<td>0.83132</td>
<td>0.83033</td>
<td>0.82857</td>
</tr>
<tr>
<td>6</td>
<td>0.73739</td>
<td>0.72776</td>
<td>0.72244</td>
</tr>
<tr>
<td>7</td>
<td>0.68188</td>
<td>0.67170</td>
<td>0.66644</td>
</tr>
<tr>
<td>8</td>
<td>0.64646</td>
<td>0.63662</td>
<td>0.64166</td>
</tr>
<tr>
<td>9</td>
<td>0.59883</td>
<td>0.58994</td>
<td>0.59924</td>
</tr>
<tr>
<td>10</td>
<td>0.54162</td>
<td>0.53355</td>
<td>0.54369</td>
</tr>
<tr>
<td>11</td>
<td>0.48522</td>
<td>0.47480</td>
<td>0.48349</td>
</tr>
<tr>
<td>12</td>
<td>0.45952</td>
<td>0.44985</td>
<td>0.45803</td>
</tr>
</tbody>
</table>

*λ is fixed at 0.5.*
Table A2: Chained hedonic indices compiled using different types of weights (in the case of color TVs, Jan. 1995 = 1.00)

<table>
<thead>
<tr>
<th></th>
<th>expenditure weight</th>
<th>quantity weight</th>
<th>intermediate weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log two-months</td>
<td>single-month</td>
<td>Box-Cox</td>
</tr>
<tr>
<td></td>
<td>single-month</td>
<td></td>
<td>Box-Cox*</td>
</tr>
<tr>
<td>two-months</td>
<td>single-month</td>
<td></td>
<td>single-month</td>
</tr>
<tr>
<td>non-log</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>96</td>
<td>0.99946</td>
<td>0.99942</td>
<td>0.99946</td>
</tr>
<tr>
<td>97</td>
<td>0.98393</td>
<td>0.98395</td>
<td>0.98393</td>
</tr>
<tr>
<td>98</td>
<td>0.97286</td>
<td>0.97287</td>
<td>0.97286</td>
</tr>
<tr>
<td>99</td>
<td>0.95866</td>
<td>0.95904</td>
<td>0.95866</td>
</tr>
<tr>
<td>10</td>
<td>0.93988</td>
<td>0.93963</td>
<td>0.93988</td>
</tr>
<tr>
<td>11</td>
<td>0.92409</td>
<td>0.92470</td>
<td>0.92409</td>
</tr>
<tr>
<td>12</td>
<td>0.91343</td>
<td>0.91460</td>
<td>0.91343</td>
</tr>
<tr>
<td>13</td>
<td>0.89426</td>
<td>0.89481</td>
<td>0.89426</td>
</tr>
<tr>
<td>14</td>
<td>0.86791</td>
<td>0.86644</td>
<td>0.86791</td>
</tr>
<tr>
<td>15</td>
<td>0.84179</td>
<td>0.84237</td>
<td>0.84179</td>
</tr>
<tr>
<td>16</td>
<td>0.82636</td>
<td>0.82692</td>
<td>0.82636</td>
</tr>
</tbody>
</table>

*λ is fixed at 0.5.
Table A2s: Chained hedonic indices compiled using different types of weights (in the case of color TVs, excluding wide-screen models, Jan. 1995 = 1.00)

<table>
<thead>
<tr>
<th></th>
<th>expenditure weight</th>
<th></th>
<th>quantity weight</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log two-months</td>
<td>single-month</td>
<td>semi-log two-months</td>
<td>single-month</td>
</tr>
<tr>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>2</td>
<td>1.00574</td>
<td>1.00572</td>
<td>1.00707</td>
<td>1.00713</td>
</tr>
<tr>
<td>3</td>
<td>0.99179</td>
<td>0.98177</td>
<td>0.98502</td>
<td>0.98601</td>
</tr>
<tr>
<td>4</td>
<td>0.98392</td>
<td>0.98376</td>
<td>0.98975</td>
<td>0.98953</td>
</tr>
<tr>
<td>5</td>
<td>0.97785</td>
<td>0.97772</td>
<td>0.98267</td>
<td>0.98236</td>
</tr>
<tr>
<td>6</td>
<td>0.96037</td>
<td>0.96021</td>
<td>0.96504</td>
<td>0.96467</td>
</tr>
<tr>
<td>7</td>
<td>0.94823</td>
<td>0.94791</td>
<td>0.95474</td>
<td>0.95435</td>
</tr>
<tr>
<td>8</td>
<td>0.94043</td>
<td>0.94015</td>
<td>0.94386</td>
<td>0.94349</td>
</tr>
<tr>
<td>9</td>
<td>0.92055</td>
<td>0.92023</td>
<td>0.92165</td>
<td>0.92121</td>
</tr>
<tr>
<td>10</td>
<td>0.89761</td>
<td>0.89730</td>
<td>0.89854</td>
<td>0.89814</td>
</tr>
<tr>
<td>11</td>
<td>0.86696</td>
<td>0.86667</td>
<td>0.86428</td>
<td>0.86376</td>
</tr>
<tr>
<td>12</td>
<td>0.85464</td>
<td>0.85432</td>
<td>0.85167</td>
<td>0.85115</td>
</tr>
<tr>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.84115</td>
<td>0.84085</td>
<td>0.84397</td>
<td>0.84351</td>
</tr>
<tr>
<td>2</td>
<td>0.84680</td>
<td>0.84653</td>
<td>0.85257</td>
<td>0.85214</td>
</tr>
<tr>
<td>3</td>
<td>0.83322</td>
<td>0.83300</td>
<td>0.83682</td>
<td>0.83644</td>
</tr>
<tr>
<td>4</td>
<td>0.83551</td>
<td>0.83525</td>
<td>0.83976</td>
<td>0.83933</td>
</tr>
<tr>
<td>5</td>
<td>0.82726</td>
<td>0.82675</td>
<td>0.83118</td>
<td>0.82962</td>
</tr>
<tr>
<td>6</td>
<td>0.81145</td>
<td>0.81092</td>
<td>0.81272</td>
<td>0.81133</td>
</tr>
<tr>
<td>7</td>
<td>0.80626</td>
<td>0.80574</td>
<td>0.81109</td>
<td>0.80962</td>
</tr>
<tr>
<td>8</td>
<td>0.79993</td>
<td>0.79944</td>
<td>0.80390</td>
<td>0.80263</td>
</tr>
<tr>
<td>9</td>
<td>0.79314</td>
<td>0.79285</td>
<td>0.79493</td>
<td>0.79366</td>
</tr>
<tr>
<td>10</td>
<td>0.77765</td>
<td>0.77714</td>
<td>0.78224</td>
<td>0.78099</td>
</tr>
<tr>
<td>11</td>
<td>0.76750</td>
<td>0.76706</td>
<td>0.77204</td>
<td>0.77079</td>
</tr>
<tr>
<td>12</td>
<td>0.76079</td>
<td>0.76036</td>
<td>0.76863</td>
<td>0.76736</td>
</tr>
<tr>
<td>97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.75384</td>
<td>0.75343</td>
<td>0.76205</td>
<td>0.76083</td>
</tr>
<tr>
<td>2</td>
<td>0.76011</td>
<td>0.75973</td>
<td>0.76912</td>
<td>0.76797</td>
</tr>
<tr>
<td>3</td>
<td>0.76573</td>
<td>0.76537</td>
<td>0.77470</td>
<td>0.77359</td>
</tr>
<tr>
<td>4</td>
<td>0.77446</td>
<td>0.77412</td>
<td>0.78604</td>
<td>0.78496</td>
</tr>
<tr>
<td>5</td>
<td>0.77180</td>
<td>0.77149</td>
<td>0.78153</td>
<td>0.78046</td>
</tr>
<tr>
<td>6</td>
<td>0.75991</td>
<td>0.75961</td>
<td>0.76923</td>
<td>0.76819</td>
</tr>
<tr>
<td>7</td>
<td>0.75015</td>
<td>0.74982</td>
<td>0.76004</td>
<td>0.75999</td>
</tr>
<tr>
<td>8</td>
<td>0.74960</td>
<td>0.74928</td>
<td>0.76149</td>
<td>0.76051</td>
</tr>
<tr>
<td>9</td>
<td>0.75295</td>
<td>0.75396</td>
<td>0.76471</td>
<td>0.76417</td>
</tr>
<tr>
<td>10</td>
<td>0.74902</td>
<td>0.74986</td>
<td>0.75912</td>
<td>0.75853</td>
</tr>
<tr>
<td>11</td>
<td>0.73565</td>
<td>0.73648</td>
<td>0.74413</td>
<td>0.74362</td>
</tr>
<tr>
<td>12</td>
<td>0.72664</td>
<td>0.72749</td>
<td>0.73807</td>
<td>0.73764</td>
</tr>
<tr>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.71617</td>
<td>0.71700</td>
<td>0.72678</td>
<td>0.72632</td>
</tr>
<tr>
<td>2</td>
<td>0.72967</td>
<td>0.73051</td>
<td>0.74267</td>
<td>0.74226</td>
</tr>
<tr>
<td>3</td>
<td>0.72360</td>
<td>0.72449</td>
<td>0.73473</td>
<td>0.73440</td>
</tr>
<tr>
<td>4</td>
<td>0.72607</td>
<td>0.72693</td>
<td>0.73822</td>
<td>0.73782</td>
</tr>
<tr>
<td>5</td>
<td>0.72343</td>
<td>0.72428</td>
<td>0.73502</td>
<td>0.73462</td>
</tr>
<tr>
<td>6</td>
<td>0.71948</td>
<td>0.72035</td>
<td>0.72948</td>
<td>0.72908</td>
</tr>
<tr>
<td>7</td>
<td>0.71768</td>
<td>0.71856</td>
<td>0.72665</td>
<td>0.72624</td>
</tr>
<tr>
<td>8</td>
<td>0.71788</td>
<td>0.71876</td>
<td>0.72973</td>
<td>0.72935</td>
</tr>
<tr>
<td>9</td>
<td>0.71699</td>
<td>0.71789</td>
<td>0.72889</td>
<td>0.72851</td>
</tr>
<tr>
<td>10</td>
<td>0.71892</td>
<td>0.72090</td>
<td>0.72526</td>
<td>0.72501</td>
</tr>
<tr>
<td>11</td>
<td>0.71988</td>
<td>0.72203</td>
<td>0.72445</td>
<td>0.72429</td>
</tr>
<tr>
<td>12</td>
<td>0.70885</td>
<td>0.71102</td>
<td>0.71714</td>
<td>0.71717</td>
</tr>
<tr>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.69626</td>
<td>0.69839</td>
<td>0.70509</td>
<td>0.70509</td>
</tr>
<tr>
<td>2</td>
<td>0.70490</td>
<td>0.70706</td>
<td>0.71613</td>
<td>0.71625</td>
</tr>
<tr>
<td>3</td>
<td>0.69558</td>
<td>0.69772</td>
<td>0.70696</td>
<td>0.70712</td>
</tr>
<tr>
<td>4</td>
<td>0.69603</td>
<td>0.69825</td>
<td>0.70656</td>
<td>0.70674</td>
</tr>
<tr>
<td>5</td>
<td>0.68759</td>
<td>0.68982</td>
<td>0.70039</td>
<td>0.70059</td>
</tr>
<tr>
<td>6</td>
<td>0.67162</td>
<td>0.67379</td>
<td>0.68209</td>
<td>0.68226</td>
</tr>
<tr>
<td></td>
<td>expenditure weight</td>
<td>non-log</td>
<td>quantity weight</td>
<td>non-log</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>---------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>semi-log two-months single-month</td>
<td>semi-log two-months single-month</td>
<td>semi-log two-months single-month</td>
<td>Box-Cox single-month</td>
</tr>
<tr>
<td>00</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
<td>1.00000</td>
</tr>
<tr>
<td>1</td>
<td>0.97448</td>
<td>0.97434</td>
<td>0.97378</td>
<td>0.97243</td>
</tr>
<tr>
<td>2</td>
<td>0.92672</td>
<td>0.92432</td>
<td>0.94919</td>
<td>0.93415</td>
</tr>
<tr>
<td>3</td>
<td>0.90835</td>
<td>0.90622</td>
<td>0.93712</td>
<td>0.91838</td>
</tr>
<tr>
<td>4</td>
<td>0.89318</td>
<td>0.89099</td>
<td>0.91597</td>
<td>0.90809</td>
</tr>
<tr>
<td>5</td>
<td>0.86411</td>
<td>0.86233</td>
<td>0.88827</td>
<td>0.88349</td>
</tr>
<tr>
<td>6</td>
<td>0.87019</td>
<td>0.86937</td>
<td>0.88665</td>
<td>0.89094</td>
</tr>
<tr>
<td>7</td>
<td>0.87049</td>
<td>0.86973</td>
<td>0.88650</td>
<td>0.88741</td>
</tr>
<tr>
<td>8</td>
<td>0.84062</td>
<td>0.83979</td>
<td>0.84965</td>
<td>0.86025</td>
</tr>
<tr>
<td>9</td>
<td>0.81273</td>
<td>0.81162</td>
<td>0.83579</td>
<td>0.83290</td>
</tr>
<tr>
<td>10</td>
<td>0.78466</td>
<td>0.78367</td>
<td>0.80891</td>
<td>0.81405</td>
</tr>
<tr>
<td>11</td>
<td>0.77121</td>
<td>0.77031</td>
<td>0.79023</td>
<td>0.80345</td>
</tr>
<tr>
<td>12</td>
<td>0.76461</td>
<td>0.76364</td>
<td>0.79949</td>
<td>0.80604</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>expenditure weight</th>
<th>non-log</th>
<th>quantity weight</th>
<th>non-log</th>
<th>intermediate weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>semi-log two-months single-month</td>
<td>semi-log two-months single-month</td>
<td>semi-log two-months single-month</td>
<td>Box-Cox single-month</td>
<td>Box-Cox single-month</td>
</tr>
<tr>
<td>01</td>
<td>0.75725</td>
<td>0.75637</td>
<td>0.77612</td>
<td>0.78903</td>
<td>0.79006</td>
</tr>
<tr>
<td>1</td>
<td>0.73495</td>
<td>0.73410</td>
<td>0.75926</td>
<td>0.76351</td>
<td>0.76499</td>
</tr>
<tr>
<td>2</td>
<td>0.70965</td>
<td>0.70873</td>
<td>0.74484</td>
<td>0.72367</td>
<td>0.72496</td>
</tr>
<tr>
<td>3</td>
<td>0.70399</td>
<td>0.70217</td>
<td>0.73049</td>
<td>0.72326</td>
<td>0.72384</td>
</tr>
<tr>
<td>4</td>
<td>0.69440</td>
<td>0.69174</td>
<td>0.72275</td>
<td>0.71409</td>
<td>0.71461</td>
</tr>
<tr>
<td>5</td>
<td>0.68386</td>
<td>0.68109</td>
<td>0.71793</td>
<td>0.69833</td>
<td>0.69883</td>
</tr>
<tr>
<td>6</td>
<td>0.68605</td>
<td>0.68341</td>
<td>0.71325</td>
<td>0.70661</td>
<td>0.70710</td>
</tr>
<tr>
<td>7</td>
<td>0.69028</td>
<td>0.68760</td>
<td>0.71270</td>
<td>0.71328</td>
<td>0.71376</td>
</tr>
<tr>
<td>8</td>
<td>0.66267</td>
<td>0.65388</td>
<td>0.69267</td>
<td>0.68009</td>
<td>0.68056</td>
</tr>
<tr>
<td>9</td>
<td>0.64242</td>
<td>0.63985</td>
<td>0.67111</td>
<td>0.65751</td>
<td>0.65788</td>
</tr>
<tr>
<td>10</td>
<td>0.62385</td>
<td>0.62584</td>
<td>0.65799</td>
<td>0.64240</td>
<td>0.64275</td>
</tr>
<tr>
<td>11</td>
<td>0.61140</td>
<td>0.60934</td>
<td>0.64984</td>
<td>0.62504</td>
<td>0.62543</td>
</tr>
</tbody>
</table>

* \( \lambda \) is fixed at 0.5.
| Item | Unit | Symbol | Description | Value | Percentage
|------|------|--------|-------------|-------|------------
| Item 1 | Unit 1 | Symbol 1 | Description 1 | Value 1 | Percentage 1 |
| Item 2 | Unit 2 | Symbol 2 | Description 2 | Value 2 | Percentage 2 |
| Item 3 | Unit 3 | Symbol 3 | Description 3 | Value 3 | Percentage 3 |
| Item 4 | Unit 4 | Symbol 4 | Description 4 | Value 4 | Percentage 4 |
| Item 5 | Unit 5 | Symbol 5 | Description 5 | Value 5 | Percentage 5 |
| Item 6 | Unit 6 | Symbol 6 | Description 6 | Value 6 | Percentage 6 |
| Item 7 | Unit 7 | Symbol 7 | Description 7 | Value 7 | Percentage 7 |
| Item 8 | Unit 8 | Symbol 8 | Description 8 | Value 8 | Percentage 8 |
| Item 9 | Unit 9 | Symbol 9 | Description 9 | Value 9 | Percentage 9 |
| Item 10 | Unit 10 | Symbol 10 | Description 10 | Value 10 | Percentage 10 |
| Item 11 | Unit 11 | Symbol 11 | Description 11 | Value 11 | Percentage 11 |
| Item 12 | Unit 12 | Symbol 12 | Description 12 | Value 12 | Percentage 12 |
The Use of Weights in Hedonic Regressions: the Measurement of Quality-Adjusted Price Changes

Mick Silver
Cardiff Business School, Cardiff University, UK

Acknowledgements are due to Erwin Diewert (University of British Columbia) and Verena Dexheimer (Federal Statistical Office, Germany) for helpful comments and the UK Office for National Statistics for their support.

Abstract: Hedonic regressions are used in the measurement of quality-adjusted price changes. Price is regressed on a set of characteristics of a sample of items and the estimated equation is used in a number of ways to undertake adjustments for quality changes. Statistical offices use the coefficients, or predicted prices, to adjust prices for quality when a new item of a different quality replaces an obsolete existing one. This is to ensure the resulting price comparisons are not tainted by quality differences. In academic studies it is more usual to include dummy variables for time in the regression equation, their coefficients providing estimates of the change in price over time having adjusted for quality changes. In these and other approaches it is important that the price (changes) are properly weighted in the calculation. It is axiomatic that the price (change) of an item with relatively low sales should not have the same effect as one with relatively high sales. Yet in spite of the widespread use of hedonic regressions little attention has been given to the proper incorporation of weights, it being simply assumed that a weighted least squares (WLS) estimator is appropriate. First, this paper shows how influence and leverage effects have a distorting effect on the weights under WLS. It second, develops and outlines a number of alternative approaches to the measurement of weighted quality-adjusted price changes and the circumstances under which each is most appropriate.

Keywords: Hedonic Regression; Quality-adjustment; Consumer Price Index; CPI; Leverage Effects.

JEL classification: C43, C81, C82, E31, M21.

1. Introduction

It is axiomatic that a measure of the aggregate change in prices should be weighted. Some items may have much larger sales, in terms of relative quantities or values, and they should be given correspondingly more emphasis in the calculation. However, items vary in quality both for a given time period and over time. There is an extensive literature on the theory and use of hedonic regressions to ‘control’ for such quality differences so that the resulting measures of price changes are unaffected by quality variation (Rosen 1974, Triplett, 1988, Griliches, 1990 and Berndt et al., 1995). One form of such quality-adjusted ‘hedonic indices’ is the time
dummy variable approach as provided by $\delta_i$ in a hedonic regression of the (log of) $m=1,...,M$ prices of items in periods $t=1,...,T$, $p_{mt}$, on their $k=1,...,K$ quality characteristics, $x_{kmt}$, $D_t$, dummy variables and errors $\varepsilon_{mt}$ where:

$$\text{ln } p_{mt} = \alpha + \sum_{k=1}^{K} \beta_k x_{kmt} + \sum_{t=2}^{T} \delta_t D_t + \varepsilon_{mt}$$

(1)

While alternative forms of weighting have received a lot of attention in choosing between index number formulae, little attention has been paid to weighted hedonic estimates, in spite of the increasing need for, and use of, such estimates as a result of rapid changes in the quality of items. An ordinary least squares regression estimator for (1) makes no use of weights, something quite unacceptable by normal index number standards if data on weights are available. The use of a weighted least squares (WLS) estimator is an obvious approach, yet its rationale is not clear. Neither is it clear whether relative quantities or sales values should be used as the weights.

A second motivating factor behind this paper is that the coefficients from hedonic regressions are used by some statistical offices to quality-adjust the prices of non-comparable replacements in the compilation of consumer price indices (CPIs) (for example, for the BLS see Liegey and Shepler, 1999). When one of the sample of matched items is no longer available, a replacement is chosen, and this may be of a different quality. The estimated coefficients $\beta_k$ (or predicted values) from a hedonic regression are used to 'correct' the price of the replaced or original item for the difference in quality, as outlined below. Yet again it is axiomatic that some weighting be applied in the estimation of such coefficients, since again some observations may have relatively high sales and other negligible ones.

This paper shows how a WLS estimator may be inappropriate and outlines and develops a number of approaches to show how, and under what circumstances, different methods should be used for the proper measurement of weighted, hedonic, quality-adjusted indices. Section 2 follows index number theory by asking what the target index should be for a hedonic index. Section 3 is based on Diewert (2002a) and asks which form the weights should take when using a WLS estimator to correspond to a target index. Section 4 develops this framework to show that leverage and influential observations affect the desired weighting structure. Section 5 provides empirical evidence based on scanner data to show that the weights from a WLS estimator may not coincide with those required for the target index formula. Section 6 proposes a number of alternative approaches to WLS, the appropriate measure depending on the extent, if at all, to which data are matched over time. Section 7 concludes.

2. Target indices

First it is necessary to establish target index number formulae to which the outcome from the hedonic analysis should correspond. Following Diewert (1976 and 1978), economic theory supports the use of superlative index number formulae. Diewert (1995 and 2002) show that these formulae can also be justified from an axiomatic, fixed basket, stochastic and Divisia approach and in practice give very similar results. The two most widely adopted superlative indices from respective geometric and arithmetic approaches are the Törnqvist price index,

$$P_T = \prod_{m=1}^{M} \left( \frac{p_{mt}}{p_{m0}} \right)^{s_m^2 + s_m^l}/2$$

(2)

$^1$ Diewert (2002a) argues that the residuals from a logarithmic formulation are less likely to be heteroskedastic.
where \( q_{mt} \) are quantities of item \( m \) in period \( t \) and \( s_m = p_m q_m / \sum_{m=1}^{M} p_m q_m \) \((r=0,1)\) their expenditure share. The Fisher index,

\[
P_F = \left( \frac{\sum_{m=1}^{M} p_m q_m}{\sum_{m=1}^{M} p_m q_m_0} \right)^{1/2} = \left( \frac{\sum_{m=1}^{M} p_m q_m}{\sum_{m=1}^{M} p_m q_m_0} \right)^{1/2} = (P_L P_P)^{1/2}
\]

is the geometric average of the Laspeyres and the Paasche price index. Equations (2) and (3) require price comparisons of matched quality; the same \( m \) items are included in the summation in both periods. If these items are of very different qualities then the matching has to be abandoned and quality differences controlled for using (1). The first concern is how to include appropriate weights in a hedonic formulation such as equation (1) to make it correspond to a target index.

3. Weighted least squares (WLS) estimator and superlative formulae

3.1 An index number approach

Diewert (2002a) makes a number of contributions. First, he argues that observations on a model should be accordingly repeated if they sell more for the estimates to be representative. A WLS estimator with quantity weights is equivalent to an OLS estimator for which observations are replicated according to the number of times they occur. The use of a quantity WLS estimator effectively treats the observations as transactions, one observation repeated for each transaction.

Second, that value weights are preferred to quantity weights: "The problem with quantity weighting is this: it will tend to give too little weight to cheap models that have low amounts of useful characteristics." (Diewert, 2002a: 8). Third, that for WLS estimates of (1), expenditure share weights should be used as opposed to the value of expenditure, to avoid inflation increasing period 1 value weights resulting in possible heteroskedastic residuals. Finally, when a model is present in both periods, the average expenditure shares, \( 1/2(s_{m0} + s_{m1}) \), should be used as weights in the WLS estimator. If only matched models exist in the data, then such estimates will be equivalent to the Törnqvist index (2). If an observation \( m \) is only available in one of the periods, its weight should be \( s_{m0} \) or \( s_{m1} \) accordingly, and the WLS estimator provides a generalisation of the Törnqvist index.

3.2 The econometric approach

WLS estimators are generally advised when the errors from estimated models are heteroskedastic and some reference is necessary to WLS in this context.

A WLS estimator of \( y_i = \beta x_i + u_i \) minimises \( \sum_{i} w_i (y_i - \beta x_i)^2 \)

For \( \hat{\beta} \) to have the smallest conditional error variance \( w_i = 1/\sigma_i^2 \) where \( \text{var}(u_i) = \sigma^2 \). OLS gives equal weight to each observation while WLS give more weight to observations with less conditional variance, thereby decreasing the sampling variance of the OLS estimator. Thus an observation from a distribution with less conditional variance is more informative (in a predictive sense), than an observation from a distribution with a higher conditional variance. A priori we cannot say whether items with larger sales will naturally have a commensurately
smaller variance, such analysis being unhelpful for the motivation for using weights in this context. Moreover, the axiomatic concern of giving more weight to prices (price changes) with higher sales is not necessarily met by a methodology whose focus is on minimising (squared) residuals.

4. Leverage and influence

It is first noted that an OLS vector of \( \beta \) estimates is a weighted average of the individual \( p \) elements, the prices of individual models,

\[
\hat{\beta} = (X^TX)^{-1}X^Tp
\]  

where the matrix \( X \) are the explanatory variable and \((X^TX)^{-1}X^T\) are the implicit weights given to the prices. Equation (4) clearly shows that the \( \hat{\beta} \) estimate is a weighted average of prices, \( p \). Consider also a WLS estimator where the explicit weights are expenditure shares:

\[
\hat{\beta} = (X^TWX)^{-1}X^TWp
\]  

It is apparent from (4) and (5) that outliers with unusual values of \( X \) will have a stronger influence in determining \( \hat{\beta} \), than observations which are one of a group clustered in a small area. Furthermore, (5) shows that the imposition of weights \( W \) allows the influence to vary with \( W \). Thus in normal index number formulae such as (2) and (3), the weights given to price changes are expenditure shares, while in the hedonic framework in (1) the results from an expenditure share weighted hedonic regression will also be determined by the residuals and \( X \) characteristics. An old model of a, for example, washing machine may have unusually poor quality characteristics, and an unusually low price given such characteristics, the relatively high residuals and leverage giving it undue influence in spite of the weights in (5).

There are two concerns. First, for considering the effect of an outlier on the hedonic estimates \( \hat{\beta} \), and second, for estimating the coefficients on the time dummy \( \delta_t \) in equation (1).

4.1 Estimating \( \hat{\beta} \)

Consider the effect of adding a, for simplicity, single unusual observation belonging to a different data generating process to the OLS regression estimate via equation (4). Following Davidson and McKinnon (1993) we compare \( \hat{\beta} \) with \( \hat{\beta}^{(0)} \) where the latter is an estimate of \( \beta \) if OLS was used on a sample omitting the new \( t^{th} \) observation. Distinguish between the leverage of the \( t^{th} \) observation, \( h_t \), and its residual \( \hat{u}_t \). The leverage for observation \( t \) is given by:

\[
h_t = X_t (X^TX)^{-1}X_t^T
\]  

where \( 0 \leq h_t \leq 1 \) 

and the difference between the hedonic coefficients with the \( t^{th} \) observation omitted and included by:

\[
\hat{\beta}^{(0)} - \hat{\beta} = \left( \frac{1}{1-h_t} \right) (X^TX)^{-1}X_t^T \hat{u}_t
\]  

Where \( h_t \) and \( \hat{u}_t \) are relatively large the effect of the \( t^{th} \) observation on at least some of \( \hat{\beta} \) is likely to be substantial. Thus high leverage \( h_t \) only potentially affects \( \hat{\beta} \); it also requires that \( \hat{u}_t \).
is not close to zero. It follows that including the $t^{th}$ observation in the regression affects the fitted value for that observation by:

$$X_t\hat{\beta} = X_t\hat{\beta}^{(i)} + \left(\frac{h_t}{1-h_t}\right)\hat{u}_t$$  \hspace{1cm} (8)

and therefore the influence, or the change in the $t^{th}$ residual by including the $t^{th}$ observation is given by:

$$\left(\frac{h_t}{1-h_t}\right)\hat{u}_t$$  \hspace{1cm} (9)

It can be shown that $h_t$ must on average equal $k/n$ where there are $k$ explanatory variables and $n$ observations. If all $h_t$ were equal to $k/n$ then every observation would have the same leverage. We can thus explore on an empirical basis the values of $\left(\frac{h_t}{1-h_t}\right)\hat{u}_t$ and $h_t$ when estimating hedonic regressions.

### 4.2 Hedonic indices

Hedonic indices can take the form of a hedonic regression with dummy variables included for time as in equation (1). Assume that instead of there being a single new $t$ observations in equations (6) to (9), there are $n_t$ which belong to each period $t=0,1$, the initial observations belonging to period 0. Bear in mind that in the previous analysis if we could identify the unusual observation, a dummy variable which took the value of 1 for the unusual observations and zero otherwise, would be included in the regression, and then $\hat{\beta} = \hat{\beta}^{(i)}$ in (7). Say the period 1 observations are only unusual in the sense that they have a constant shift parameter $\delta_1$ (in logs) applied to them; they are $(\exp(\delta_1)-1)$ percent higher than in period 0. Then the dummy variable hedonic indices in equation (1) will capture the quality-adjusted price change. The appropriate weights for a WLS estimator of $\delta_1$ in (1) to equate with the average expenditure-share weighted Tönnqvist index (2) are thus expenditure-share weights. As in (5), a potential problem still remains with influential observations, as opposed to the simple shift parameter.

A simple illustrative example is provided in Annex 2 on influence effects, while in section 5 we turn to examining whether influence matters in our calculations; whether quantity/expenditure share weights and influence differ.

### 5. Some empirical evidence

It may be argued that older/newer models/brands of a product are likely to have unusual characteristics, prices and thus residuals, and therefore influence over and above their expenditure share weights (Berndt et al., 2001; Pakes, 2002). However, if the hedonic regression controls for the effect on prices of the unusual features the residuals may be low and, via (9), their influence. The alignment of weights and influence is an empirical matter.

Table 1 provides some evidence and shows average and standard deviation leverages (6), $h_t$; residuals $\hat{u}_t$; and influence (9) for successive expenditure share weighted quartile quantity and quartile expenditure shares. If weights and influence diverge, the average influence in each
quartile would differ. The results are from monthly scanner data on washing machines for 1998. Hedonic OLS regressions were run each month for about 550 models of washing machines, over 6,000 observations. The regressions fitted the data relatively well by the usual criteria, the average $R^2=0.80$ with the signs on the coefficients according with expectations. Leverage, (absolute) residuals and influence values were calculated each month and the means and standard deviations calculated for each quartile group, the results being averaged over the 12 months. The data are outlined in the Annex and the monthly regressions and monthly results are available from the authors on request.

Table 1 shows for both quantity and value quartile shares a clear inverse relationship between the relative sales (expenditure) shares and mean influences: observations with higher sales share had, for OLS regressions, less influence. Their residuals and leverage were lower. F-tests for equality of means over quartile groups was rejected at the 1% level (for value shares leverage, residuals and influence respectively: $F=4.45$, $p$-value=0.002; $F=3.90$, $p=0.004$ and $F=4.10$, $p=0.003$ and for quantity shares $F=6.14$, $p$-value=0.000; $F=4.49$, $p=0.001$ and $F=4.93$, $p=0.001$. Influence and expenditure weighting did not coincide.

Table 1: Summary statistics on influence and residuals for weight distribution

<table>
<thead>
<tr>
<th>Sales shares</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean leverage residuals influence</td>
<td>Mean leverage residuals influence</td>
</tr>
<tr>
<td>under Q1: q-share</td>
<td>0.0958 0.1537 -0.0163</td>
<td>0.1215 0.1644 0.0778</td>
</tr>
<tr>
<td>Q1 to median: q-share</td>
<td>0.0773 0.1149 -0.0115</td>
<td>0.1022 0.1008 0.0522</td>
</tr>
<tr>
<td>Median to Q3: q-share</td>
<td>0.0760 0.1072 -0.0081</td>
<td>0.0975 0.0928 0.0415</td>
</tr>
<tr>
<td>Q3 to Q4: q-share</td>
<td>0.0525 0.0966 -0.0061</td>
<td>0.0521 0.0785 0.0125</td>
</tr>
<tr>
<td>All: q-share</td>
<td>0.0756 0.1185 -0.0106</td>
<td>0.1001 0.1183 0.0610</td>
</tr>
<tr>
<td>under Q1: v-share</td>
<td>0.0870 0.1549 -0.0169</td>
<td>0.1055 0.1693 0.0550</td>
</tr>
<tr>
<td>Q1 to median: v-share</td>
<td>0.0847 0.1151 -0.0112</td>
<td>0.1254 0.0990 0.0422</td>
</tr>
<tr>
<td>Median to Q3: v-share</td>
<td>0.0765 0.1064 -0.0095</td>
<td>0.0910 0.0933 0.0500</td>
</tr>
<tr>
<td>Q3 to Q4: v-share</td>
<td>0.0541 0.0977 -0.0062</td>
<td>0.0540 0.0765 0.0110</td>
</tr>
<tr>
<td>All: v-share</td>
<td>0.0756 0.1185 -0.0110</td>
<td>0.1001 0.1183 0.0513</td>
</tr>
</tbody>
</table>

n 6694
$R^2$-adj (mean) 0.8017

6. Proposed alternative approaches

Given concern about the use of WLS for hedonic adjustment a number of alternative approaches are proposed, their suitability being related to the need to be representative of the universe of models being sold.

6.1 Matched models

The desired target measure of price change for a semi-log hedonic function is (2), the superlative geometric Törnqvist index (see Feenstra (1995) who shows that for a Fisher index (3) a linear hedonic form is required). Our trouble with the dummy time variable method (1)
is that the OLS estimator excludes the expenditure share weights and the WLS estimator is influenced by factors other than these weights. Consider the following:

\[
\left[ \prod_{m=1}^{M} \left( \frac{\hat{P}_{m1}}{P_{m0}} \right)^{\gamma_{m1}} \prod_{m=1}^{M} \left( \frac{P_{m1}}{\hat{P}_{m0}} \right)^{\gamma_{m0}} \right]^{1/2}
\]

(10a)

\[
\hat{P}_{m0} = p_{m0} \exp\left( \sum \beta_0 (z_{mk1} - z_{mk0}) \right)
\]

\[
\hat{P}_{m1} = p_{m1} \exp\left( -\sum \beta_1 (z_{mk1} - z_{mk0}) \right)
\]

(10b)

or

\[
\hat{P}_{m0} = \exp[\alpha_0 + \sum \beta_0 z_{mk1}]
\]

\[
\hat{P}_{m1} = \exp[\alpha_1 + \sum \beta_1 z_{mk0}]
\]

(10c)

where \( s_{mt} \) are expenditure shares, \( p_{mt} \) are price, and \( z_{mk} \) are \( k \) characteristics with associated \( \beta_{kt} \) derived from a semi-log hedonic regressions over \( m=1...M \) product varieties (models) for each period \( t=0,1 \).

First, assume the data can be organised into matched pairs of the same model over time, akin to the method used by statistical offices in the compilation of CPIs. Then only (10a) is appropriate. Equation (10a) is equivalent to the target index (2) since \( \hat{P}_{mt} = p \), for \( t=0,1 \). However, the monitoring of matched prices not only requires additional information on the matching, but is also potentially biased since the sample of matched models excludes ‘old’ models available in period 0, but not in 1, and also excludes the ‘new’ ones introduced in period 1, but of course not available in period 0 (see Silver and Heravi (2002) for evidence on the bias).

6.2 Patched matched models using replacement items

Replacement models for obsolete ones in period 1 may be found, though being new, their quality may differ. A quality adjustment to the original or replacement price can be undertaken as in (10b) using changes in the quality of the two models picked up via changes in their characteristics \( (z_{kt} - z_{kt-1}) \) which are multiplied by estimates of their associated marginal hedonic values \( \beta_{kt} \), and summed. Thus (10a) and (10b) are appropriate. Note that \( \hat{P}_{mt} \) corrects the observed prices \( \hat{P}_{mt} \) for changes in the characteristics between the two periods, corresponding to the “explicit quality adjustment” described by Triplett (1990:39); see also Silver and Heravi (2001).

6.3 Patched matched models using imputations

If replacement items are unavailable the missing old or new prices may be imputed using (10c). Thus (10a) and (10c) are appropriate. For example, if the model is in the sample in period 0, but not 1, the imputed price for the model in period 1 is its period 0 characteristics evaluated at period 1 hedonic prices.\(^2\) Equation (10) may be more cumbersome than equation (1), but benefits from employing a pure weighting system, untainted by undue influence.

\(^2\) Diewert (2002a) shows how a further adjustment may be necessary since predicted and actual prices are being compared.
Equation (10a) has been shown to be suitable for matched data, with 10(b) for matched data when an observation is missing in any period, but there is a replacement of different quality and with 10(c) for matched data without replacements using imputations. Yet there remains the question of whether to use a WLS or OLS estimator for the parameters in (10b or c). Two concerns impose: efficiency and representativity. A concern in (10b or c) in using WLS is not the appropriateness of the weighting, for this is captured in (10a), but the efficiency of $\beta$. It was argued in section 3.2 on econometric grounds that an appropriate variable for use as weights in WLS was one that was strongly related to the residual variance. The question of choice between value or quantity weights is thus dictated by the strength of their relationship to the conditional residual variation. Thus price changes in (10a) are expenditure share weighted while the weights for (10b or c) are an empirical matter. However, another criterion is how representative the hedonic regression is for the adjustments being made. Assume there are many new observations for a brand, their characteristics being say superior to other brands. The analysis in section 4 shows that they will have further influence on the estimated $\beta$ over and above that due to expenditure share weights. If, by and large, the new non-comparable replacements are of this new brand, then this ‘bias’ is arguably in the right direction, giving more emphasis to the branded observations whose quality is to be adjusted. In principle, degrees of freedom permitting, separate parameters may be estimated for each brand by way of slope and intercept brand dummies as a better mechanism for ‘tailoring’ the coefficients to the brand-specific prediction.

6.4 Unmatched data, the time dummy hedonic and deletion

A concern with matching, especially over relatively long periods, is that it leads to a deterioration of the sample. For example, Koskimäki and Vartia (2001) attempted to match prices of models of personal computers (PCs) over three two-month periods (spring, summer and fall) using a sample of prices collected as part of standard price collection for the Finish CPI. Of the 83 spring prices only 55 matched pairs could be made with the summer, and then only 16 continued through to the fall (see also Silver and Heravi, 2002). The use of imputations such as (10b) or (10c) on this scale is not desirable. We return to the hedonic time dummy variable approach in equation (1). It allows popular models to be included in the sample in each month without any restriction as to whether they were previously or subsequently included and matched. Bear in mind that the use of replacements in (10b) only allows sampling from the ‘dynamic’ population when a replacement is required, while that of (10c) is based on imputations from the deteriorating original matched sample. Both are unsatisfactory. The hedonic adjustment in (1) allows re-sampling each month from all models and for items such as consumer durables, where there is a high turnover of models of different qualities, this is a highly desirable property. However, as discussed, the implicit weighting system in an expenditure share WLS estimate of (1) may be inappropriate and the OLS one even more so. As such it is proposed that the time dummy hedonic method (1) is only used when there is substantial sample degradation and that influence measures (9) and residuals be computed and observations with relatively low weights and high influence values be deleted, and (1) re-estimated.

6.5 Unmatched data and the superlative hedonic approach

If data are not matched, they are re-sampled each period to be representative of the universe of new and old models, the time dummy method has been proposed with some deletion of observations with undue influence. But the method remains problematic if a model is has a relatively high weight, so cannot be excluded, but its influence is over and above that due to is
weight. One approach is to use calculated influence variable to adjust the weights. An alternative is to use a development to the formulation 10(a) and 10(b) akin to stratified random sampling. First stratification variables are selected which are related to price changes, say screen size and makes for television sets. Then, using (10a), weighted indices are calculated for price changes in these ‘core’ stratum, that is the prices in (10a) are the (geometric) mean prices in each stratum, for example a Sony 21 inch television set. But over time the quality of items in each stratum will change by other ‘non-core’ quality characteristics, such as the possession of stereo, wide-screen etc. Such changes in the quality of the average prices being compared are then controlled for in 10(b). This approach was adopted for weighted price comparisons using scanner data for television sets over time and across countries (Silver and Heravi (2001) and Heravi, Heston and Silver, 2003 respectively).

7. Conclusions

The conclusions for estimating hedonic quality-adjusted price changes when weights are available are:

• Some weighting system is better than none, and a WLS estimator is preferred to an OLS one (section 3).

• The use of expenditure share weights for hedonic time dummy regression estimates is preferred to relative quantities (section 3 and 4).

• WLS estimators may not give the appropriate weights required by the target superlative indices, some observations having undue influence effects thus contributing to the effective weights (section 4).

• The empirical work shows an inverse relationship between expenditure share weights and influence (section 5).

• The use of the proposed hedonically-adjusted predicted prices in an explicitly weighted superlative framework (10a) is preferred, the nature of the adjustment depending on whether replacement models are (10b) or are not (10c) available (section 6).

• When matching and imputations/adjustments are undertaken on a matched sample that has substantially deteriorated, the approach in (10) may be inappropriate. Triplett (2002) has warned of such selectivity or ‘out-of-sample bias’ and Silver and Heravi (2002) have demonstrated its nature and substantial effect. In such circumstances the time dummy variable hedonic approach (1) should be reconsidered. It has the advantage of allowing re-sampling each month from the dynamic universe of items, rather than from just the matched ones or matched/replacement ones. It should, data permitting, be undertaken with a WLS estimator, though observations with relatively low weights and high influence values should be deleted and the regression re-run (section 6).

• Alternatively, price can be re-sampled each period, but explicit weights can be imposed on average price changes of more loosely defined cells of core characteristics or stratum, with quality changes being undertaken within each stratum. This at least imposes an appropriate weighting structure between the strata.
Data Annex for Table 1

The data were monthly observations for four outlet types for 1998 amounting to 7,750 observations (models sold in an outlet type) amounting to 1.5 million transactions worth over £0.5 billion. The price was the unit value of sales of a model in an outlet-type. The characteristics included: (i) Manufacturer (make) – dummy variables for about 20 makes; (ii) type of machine: 5 types – top-loader; twin tub; washing machine (WM); washer dryer (WD) with and without computer; WD with / without condensors; (iii) drying capacity of WD; (iv) height of machines in cms; (v) width; (vi) spin speeds: 5 main - 800rpm, 1000rpm, 1100rpm, 1200rpm and 1400rpm; (vii) water consumption; (viii) load capacity; (ix) energy consumption (kWh per cycle); (x) free standing, built-under and integrated; built-under not integrated; built-in and integrated; (xi) vintage; (xii) outlet-types: multiples, mass merchandisers, independents, multiples.

Annex 2: Illustrative example

Consider Table 2 which contains illustrative data for price ($P$), a quality characteristic $x$ which is on average the same over the two time periods, $t = 0, 1$. Consider the first 18 observations in Table 2. They are generated for periods 0 and 1 respectively from:

$$P_0 = 100 (1.06)^{x_0} \quad \text{and} \quad P_1 = 100 (1.06)^{x_1}$$

i.e. $\ln P_0 = \ln(100) + x_0 \ln(1.06)$ and $\ln P_1 = \ln(105) + x_1 \ln(1.06)$ (11)

Throughout the range of $x$ the price change for those observations to 5%. The last two observations are generated at $x = 5$, observation 19 in period 1, $P_{119} = 115 (1.06)^5$ having a 15% price increase compared with $P_{120} = 100 (1.06)^5$. The first 18 observations are equally weighted accounting for 72% of the weight, the remaining two observations each accounting for 14%.

Table 2: Illustrative Data

<table>
<thead>
<tr>
<th>Observation</th>
<th>Weight</th>
<th>$P$</th>
<th>$x$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.04</td>
<td>106.600</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.04</td>
<td>112.360</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>119.102</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td>126.248</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>133.823</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0.04</td>
<td>141.852</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.04</td>
<td>150.363</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0.04</td>
<td>159.385</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0.04</td>
<td>168.948</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0.04</td>
<td>111.300</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.04</td>
<td>117.978</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.04</td>
<td>125.057</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0.04</td>
<td>132.560</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.04</td>
<td>140.514</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.04</td>
<td>148.945</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>0.04</td>
<td>157.881</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>0.04</td>
<td>167.354</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>0.04</td>
<td>177.395</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>0.14</td>
<td>153.896</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>0.14</td>
<td>133.823</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
Consider the unweighted case. Since the observations are on average of constant quality over time, the aggregate geometric mean price change is \((1.05)^{(1/20)} (1.15)^{20} = 1.0596\), i.e. \(5.96\%\). An OLS regression of \(P\) on \(x\) and \(t\) with \(P_{19}\) evaluated at \(x=5\) finds a coefficient of 0.057886 and a percentage price change of \((e^{0.057886} - 1)100\) of 5.96. [The coefficients from a semi-logarithmic form are not unbiased and require a correction of \(\frac{1}{2}\) (standard error)^2-Goldberger (1968). However, the adjustment in this example is negligible]. The dummy time variable method works. However, if the higher price observation in period 19 had unusual characteristics, say \(x_{19}=9\) or 15 with \(P_{19}\) accordingly evaluated at a 15% increase using (11), then the OLS price change estimate would be wrong at 5.86 and 5.55% respectively. In each of these three cases the influence of \(P_{19}\) is extremely high; at \(x=5\), the influence of observation 19, \(i_{19}=0.0091\), its mean \(\bar{i}=0.0012\) and standard deviation \(s_{i}=0.002\), and for \(x=9\): \(i_{19}, \bar{i}\) and \(s_{i}\) are 0.0178, 0.002 and 0.0039, and for \(x=15\): 0.044, 0.0036 and 0.0096 respectively. \(i_{19}/\bar{i}\) increases from 7.5, to 8.6 and 12.0 as \(x\) takes on more extreme values, from 5,9 to 15.

Table 3: Results from \(P = \beta_0 c + \beta_1 x + \beta_2 T + \beta_3 D + \varepsilon\) illustrative data

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Percentage change: ((e^\beta -1)100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant (c): (\beta_0)</td>
<td>4.60517</td>
<td>0.011623</td>
<td>0</td>
</tr>
<tr>
<td>(x): (\beta_1)</td>
<td>0.058269</td>
<td>0.0019108</td>
<td>6.00</td>
</tr>
<tr>
<td>OLS ((z_{19}=5))</td>
<td>T: (\beta_2)</td>
<td>0.057886</td>
<td>5.96</td>
</tr>
<tr>
<td>((z_{19}=9))</td>
<td>T: (\beta_2)</td>
<td>0.056912</td>
<td>5.86</td>
</tr>
<tr>
<td>((z_{19}=15))</td>
<td>T: (\beta_2)</td>
<td>0.053988</td>
<td>5.55</td>
</tr>
<tr>
<td>OLS ((z_{19}=5))</td>
<td>T: (\beta_2)</td>
<td>0.048789</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>D: (\beta_3)</td>
<td>0.090972</td>
<td>9.52</td>
</tr>
<tr>
<td>OLS ((z_{19}=5))</td>
<td>T: (\beta_2)</td>
<td>0.074261</td>
<td>7.71</td>
</tr>
</tbody>
</table>

The further inclusion of a dummy variable \(D=1\) for observation 19 leads to a regression model with \(R^2bar=1.00\) and coefficients on \(T\) and \(D\) of 0.048789 and 0.090972 (Table 3). The equation is perfectly specified, \(y_m = \hat{y}_m\); the increase in period 1 being \(\exp(0.048789) = 1.05\) excluding observation 19 and \(\exp(0.048789+0.090972) = 1.15\) for observation 19 itself. Yet these results by themselves do not allow us to deduce the quality-adjusted price. Table 2 also includes data on relative expenditure weights, assumed to be constant over time. The Törnqvist index is \((1.05)^{0.72} (1.15)^{0.28} = 1.07709\), i.e. a 7.71% increase. A WLS regression also yields an estimate of 7.71% when \(P_{19}\) is evaluated at \(x=5\), as in Table 1. Again the estimate will change as \(x\) changes, but more so because of its increased weight. For \(x=9\) and \(x=15\), the WLS price increase is only 7.02 and 5.87% respectively.
References


Session 3 - Elementary aggregation, superlative indexes

Chair: Keith Woolford, Australian Bureau of Statistics

Summary of session

Aggregate index outcomes can be significantly influenced by the choice of index number formula at both the elementary aggregate level and at higher levels of aggregation. In general, the magnitude of any differences delivered by the alternative formulae is a function of the dispersion in the rates of price change or of the change in price dispersion.

The first two papers in this session (by Balk and Silver respectively) are concerned primarily with choosing the formula to be used to compile indexes for elementary aggregates. The third paper, presented by Greenlees, focuses on higher level aggregation and describes the calculation of the new superlative consumer price index compiled by the U.S. Bureau of Labor Statistics (the C-CPI-U).

Balk investigates alternative elementary aggregate formula from a sampling perspective. The population for which the elementary aggregate index is to be constructed is classified as either homogeneous or heterogeneous and some practical guidelines for making this distinction are provided. Alternative sample index formulae are then assessed against relevant alternative target population price indexes using feasible alternative sample designs (simple random sampling and sampling with probability proportional to size). For each sample/population index formula pair, Balk categorises any relative bias in the sample formula into a technical component (which trends to zero as the sample size increases) and a structural component (which is independent of sample size).

In developing his sample approach, Balk outlines a logical progression of steps that should be followed by the prices statistician. First, classify the elementary aggregate as being homogeneous or not. Second, select the appropriate population target price index and finally, according to the sample design, choose that sample price index formula which provides the least biased (or approximately unbiased) estimator of the population target price index.

Silver outlines the nature of the relationships between the various formulae in terms of price dispersion and tries to explain the existence of price dispersion and its persistence over time by reference to search cost and menu cost theories and signal extraction models. He then proceeds to investigate the phenomena using a scanner data set containing some 70,000 observations on television sets over 51 months. Hedonic techniques are used to construct 'heterogeneity-controlled' indexes to compare with the more traditional formulae.

The paper provides some interesting insights into price dispersion both within months and over time and concludes by drawing out some practical advice to assist in choosing the elementary aggregate formula and in the selection and replacement of representative items.

Greenlees provides a comprehensive description of the methodology used to calculate the Chained Consumer Price Index for All Urban Consumers (C-CPI-U) that has been published by the U.S. Bureau of Labor Statistics since August 2002. This superlative price index is
constructed using the Tomqvist formula and is published in parallel with the more longstanding CPIs compiled for Urban Wage Earners and Clerical Workers (CPI-W) and All Urban Consumers (CPI-U). The CPI-W and CPI-U are compiled using the more conventional Laspeyres type fixed-base formula and are not generally revisable (at least in response to new weighting information becoming available). By comparison, the C-CPI-U is first published in preliminary form and is subject to two subsequent revisions as more recent expenditure data becomes available.

In addition to describing how the index is calculated, the paper discusses the various issues that were addressed in developing this new measure and comments on differences in outcomes compared to the CPI-U. The latter appear to be rather high (0.8 percent in 2000 and 0.3 percent in 2001).

**Recommendations for statistical agencies**

These three papers provide a number of valuable insights into the issues involved in selecting the most appropriate formula for elementary and higher level index aggregation. Three recommendations could be drawn from this material and the discussions:

1. The choice of elementary aggregate formula is important, especially when the price dispersion within the elementary aggregate is high. It should be made carefully according to the way in which representative products are sampled.

2. One should exercise care in using the relative of arithmetic mean prices formula (Dutot) for other than homogeneous commodities.

3. The method used by the U.S. Bureau of Labor Statistics to calculate the C-CPI-U and the issues addressed in its design are a sound starting point for any national statistical agency contemplating the construction of a superlative index.
Price Indexes for Elementary Aggregates: the Sampling Approach

Bert M. Balk

Statistics Netherlands and Erasmus Research Institute of Management, Erasmus University Rotterdam

The author acknowledges helpful comments from Jörgen Dalén, Erwin Diewert, Jan de Haan, Peter Hill, Paul Knottnerus, and, especially, Mick Silver.

Abstract: At the lowest level of aggregation of a CPI or PPI quantity information is usually unavailable and nothing but matched samples of prices are used for the index computation. Familiar indexes used at this level of aggregation are those of Dutot, Carli, and Jevons. An important, yet often overlooked characteristic of these and similar indexes is that they are sample statistics, whose properties can be studied from the sampling point of view. This paper provides a systematic study of this topic and concludes with a number of recommendations for statistical practice.

Keywords: CPI, PPI, elementary aggregate, price index.

1. Introduction

Mainstream (bilateral) index number theory applies to aggregates consisting of a finite set of commodities. Two basic assumptions are that the set of commodities does not change between the two periods compared, and that all the price and quantity data which are necessary for the computation of an index are available to the statistician. In this paper we are concerned with what to do when the second of these assumptions is or cannot be fulfilled. There are, of course, various kinds of unavailability of data. The situation we will consider in particular in this paper is that nothing but price data are available for a sample of commodities and/or respondents.

Since such a situation materializes at the very first stage of the computation of any official price index, such as a Consumer Price Index (CPI) or a Producer Price Index (PPI), we are dealing here with an issue of great practical significance.

The usual approach to the problem of unavailable quantity data is to consider price indexes which are functions of prices only. The main formulas discussed in the literature and used in practice are

- the ratio of arithmetic average prices (the formula of Dutot),
- the arithmetic average of price relatives (the formula of Carli),
- the geometric average of price relatives = the ratio of geometric average prices (the formula of Jevons).

The views expressed in this paper are those of the author and do not necessarily reflect any policy of Statistics Netherlands.
The appropriateness of these formulas has been studied by various methods. Following the early contribution of Eichhorn and Voeller (1976), Dalén (1992) and Diewert (1995) studied their properties from an axiomatic point of view. Additional insights were obtained by deriving (approximate) numerical relations between these formulas, and by combining these relations with more or less intuitive economic reasoning. Balk's (1994) approach was to see which assumptions would validate these formulas as estimators of true but unknown population price indexes, which by definition are functions of prices and quantities. A summary of the state of affairs, written from the CPI perspective, but easily generalizable to other perspectives, was recently provided by Diewert (2002).

This paper develops the sampling approach. In section 2 it is argued that, although not known to the statistician, all the detailed price and quantity data of the commodities and respondents pertaining to the aggregate under consideration exist in the real world. Section 3 then argues that the first task faced by the statistician is to decide on the nature of the aggregate (homogeneous or heterogeneous) and on the target price index (the unit value index or some superlative or non-superlative price index). Next the sampling design comes into the picture. With help of these two pieces of information, one can judge the various estimators with respect to their performance. This is the topic of section 4, which is on homogeneous aggregates, and sections 5 - 7, which are on heterogeneous aggregates and superlative target price indexes. Section 8 adds to this topic with some micro-economic considerations on the choice of a sample price index. Section 9 discusses the not unimportant case where, for operational reasons, a non-superlative price index was chosen as target. Section 10 surveys the behaviour of the various sample price indexes with respect to the Time Reversal Test, and reviews the (approximate) numerical relations between them. Section 11 summarizes the key results and concludes with practical advice.

2. Setting the stage

The aggregates covered by a CPI or a PPI are usually arranged in the form of a tree-like hierarchy (such as COICOP or NACE). Any aggregate is a set of economic transactions pertaining to a set of commodities. Commodities can be goods or services. Every economic transaction relates to the change of ownership (in the case of a good) or the delivery (in the case of a service) of a specific, well-defined commodity at a particular place and date, and comes with a quantity and a price. The price index for an aggregate is calculated as a weighted average of the price indexes for the subaggregates, the (expenditure or sales) weights and type of average being determined by the index formula. Descent in such a hierarchy is possible as far as available information allows the weights to be decomposed. The lowest level aggregates are called elementary aggregates. They are basically of two types:

- those for which all detailed price and quantity information is available;
- those for which the statistician, considering the operational cost and/or the response burden of getting detailed price and quantity information about all the transactions, decides to make use of a representative sample of commodities and/or respondents.

Any actual CPI or PPI, considered as a function which transforms sample survey data into an index number, can be considered as a stochastic variable, whose expectation ideally equals its population counterpart. The elementary aggregates then serve as strata for the sampling procedure. We are of course particularly interested in strata of the second type.
The practical relevance of studying this topic is large. Since the elementary aggregates form the building blocks of a CPI or a PPI, the choice of an inappropriate formula at this level can have a tremendous impact higher-up in the aggregation tree.

The detailed price and quantity data, although not available to the statistician, nevertheless - at least in principle - exist in the outside world. It is thereby frequently the case that at the respondent level (outlet or firm) already some aggregation of the basic transaction information has been executed, usually in a form that suits the respondent’s financial and/or logistic information system. This could be called the basic information level. This is, however, in no way a naturally given level. One could always ask the respondent to provide more disaggregated information. For instance, instead of monthly data one could ask for weekly data; whenever appropriate, one could ask for regional instead of global data; or, one could ask for data according to a finer commodity classification. The only natural barrier to further disaggregation is the individual transaction level. 2

Thus, conceptually, for all well-defined commodities belonging to a certain elementary aggregate and all relevant respondents there exists information on both the quantity sold and the associated average price (unit value) over a certain time period. Let us try to formalize this somewhat. The basic information – which in principle exists in the outside world – is of the form \( \{(p_n^t, q_n^t), n = 1, \ldots, N\} \) where \( t \) denotes a time period, the elements of the population of (non-void) pairs of well-defined commodities and respondents, henceforth called elements, are labelled from 1 to \( N \), \( p_n^t \) denotes the price and \( q_n^t \) denotes the quantity of element \( n \) at time period \( t \). It may be clear that \( N \) may be a very large number, since even at very low levels of aggregation there can be thousands of elements involved. We repeat that it will be assumed that the population does not change between the time periods considered.

It is assumed that we must compare a later period 1 to an earlier period 0. The later period will be called comparison period and the earlier period base period. The conceptual problem is to split the value change

\[
\frac{\sum_{n=1}^{N} p_n^1 q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^0}
\]

multiplicatively into a price index and a quantity index. This is traditionally called the index number problem. Both indexes should depend only on the prices and quantities of the two periods.

3. Homogeneity or heterogeneity

There is now an important conceptual choice to make. In the statistician’s parlance this is known as the ‘homogeneity or heterogeneity’ issue. Although in the literature lots of words have been devoted to this issue, at the end of the day the whole problem can be reduced to the rather simple looking operational question:

\[
\text{(2) Does it make (economic) sense to add up the quantities } q_n^t \text{ of the elements } n=1,\ldots,N? \]

---

2 See Balk (1994) for a similar approach.
If the answer to this question is ‘yes’, then the appropriate, also called target, price index for the elementary aggregate is the unit value index

\[
P_{\text{U}} = \frac{\sum_{n=1}^{N} p_{n}^{1} q_{n}^{1}}{\sum_{n=1}^{N} p_{n}^{0} q_{n}^{0}},
\]

that is, the average comparison period price divided by the average base period price. The corresponding quantity index is the simple sum or Dutot index

\[
Q_{D} = \sum_{n=1}^{N} q_{n}^{1} / \sum_{n=1}^{N} q_{n}^{0}.
\]

When the quantities are additive, we are obviously dealing with a situation where a certain commodity at the same time is sold or bought at different places at different prices. Put otherwise, we are dealing with pure price differences. These can be caused by market imperfections, such as price discrimination, consumer ignorance, or rationing. Economic theory seems to preclude this possibility since it states that in equilibrium “the law of one price” must hold. In reality, however, market imperfections are the rule rather than the exception. But also physical restrictions can play a role. Although, for instance, the “representative” consumer is assumed to be fully informed about all prices and to have immediate and costless access to all outlets throughout the country, the sheer physical distance between the outlets precludes “real” consumers from exploiting this magical possibility; thus, price differences exist where they, according to a representative-agent-based theory, are not supposed to exist.

If the answer to question (2) is ‘no’, then there are various options available for the target price index. First of all, the axiomatic as well as the economic approach to index number theory leads to the conclusion that the target price index should be some superlative index. According to the recent survey by Diebert (2002: Section 5), three price indexes appear to be particularly relevant. The first is the Törnqvist price index

\[
P_{T} = \prod_{n=1}^{N} \left( \frac{p_{n}^{1}}{p_{n}^{0}} \right)^{s_{n}^{t}} / \sum_{n=1}^{N} p_{n}^{1} q_{n}^{1} / \sum_{n=1}^{N} p_{n}^{0} q_{n}^{0},
\]

where \(s_{n}^{t} = p_{n}^{1} q_{n}^{1} / \sum_{n=1}^{N} p_{n}^{1} q_{n}^{1} \) (\(t=0,1\)) is element \(n\)’s value share in period \(t\). This price index is a weighted geometric average of the price relatives, the weights being average value shares. The corresponding quantity index is defined as

\[
\tilde{Q}_{T} = \left( \sum_{n=1}^{N} p_{n}^{1} q_{n}^{1} / \sum_{n=1}^{N} p_{n}^{0} q_{n}^{0} \right) / P_{T}.
\]

Balk (1998) shows that the unit value index satisfies the conventional axioms for a price index, except the commensurability axiom and the proportionality axiom. However, when the elements are commensurate, the commensurability axiom reduces to \(P(\lambda p^{1}, \lambda q^{1}, \lambda p^{0}, \lambda q^{0}) = P(p^{1}, q^{1}, p^{0}, q^{0})(\lambda > 0)\), which clearly is satisfied.
The second superlative price index is the Fisher index,

\[ P_F = \left( \frac{\sum_{n=1}^{N} p_n^0 q_n^0}{\sum_{n=1}^{N} p_n^1 q_n^0} \right)^{1/2} \left( \frac{\sum_{n=1}^{N} p_n^1 q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^1} \right)^{1/2} = (P_L P_P)^{1/2}, \]

which is the geometric average of the Laspeyres and the Paasche price indexes. In this case the quantity index is given by

\[ Q_F = \left( \frac{\sum_{n=1}^{N} p_n^0 q_n^0}{\sum_{n=1}^{N} p_n^1 q_n^0} \right)^{1/2} \left( \frac{\sum_{n=1}^{N} p_n^1 q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^1} \right)^{1/2} = (Q_L Q_P)^{1/2}, \]

which is the geometric average of the Laspeyres and the Paasche quantity indexes. The third superlative price index is the Walsh index, defined as

\[ P_W = \frac{\sum_{n=1}^{N} p_n^1 (q_n^0 q_n^1)^{1/2}}{\sum_{n=1}^{N} p_n^0 (q_n^0 q_n^1)^{1/2}}, \]

in which case the quantity index is defined by

\[ Q_W = \left( \frac{\sum_{n=1}^{N} p_n^1 q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^0} \right)/P_W. \]

The Walsh price index is a member of the class of so-called basket price indexes, that is, price indexes which compare the cost of a certain basket of quantities in the comparison period to the cost in the base period. The Laspeyres and Paasche price indexes are typical examples: the first employs the base period basket and the second the comparison period basket. The basket of the Walsh price index is an artificial one, namely consisting of the geometric averages of the quantities of the two periods.

Many statistical offices, however, are forced by operational reasons to define a non-superlative price index as target. In general their target appears to have the form of a Lowe price index

\[ P_{Lo} = \frac{\sum_{n=1}^{N} p_n^1 q_n^b}{\sum_{n=1}^{N} p_n^0 q_n^b}, \]

where \( b \) denotes some period prior to the base period 0. The corresponding quantity index is then defined by

\[ Q_{Lo} = \left( \frac{\sum_{n=1}^{N} p_n^1 q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^0} \right)/P_{Lo}. \]

Notice that the five price indexes considered above all satisfy the Time Reversal test, that is, using obvious notation, \( P(p^1, q^1, p^0, q^0) = 1/P(p^0, q^0, p^1, q^1) \).
It could be that the statistician is unable to decide between a simple 'yes' or 'no' reply to (2), that is, he or she finds that for certain subsets of the elementary aggregate \( \{1,\ldots,N\} \) it makes sense to add up the quantities whereas for the remainder this does not make sense. Then the aggregate should be split into subsets to which either the 'yes' or the 'no' answer applies. If this splitting appears to be not feasible then the 'no' answer should take precedence over the 'yes' answer. Thus, conceptually, we have to deal with but two cases.

Having defined the target price (and quantity) index, the statistician must face the basic problem that not all the information on the prices and quantities of the elements is available. The maximum he or she can obtain is information \( \{(p_{ni}, q_{ni}) ; t = 0,1; n \in S\} \) for a sample \( S \subset \{1,\ldots,N\} \). More realistic, however, is the situation where the information set has the form \( \{p_n' ; t = 0,1; n \in S\} \), that is, nothing but a matched sample of prices is available. From this sample information the population price index (or quantity index) must be estimated. This is the point where the theory of finite population sampling will appear to be helpful for obtaining insight into the properties of the various estimators.

At the outset we must notice that in practice the way the sample \( S \) is drawn usually remains hidden in a certain darkness. The main problem is that there is no such thing as a sampling frame. Knowledge about the composition of the elementary aggregate, in the form of an exhaustive listing of all its elements, is usually absent. There is only, more or less ad hoc, evidence available about particular elements belonging or not belonging to this aggregate. In order to use the theory of finite population sampling, however, we must make certain assumptions about the sampling design.

In the remainder of this paper we will consider two scenarios. Each of these is believed to be more or less representative of actual statistical practice. The first scenario assumes that \( S \) is a simple random sample, which means that each element of the population has the same probability of being included in the sample. This so-called (first order) inclusion probability is \( \Pr(n \in S) = \zeta(S)/N \), where \( \zeta(S) \) denotes the (prespecified) sample size.

In the second scenario the more important elements of the population have a correspondingly larger probability of being included in the sample than the less important elements. This will be formalized by assuming that the elements were drawn with probability proportional to size, where size denotes some measure of importance. If the size of element \( n \) is denoted by a positive scalar \( a_n (n=1,\ldots,N) \), then the probability that element \( n \) is included in the sample \( S \) is \( \Pr(n \in S) = \zeta(S)a_n / \sum_{n=1}^{N} a_n \). Without loss of generality, it will be assumed that \( \Pr(n \in S) < 1 \) for \( n=1,\ldots,N \). Notice that in both scenarios \( \sum_{n=1}^{N} \Pr(n \in S) = \zeta(S) \).

The general question we first consider is whether it is at all possible to find an estimator for \( P_u, P_T, P_F \), or \( P_w \) which uses, in addition to base period and comparison period price information, nothing but base period quantity information. Within the first scenario, the answer is obviously 'no', because each of these target indexes depends also on comparison period quantity information, and this information is nowhere involved in the sampling design.

\[\text{Elements for which initially this probability would turn out to be greater than or equal to 1 are selected with certainty and from the remaining set of elements a sample is drawn.}\]
There is no free lunch here. Put otherwise, any estimator which is based on the data set \( \{ p_n^0, q_n^0, p_n^1; n \in S \} \) will necessarily be biased. This conclusion obviously extends to any estimator which is based on the (smaller) data set \( \{ p_n^0, p_n^1; n \in S \} \). Within the second scenario, however, the answer depends on the extent to which comparison period quantity information can be assumed to be included in the size measure.

4. Homogeneous aggregates

Suppose we deal with a homogeneous aggregate. Then the target (or population) price index is the unit value index \( P_u \). If the total base period value \( \sum_{n=1}^{N} p_n^0 q_n^0 \) as well as the total comparison period value \( \sum_{n=1}^{N} p_n^1 q_n^1 \) is known, the obvious route to take – see expression (3) – is to estimate the Dutot quantity index \( Q_D \). A likely candidate is its sample counterpart

\[
(13) \quad \hat{Q}_D = \sum_{n \in S} q_n^1 / \sum_{n \in S} q_n^0 .
\]

Suppose that \( S \) is a simple random sample and recall that the inclusion probabilities are \( \Pr(n \in S) = \zeta(S)/N \), where \( \zeta(S) \) denotes the sample size. Then the expected value of the sample Dutot quantity index is

\[
(14) \quad E(\hat{Q}_D) = E \left( \frac{(1/\zeta(S)) \sum_{n \in S} q_n^1}{(1/\zeta(S)) \sum_{n \in S} q_n^0} \right) = \frac{(1/\zeta(S)) \sum_{n=1}^{N} q_n^1 \Pr(n \in S)}{(1/\zeta(S)) \sum_{n=1}^{N} q_n^0 \Pr(n \in S)} = \frac{(1/N) \sum_{n=1}^{N} q_n^1}{(1/N) \sum_{n=1}^{N} q_n^0} = Q_D .
\]

Expression (14) means that \( \hat{Q}_D \) is an approximately unbiased estimator of the population Dutot quantity index \( Q_D \). The bias\(^5\) is of technical nature and will approach zero when the sample size gets larger.

Consider now the sample Carli quantity index, defined as

\[
(15) \quad \hat{Q}_c = \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{q_n^1}{q_n^0} \right) .
\]

Assume that the elements were drawn with probability proportional to size, whereby the size of element \( n \) is defined as its base period quantity share \( q_n^0 / \sum_{n=1}^{N} q_n^0 \). Thus the probability that element \( n \) is included in the sample is equal to \( \Pr(n \in S) = \zeta(S) q_n^0 / \sum_{n=1}^{N} q_n^0 \). Then the expected value of the sample Carli quantity index is equal to

\[^5\text{The bias is due to the fact that we approximate } E(xy) \text{ by } E(x)E(y). \text{ A Taylor series expansion yields that to the second order } E(xy) = (\leq) E(x)E(y) \text{ if and only if } 1 > (\leq) \rho(x,y)cv(x)/cv(y), \text{ where } \rho(.,.) \text{ is the correlation coefficient and } cv(.) \text{ the coefficient of variation. The bias will typically be positive, whereby its magnitude depends on the value of } E(x)/E(y).\]
(16) \[ E(\hat{Q}_C) = (1/\varsigma(S)) \sum_{n=1}^{N} (q_n^1 / q_n^0) \Pr(n \in S) = \sum_{n=1}^{N} (q_n^1 / \sum_{m=1}^{N} q_m^0)(q_n^1 / q_n^0) = Q_D. \]

Put otherwise, under this sampling design, the sample Carli quantity index is an unbiased estimator of the population Dutot quantity index.

Let the total comparison period value now be unknown to the statistician and consider the sample unit value index

\[ \hat{P}_U = \frac{\sum_{n \in S} p_n^1 q_n^1 / \sum_{n \in S} q_n^1}{\sum_{n \in S} p_n^0 q_n^0 / \sum_{n \in S} q_n^0}. \]

This presupposes that the sample is of the form \( \{ (p_n^t, q_n^t); t = 0, 1; n \in S \} \), that is, for every sampled element one disposes of its value and its quantity in both periods. Then one can show, in much the same way as was done in expression (14), that under simple random sampling the sample unit value index is an approximately unbiased estimator of the target unit value index \( P_U \). Likewise, it appears that

\[ (\sum_{n=1}^{N} p_n^0 q_n^0) \sum_{n \in S} p_n^1 q_n^1 \sum_{n \in S} q_n^1 \]

is an approximately unbiased estimator of the aggregate’s total comparison period value \( \sum_{n=1}^{N} p_n^t q_n^t \). Notice that (18) has the form of a ratio estimator.

Suppose next that only sample prices are available, that is, the sample is of the form \( \{ p_n^t, p_n^0, n \in S \} \), and consider the sample Dutot price index, defined as

\[ \hat{P}_D = \frac{\sum_{n \in S} p_n^1}{\sum_{n \in S} p_n^0} = \frac{(1/\varsigma(S)) \sum_{n \in S} p_n^1}{(1/\varsigma(S)) \sum_{n \in S} p_n^0}. \]

The second part of this expression provides the familiar interpretation of the sample Dutot price index as a ratio of unweighted average sample prices. Under probability proportional to size sampling, whereby again the size of element \( n \) is defined as its base period quantity share, it is easily seen that, apart from a technical bias,

\[ E(\hat{P}_D) = E \left( \frac{\sum_{n \in S} p_n^1}{\sum_{n \in S} p_n^0} \right) \approx \frac{E(1/\varsigma(S)) \sum_{n \in S} p_n^1}{E(1/\varsigma(S)) \sum_{n \in S} p_n^0} = \frac{\sum_{n=1}^{N} p_n^1 q_n^1 / \sum_{n=1}^{N} q_n^1}{\sum_{n=1}^{N} p_n^0 q_n^0 / \sum_{n=1}^{N} q_n^0}. \]

The denominator of the right hand side ratio is the same as the denominator of the unit value index \( P_U \). The numerators, however, differ: the target index uses comparison period quantity

\[ \text{Clearly, taking the average of prices is the counterpart of the adding-up of quantities, i.e. the first makes sense if and only if the second does.} \]
shares as weights whereas \( E(\hat{P}_D) \) yields base period quantity shares as weights. Thus the sample Dutot price index will in general be a biased estimator of the unit value index. The relative bias amounts to

\[
(21) \quad \frac{E(\hat{P}_D)}{P_U} \approx \frac{\sum_{n=1}^{N} P_n^1 q_n^0 / \sum_{n=1}^{N} q_n^0}{\sum_{n=1}^{N} P_n^0 q_n^1 / \sum_{n=1}^{N} q_n^1}.
\]

The relative bias of the sample Dutot price index thus consists of two components, a technical part which vanishes as the sample size gets larger and a structural part which is independent of the sample size. This structural part is given by the right hand side of expression (21). It vanishes if the (relative) quantities in the comparison period are the same as those in the base period, which is unlikely to happen in practice. The result, expressed by (20), was mentioned by Balk (1994; 139) and Diewert (2002; Section 7.4).

5. Heterogeneous aggregates and the Törnqvist price index

We now turn to the more important situation where we deal with a heterogeneous aggregate. Suppose that the Törnqvist price index \( P_T \) is decided to be the target and consider its sample analogue

\[
(22) \quad \hat{P}_T = \prod_{n \in S} \left( \frac{P_n^1}{P_n^0} \right)^{(s_n^1 + s_n^0)/2},
\]

where \( s_n^t = p_n^t q_n^t / \sum_{n \in S} p_n^t q_n^t \) \((t=0,1)\) is element \( n \)'s sample value share. It is clear that the sample must be of the form \( \{ (p_n^t q_n^t, P_n^t); t=0,1; n \in S \} \), that is, for each sample element we dispose of its value and its price in both periods. Under the assumption that each element of the population has the same probability of being included in the sample, namely \( \varsigma(S)/N \), it appears that

\[
(23) \quad E(\ln \hat{P}_T) = \frac{1}{2} E \left( \frac{\sum_{n \in S} P_n^0 q_n^0 \ln(p_n^1 / p_n^0)}{\sum_{n \in S} P_n^0 q_n^0} + \frac{\sum_{n \in S} P_n^1 q_n^1 \ln(p_n^1 / p_n^0)}{\sum_{n \in S} P_n^1 q_n^1} \right)
\]

\[
= \frac{1}{2} \left( \frac{E((1/\varsigma(S))\sum_{n \in S} P_n^0 q_n^0 \ln(p_n^1 / p_n^0))}{E((1/\varsigma(S))\sum_{n \in S} P_n^0 q_n^0)} + \frac{E((1/\varsigma(S))\sum_{n \in S} P_n^1 q_n^1 \ln(p_n^1 / p_n^0))}{E((1/\varsigma(S))\sum_{n \in S} P_n^1 q_n^1)} \right)
\]

\[
= \frac{1}{2} \left( \frac{(1/N)\sum_{n=1}^{N} P_n^0 q_n^0 \ln(p_n^1 / p_n^0)}{(1/N)\sum_{n=1}^{N} P_n^0 q_n^0} + \frac{(1/N)\sum_{n=1}^{N} P_n^1 q_n^1 \ln(p_n^1 / p_n^0)}{(1/N)\sum_{n=1}^{N} P_n^1 q_n^1} \right)
\]

\[
= \ln P_T.
\]
This means that \( \ln \hat{P}_T \) is an approximately unbiased\(^7\) estimator of \( \ln P_T \). Employing Jensen’s Inequality\(^8\), one obtains

\[
(24) \quad E(\hat{P}_T) \geq P_T,
\]

that is, the sample Törnqvist price index has an upward bias relative to its population counterpart. However, this bias is of technical nature and will approach zero when the sample size gets larger.

The previous result critically depends on the availability of sample quantity or value information. Suppose that we cannot obtain these data and consider the sample Jevons price index\(^9\)

\[
(25) \quad \hat{P}_J = \prod_{n \in S} \left( \frac{p_n^1}{p_n^0} \right)^{1/\zeta(S)}.
\]

Under probability proportional to size sampling, whereby the size of element \(n\) is now defined as its base period value share \(s_n^0\), resulting in \(Pr(n \in S) = \zeta(S)s_n^0\), it is easily seen that

\[
(26) \quad E(\ln \hat{P}_J) = E\left(\frac{1}{\zeta(S)} \sum_{n \in S} \ln(p_n^1 / p_n^0)\right) = \sum_{n=1}^N s_n^0 \ln(p_n^1 / p_n^0) = \ln\left(\prod_{n=1}^N \left(\frac{p_n^1}{p_n^0}\right)^{s_n^0}\right).
\]

By employing Jensen’s Inequality, this leads to the result that

\[
(27) \quad E(\hat{P}_J) \geq \prod_{n=1}^N \left(\frac{p_n^1}{p_n^0}\right)^{s_n^0} = P_{GL}.
\]

At the right hand side we have obtained the so-called Geometric Laspeyres population price index, which in general will differ from the Törnqvist population price index. The relative bias of the sample Jevons price index with respect to the Törnqvist population price index is

\[
(28) \quad \frac{E(\hat{P}_J)}{P_T} \geq \prod_{n=1}^N \left(\frac{p_n^1}{p_n^0}\right)^{s_n^0} \left(s_n^2 - s_n^1\right)^{1/2}.
\]

The relative bias of the sample Jevons price index thus consists of two components, a technical part which vanishes as the sample size gets larger and a structural part which is independent of the sample size. This structural part is given by the right hand side of expression (28). It vanishes when base period and comparison period value shares are equal, which is unlikely to occur in practice.

---

\(^7\) The bias is positive, since \(cv(p_n^1 q_n^1 \ln(p_n^1 / p_n^0)) \leq cv(p_n^0 q_n^0)\) and \(\rho(p_n^1 q_n^1 \ln(p_n^1 / p_n^0), p_n^0 q_n^0) \leq 1\) \((i=0, 1)\).

\(^8\) Jensen’s Inequality says that \(E(f(x)) \geq f(E(x))\) when \(f\) is a convex function of one variable and the expectation of \(x\) exists. This can be shown by expanding \(f(x)\) as a Taylor series around \(E(x)\) and taking the expectation.

\(^9\) See also Bradley (2001, 379).
Instead of defining the size of element \( n \) as its base period value share \( s_n^0 \), one could as well define its size as being \( (s_n^0 + s_n^1)/2 \), the arithmetic mean of its base and comparison period value share. Then we obtain, instead of (27),

\[
E(\hat{P}_f) \geq \prod_{n=1}^{N} \left( \frac{p_n^{'1}}{p_n^{'0}} \right)^{(s_n^0 + s_n^1)/2} = P_f,
\]

which means that the sample Jevons price index is an approximately unbiased estimator of the population Törnqvist price index. The bias will now vanish when the sample size gets larger. This result was mentioned by Diewert (2002; Section 7.4).

6. Heterogeneous aggregates and the Fisher price index

Suppose that instead of the Törnqvist price index one has decided that the Fisher price index (7) should be the target. Suppose further that our sample information consists of prices and quantities. The sample analogue of the population Fisher price index is

\[
\hat{P}_F = \left( \frac{\sum_{n \in S} p_n^{'0} q_n^0 \sum_{n \in S} p_n^{'1} q_n^1}{\sum_{n \in S} p_n^{'0} q_n^0 \sum_{n \in S} p_n^{'1} q_n^1} \right)^{1/2} = \left( \frac{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0 (1/\varphi(S)) \sum_{n \in S} p_n^{'1} q_n^1}{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0 (1/\varphi(S)) \sum_{n \in S} p_n^{'1} q_n^1} \right)^{1/2}.
\]

Then

\[
\ln \hat{P}_F = \frac{1}{2} \left[ \ln \left( \frac{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0}{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0} \right) - \ln \left( \frac{1/\varphi(S) \sum_{n \in S} p_n^{'1} q_n^1}{1/\varphi(S) \sum_{n \in S} p_n^{'1} q_n^1} \right) + \ln \left( \frac{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0}{1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0} \right) \right].
\]

Applying Jensen's Inequality and assuming that the sample was drawn such that the probability that element \( n \) is included in the sample is equal to \( \varphi(S)/N \), we get

\[
E(\ln(1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0)) \leq \ln(E((1/\varphi(S) \sum_{n \in S} p_n^{'0} q_n^0)) = \ln((1/N) \sum_{n=1}^{N} p_n^{'0} q_n^0)
\]

and similar expressions for the other three parts of the right hand side of expression (31). We might expect that the four biases cancel \(^{10}\), and hence

\[
E(\ln \hat{P}_F) \approx \ln P_f.
\]

Again applying Jensen’s Inequality, we see that

\[
E(\hat{P}_F) \geq P_f,
\]

\(^{10}\) To the second order, the bias involved in (32) is \( -(1/2)(cv(p_n^{'0} q_n^0))^2/\varphi(S) \). For the other three parts of (31) the biases have a similar structure.
which means that under simple random sampling the sample Fisher price index has an upward bias relative to its population counterpart. This bias, however, will approach zero when the sample size gets larger.

Suppose now that only sample prices are available, and consider the sample Carli price index,

\[ \hat{P}_c = \frac{1}{\zeta(S)} \sum_{n \in S} P_n^1 / P_n^0. \]

Under probability proportional to size sampling, whereby the size of element \( n \) is defined as its base period value share \( s_n^0 \), we immediately see that

\[ E(\hat{P}_c) = \sum_{n=1}^{N} s_n^0 (P_n^1 / P_n^0) = \sum_{n=1}^{N} P_n^1 q_n^0 / P_n^0 = P_L. \]

Thus the expected value of the sample Carli price index appears to be equal to the population Laspeyres price index. This result was already mentioned by Balk (1994; 139); see also Diewert (2002; Section 7.4). The relative bias of the sample Carli price index with respect to the population Fisher price index appears to be

\[ \frac{E(\hat{P}_c)}{P_F} = \left( \frac{P_L}{P_P} \right)^{1/2}, \]

which is the square root of the ratio of the population Laspeyres price index and the population Paasche price index. Notice that this bias is of structural nature, \( i.e. \) will not disappear when the sample size gets larger.

Note that the population Fisher price index can be written as

\[ P_F = \left( \sum_{n=1}^{N} s_n^0 (P_n^1 / P_n^0) \right)^{1/2} \left( \sum_{n=1}^{N} s_n^0 (P_n^1 / P_n^0)^{-1} \right)^{-1/2}. \]

We now consider whether, following a suggestion of Fisher (1922; 472; formula 101), the Carruthers-Sellwood-Ward (1980) - Dalén (1992) sample price index

\[ \hat{P}_{CSWD} = \left( \frac{1}{\zeta(S)} \sum_{n \in S} (P_n^1 / P_n^0) \right)^{1/2} \left( \frac{1}{\zeta(S)} \sum_{n \in S} (P_n^1 / P_n^0)^{-1} \right)^{-1/2} \]

under some sampling design might be a suitable estimator of the population Fisher price index. The CSWD sample price index is the geometric average of the sample Carli price index (35) and the sample Harmonic (or Coggeshall) price index

\[ \hat{P}_h = \left( \frac{1}{\zeta(S)} \sum_{n \in S} (P_n^1 / P_n^0)^{-1} \right)^{-1}. \]

Thus, consider
(41) \[ \ln \hat{P}_{\text{CSWD}} = \frac{1}{2} \ln \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right) \right) - \frac{1}{2} \ln \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right)^{-1} \right) . \]

Under probability proportional to size sampling, whereby the size of element \( n \) is defined as its base period value share \( s_n^0 \), and again using Jensen's Inequality, we see that

(42) \[ E \left( \ln \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right) \right) \right) \leq \ln \left( E \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right) \right) \right) = \ln \left( \sum_{n=1}^{N} s_n^0 \left( \frac{p_n^1 / p_n^0}{p_n^0} \right) \right) = \ln P_L . \]

Similarly,

(43) \[ -E \left( \ln \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right)^{-1} \right) \right) \geq \ln \left( E \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 / p_n^0}{p_n^0} \right)^{-1} \right) \right) = \ln \left( \sum_{n=1}^{N} s_n^0 \left( \frac{p_n^1 / p_n^0}{p_n^0} \right)^{-1} \right) = \ln P_{HL} , \]

where \( P_{HL} \) is called the population Harmonic Laspeyres price index. Combining these two inequalities, one might expect that the two biases cancel\(^{11}\) and thus

(44) \[ E(\ln \hat{P}_{\text{CSWD}}) \approx \frac{1}{2} (\ln P_L + \ln P_{HL}) = \ln(P_LP_{HL})^{1/2} , \]

or, again using Jensen's Inequality,

(45) \[ E(\hat{P}_{\text{CSWD}}) \geq (P_LP_{HL})^{1/2} . \]

The right hand side of this inequality clearly differs from the population Fisher price index. The relative bias of the CSWD sample price index with respect to the population Fisher price index is

(46) \[ \frac{E(\hat{P}_{\text{CSWD}})}{P_F} \geq \left( \frac{P_{HL}}{P_p} \right)^{1/2} . \]

Notice that the relative bias consists of two components, a technical component which vanishes as the sample size gets larger and a structural component which is independent of the sample size.

Instead of defining the size of element \( n \) as its base period value share \( s_n^0 \), one could as well define its size as being \( (s_n^0 + s_n^1) / 2 \), the arithmetic mean of its base and comparison period value share. Then we obtain, instead of (42),

\(^{11}\) To the second order the bias involved in (42) is \( -(1/2)(cv(p_n^1 / p_n^0))^2 / \zeta(S) \), and the bias involved in (43) is \( +1/2)(cv(p_n^1 / p_n^0))^2 / \zeta(S) \).
where

\[ P_{PAL} = \sum_{n=1}^{N} s_n^1 \left( \frac{p_n^1}{p_n^0} \right) \]

is the population Palgrave price index. Similarly, instead of (43) we get

\[ -E \left( \ln \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1}{p_n^0} \right) \right) \right) \geq \ln \left( \sum_{n=1}^{N} \frac{1}{2} \left( s_n^0 + s_n^1 \right) \left( \frac{p_n^1}{p_n^0} \right)^{-1} \right) = \ln \left( \frac{P_{HL}^{-1} + P_p^{-1}}{2} \right) \]

Combining these two inequalities, one might expect that the two biases cancel and thus

\[ E(\ln \hat{P}_{CSWD}) \approx \ln \left( \frac{P_L + P_{PAL}}{P_{HL}^{-1} + P_p^{-1}} \right)^{1/2}, \]

or, using Jensen's Inequality,

\[ E(\hat{P}_{CSWD}) \geq \left( \frac{P_L + P_{PAL}}{P_{HL}^{-1} + P_p^{-1}} \right)^{1/2} = P_F \left( \frac{1 + P_{PAL}/P_L}{1 + P_p/P_{HL}} \right)^{1/2}. \]

Notice that \( P_p/P_{HL} \) is the temporal antithesis of \( P_{PAL}/P_L \). We may therefore expect that numerator and denominator of the right hand side multiplicative factor will approximately cancel. Thus, under the probability proportional to size sampling design defined immediately before expression (47), the CSWD sample price index turns out to be an approximately unbiased estimator of the population Fisher price index.

We finally consider the following modification of the CSWD sample price index:

\[ \hat{P}_B = \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1}{p_n^0} \right) \right)^{1/2} \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{q_n^1}{q_n^0} \right) \right)^{-1/2} \left( \frac{1}{\zeta(S)} \sum_{n \in S} \left( \frac{p_n^1 p_n^0}{q_n^1 q_n^0} \right) \right)^{1/2}. \]

This is the product of a sample Carli price index, a sample Harmonic quantity index, and a sample Carli value index. It is straightforward to show, using the same reasoning as the previous paragraphs, that under probability proportional to size sampling, whereby the size of element \( n \) is defined as its base period value share \( s_n^0 \),

\[ E(\ln \hat{P}_B) \approx \frac{1}{2} \left[ \ln P_L - \ln Q_L + \ln \left( \frac{\sum_{n=1}^{N} p_n^1 q_n^0}{\sum_{n=1}^{N} p_n^0 q_n^0} \right) \right] = \ln P_F, \]

and thus
However, it is clear that the computation of $\hat{P}_B$ requires more information than the computation of $\hat{P}_{CSWD}$, namely all sample quantity relatives.

7. Heterogeneous aggregates and the Walsh price index

Suppose that the Walsh price index (9) was chosen as the target and that our sample information consists of prices and quantities. The sample analogue of the population Walsh price index is

$$\hat{P}_W = \frac{\sum_{n \in S} P_n^1 (q_n^0 q_n^1)^{1/2}}{\sum_{n \in S} P_n^0 (q_n^0 q_n^1)^{1/2}}.$$  

Suppose again that $S$ is a simple random sample. Then we find that

$$E(\hat{P}_W) = E\left(\frac{(1/\varsigma(S))\sum_{n \in S} P_n^1 (q_n^0 q_n^1)^{1/2}}{(1/\varsigma(S))\sum_{n \in S} P_n^0 (q_n^0 q_n^1)^{1/2}}\right) \approx \frac{E((1/\varsigma(S))\sum_{n \in S} P_n^1 (q_n^0 q_n^1)^{1/2})}{E((1/\varsigma(S))\sum_{n \in S} P_n^0 (q_n^0 q_n^1)^{1/2})}$$

$$= \frac{(1/N)\sum_{n=1}^{N} P_n^1 (q_n^0 q_n^1)^{1/2}}{(1/N)\sum_{n=1}^{N} P_n^0 (q_n^0 q_n^1)^{1/2}} = P_W,$$

which means that the sample Walsh price index is an approximately unbiased estimator of the population Walsh price index.

Suppose now that only sample prices are available. The population Walsh price index can be written as a quadratic mean of order 1 index,

$$P_W = \frac{\sum_{n=1}^{N} (s_n^0 s_n^1)^{1/2} (p_n^1 / p_n^0)^{1/2}}{\sum_{n=1}^{N} (s_n^0 s_n^1)^{1/2} (p_n^1 / p_n^0)^{-1/2}},$$

which suggests the following sample price index

$$\hat{P}_{BW} = \frac{\sum_{n \in S} (P_n^1 / P_n^0)^{1/2}}{\sum_{n \in S} (P_n^1 / P_n^0)^{-1/2}}.$$  

Since there is in the literature no name attached to this formula, expression (58) will be baptized as the Balk-Walsh sample price index. Under a probability proportional to size sampling design, whereby the size of element $n$ is defined as $(s_n^0 s_n^1)^{1/2}$, the geometric mean of its base and comparison period value share, we find that
(59) \[ E(\hat{P}_{BW}) = E\left(\frac{1}{\zeta(S)}\sum_{n \in S} \frac{P_n^1}{P_n^0}^{\frac{1}{\sigma}}\right) \approx \frac{E\left(\frac{1}{\zeta(S)}\sum_{n \in S} \frac{P_n^1}{P_n^0}^{\frac{1}{\sigma}}\right)}{E\left(\frac{1}{\zeta(S)}\sum_{n \in S} \frac{P_n^1}{P_n^0}^{\frac{1}{\sigma}}\right)} = \frac{(1/N) \sum_{n=1}^{N} (s_n^0 s_n^1)^{1/\sigma} (p_n^1/p_n^0)^{-1/2}}{(1/N) \sum_{n=1}^{N} (s_n^0 s_n^1)^{1/\sigma} (p_n^1/p_n^0)^{-1/2}} = P_w. \]

Thus, under this sampling design, the Balk-Walsh sample price index appears to be an approximately unbiased estimator of the population Walsh price index. The bias will approximate zero when the sample size gets larger.

With help of expression (59) it is easy to demonstrate that, if the size of element \( n \) had been defined as its base period value share, \( s_n^0 \), the expectation of the Balk-Walsh sample price index would be unequal to the population Walsh price index.

8. Heterogeneous aggregates: micro-economic considerations on the choice of the sample price index

The previous three sections demonstrated that, when nothing but sample prices are available and the sampling design is restricted to one that uses only base period value share information, it is impossible to estimate any of the population superlatice price indexes unbiasedly. Basically, we are left with a number of second-best alternatives, namely the sample Jevons (25), Carli (35), Harmonic (40), Carruthers-Sellwood-Ward-Dalén (39), and Balk-Walsh (58) price indexes. Is one of these to be preferred?

To assist in the choice, we consider the sample Generalized Mean price index, which is defined as

\[ (60) \quad \hat{P}_{GM}(\sigma) = \left(\frac{1}{\zeta(S)}\sum_{n \in S} (P_n^1/P_n^0)^{1-\sigma}\right)^{1/(1-\sigma)} \quad (\sigma \neq 1) \]
\[ = \prod_{n \in S} (P_n^1/P_n^0)^{1/\zeta(S)} \quad (\sigma = 1). \]

It is immediately seen that \( \hat{P}_j = \hat{P}_{GM}(1) \), \( \hat{P}_c = \hat{P}_{GM}(0) \), and \( \hat{P}_h = \hat{P}_{GM}(2) \), whereas \( \hat{P}_{CSWD} = [\hat{P}_{GM}(0)\hat{P}_{GM}(2)]^{1/2} \), and \( \hat{P}_{BW} = [\hat{P}_{GM}(1/2)\hat{P}_{GM}(3/2)]^{1/2} \). However, since the Generalized Mean price index is a monotonous function of \( \sigma \), we may conclude that \( \hat{P}_{CSWD} \approx \hat{P}_{BW} \approx \hat{P}_{GM}(1) \). Thus these five sample price indexes are members of the same family.

Under probability proportional to size sampling, whereby the size of element \( n \) is defined as its base period value share \( s_n^0 \), one obtains that

\[ (61) \quad E(\hat{P}_{GM}(\sigma)^{1-\sigma}) = \sum_{n=1}^{N} s_n^0 (p_n^1/p_n^0)^{-\sigma}. \]

To apply Jensen's Inequality, we must distinguish between two cases. If \( \sigma \leq 0 \) we obtain
(62) \( E(\hat{P}_{GM}(\sigma)) \leq \left( \sum_{n=1}^{N} s_n^0 \left( \frac{p_n^1}{p_n^0} \right)^{1-\sigma} \right)^{1/(1-\sigma)} \equiv P_{LM}(\sigma) \),

whereas if \( \sigma \geq 0 \) we obtain

(63) \( E(\hat{P}_{GM}(\sigma)) \geq \left( \sum_{n=1}^{N} s_n^0 \left( \frac{p_n^1}{p_n^0} \right)^{1-\sigma} \right)^{1/(1-\sigma)} \equiv P_{LM}(\sigma) \quad (\sigma \neq 1) \)

\[
E(\hat{P}_{GM}(1)) \geq \prod_{n=1}^{N} \left( \frac{p_n^1}{p_n^0} \right)^{s_n^0} \equiv P_{LM}(1),
\]

where \( P_{LM}(\sigma) \) is the Lloyd-Moulton population price index. Economic theory teaches us that this index is exact for a Constant Elasticity of Substitution revenue function (for the producers' output side) or cost function (for the producers' input side or the consumer). The parameter \( \sigma \) is thereby to be interpreted as the (average) elasticity of substitution. At their output side, producers are supposed to maximize revenue, which implies a non-positive elasticity of substitution. Producers at their input side and consumers, however, are supposed to minimize cost, which implies a non-negative elasticity of substitution.

In particular, the conclusion must be that, under the sampling design here assumed, the sample Jevons, Harmonic, CSWD, and Balk-Walsh price indexes are inadmissible for the producer output side since the expected value of each of these indexes would exhibit a positive substitution elasticity. The sample Carli price index is admissible, even unbiased, but would imply a zero substitution elasticity.

9. Heterogeneous aggregates and the Lowe price index

We now turn to the more realistic case in which the Lowe price index (11) is defined to be the target. The population Lowe price index can be written as a ratio of two Laspeyres price indexes

\[
P_{Lo} = \frac{\sum_{n=1}^{N} p_n^1 q_n^b / \sum_{n=1}^{N} p_n^1 q_n^b}{\sum_{n=1}^{N} p_n^0 q_n^b / \sum_{n=1}^{N} p_n^0 q_n^b} = \frac{\sum_{n=1}^{N} s_n^0 (p_n^1 / p_n^0)}{\sum_{n=1}^{N} s_n^0 (p_n^0 / p_n^0)},
\]

where \( s_n^b \) is element \( n \)'s value share in period \( b \) (\( n=1, \ldots, N \)), which is assumed to be some period prior to the base period. This suggests the following sample price index\(^{13}\)

\[
\hat{P}_{Lo} = \frac{\sum_{n \in S} p_n^1 / p_n^b}{\sum_{n \in S} p_n^0 / p_n^b},
\]

which is the ratio of two sample Carli price indexes. Indeed, under a probability proportional to size sampling design, whereby the size of element \( n \) is defined as \( s_n^b \), that is its period \( b \) value share, it is easily demonstrated that

\(^{12}\) For the consumer case, see Balk (2000).

\(^{13}\) See also Bradley (2001, 377). Note that Bradley uses 'modified Laspeyres index' instead of 'Lowe index'.
The bias might be expected to be positive, since in a situation of monotone price changes it will be the case that $cv(p^n_i / p^n_b) \approx cv(p^n_o / p^n_b)$, whereas $\rho(p^n_i / p^n_b, p^n_o / p^n_b) < 1$.

Alternatively and more consistent with practice, one could consider the so-called price-updated period $b$ value shares, defined as

$\begin{equation}
(67)
S_n^{b(o)} = \frac{S_n^b (p^n_0 / p^n_b)}{\sum_{n=1}^N S_n^b (p^n_0 / p^n_b)} = \frac{P_n^b q_n^b}{\sum_{n=1}^N P_n^b q_n^b} \quad (n=1, \ldots, N).
\end{equation}$

Under a probability proportional to size sampling design, whereby the size of element $n$ is now defined as $S_n^{b(o)}$, that is its price-updated period $b$ value share, it is immediately seen that

$\begin{equation}
(68)
E(\hat{P}_C) = \sum_{n=1}^N S_n^{b(o)} (p^n_i / p^n_o) = P_{Lo}^b,
\end{equation}$

that is, the sample Carli price index is an unbiased estimator of the population Lowe price index. However, if the size of element $n$ was defined as $S_n^b$, that is its period $b$ value share itself, one would have obtained

$\begin{equation}
(69)
E(\hat{P}_C) = \sum_{n=1}^N S_n^b (p^n_i / p^n_o),
\end{equation}$

which, unless the prices have not changed between the periods $b$ and $0$, differs from the population Lowe price index.

10. The Time Reversal test and some numerical relations

When there is nothing but sample price information available, that is, the sample has the form \( \{p^n_t; t = 0,1; n \in S\} \), then the menu of sample price indexes appears to be limited. For a homogeneous aggregate only the sample Dutot price index (19) is available. Note that this index, like the population unit value index, satisfies the Time Reversal test, that is, using obvious notation,

$\begin{equation}
(70)\ \hat{P}_D(p^1, p^0) \hat{P}_D(p^0, p^1) = 1.
\end{equation}$

However, as has been shown, under a not unreasonable sampling design, the sample Dutot price index is a biased estimator of the target unit value index.
For a heterogeneous aggregate one has, depending on the definition of the target price index, the choice between the sample Carli price index (35), the sample Jevons price index (25), the sample Harmonic price index (40), the sample CSWD price index (39), the sample Balk-Walsh price index (58), and the sample Lowe price index (65). The first three indexes are special cases of the sample Generalized Mean price index (60), respectively for \( \sigma = 0, 1, 2 \). Since the Generalized Mean price index is monotonously increasing in \( 1 - \sigma \), we obtain the general result that

\[
\hat{P}_{GM}(p^1, p^0; \sigma) \hat{P}_{GM}(p^0, p^1; \sigma) \geq 1 \quad \text{for } \sigma < 1
\]

(71)

\[
\hat{P}_{GM}(p^1, p^0; \sigma) \hat{P}_{GM}(p^0, p^1; \sigma) \leq 1 \quad \text{for } \sigma > 1,
\]

(72)

which means that the GM price index fails the Time Reversal Test. In particular, the Carli price index and the Harmonic price index fail the Time Reversal test, that is,

\[
\hat{P}_c(p^1, p^0) \hat{P}_c(p^0, p^1) \geq 1,
\]

(73)

\[
\hat{P}_H(p^1, p^0) \hat{P}_H(p^0, p^1) \leq 1.
\]

(74)

The Jevons price index, as well as the CSWD price index and the Balk-Walsh price index satisfy the Time Reversal test, as one verifies immediately. As has been shown, under a not unreasonable sampling design, these three sample price indexes are (approximately) unbiased estimators of the Lloyd-Moulton population price index with \( \sigma = 1 \).

The sample Lowe price index also satisfies the Time Reversal Test. This index is, under a not unreasonable sampling design, an (approximately) unbiased estimator of the population Lowe price index.

We now turn to numerical relations between all these indexes. It is well known that

\[
\hat{P}_H \leq \hat{P}_J \leq \hat{P}_C,
\]

(75)

and thus we might expect that

\[
\hat{P}_{CSWD} = (\hat{P}_H \cdot \hat{P}_C)^{1/2} \approx \hat{P}_J.
\]

(76)

The magnitudes of the differences between the indexes depend on the variance of the price relatives \( p^1_n / p^0_n \). When all the price relatives are equal, the inequalities (75) turn into equalities. In fact, Dalén (1992) and Diewert (1995) showed that, to the second order, the following approximations hold:

\[
\hat{P}_J \approx \hat{P}_C(1 - \frac{1}{2}\text{var}(\varepsilon))
\]

(77)

\[
\hat{P}_H \approx \hat{P}_C(1 - \text{var}(\varepsilon))
\]

(78)
(79) \( \hat{P}_{CSWD} \approx \hat{P}_C (1 - \frac{1}{2} \text{var}(\varepsilon)) \),

where \( \text{var}(\varepsilon) = (1/\zeta(S)) \sum_{n \in S} \varepsilon_n^2 \) and \( \varepsilon_n = (p_n^1/ p_n^0 - \hat{P}_C)/ \hat{P}_C \) \( (n \in S) \). In the same way\(^{14}\) one can show that

(80) \( \hat{P}_{BW} \approx \hat{P}_C (1 - \frac{1}{2} \text{var}(\varepsilon)) \).

Thus we may conclude that the sample Jevons price index, the sample CSWD price index, and the sample Balk-Walsh price index approximate each other to the second order. From the point of view of simplicity, the sample Jevons price index obviously gets the highest score.

To obtain some insight into the relation between the sample Lowe price index (65) and the sample Carli price index (35), we write the first as

(81) \( \hat{P}_{Lo} = \frac{\sum_{n \in S} (p_n^0/p_n^b)(p_n^b/p_n^0)}{\sum_{n \in S} p_n^0/p_n^b} \).

Consider now the difference \( \hat{P}_{Lo} - \hat{P}_C \). By straightforward manipulation of this expression one can show that

(82) \( \hat{P}_{Lo} = \hat{P}_C (1 + \text{cov}(\delta, \varepsilon)) \),

where \( \text{cov}(\delta, \varepsilon) = (1/\zeta(S)) \sum_{n \in S} \delta_n \varepsilon_n \), \( \delta_n = (p_n^0/p_n^b - \hat{P}_C(p_n^0, p_n^b))/ \hat{P}_C(p_n^0, p_n^b) \), and \( \varepsilon_n = (p_n^1/ p_n^0 - \hat{P}_C(p_n^1, p_n^0))/ \hat{P}_C(p_n^1, p_n^0) \) \( (n \in S) \). Thus the difference between these two sample price indexes depends on the covariance of the relative price changes between the periods \( b \) and 0 and those between the periods 0 and 1. Whether this difference is positive or negative, large or small, is an empirical matter.

Since in practice the sample Dutot price index appears to be used frequently in the case of heterogeneous aggregates, it might be of some interest to discuss the relation between this index and the sample Jevons index. The first is a ratio of arithmetic average prices whereas the second is a ratio of geometric average prices. In order to see their relation, we write the Jevons index as

(83) \( \ln \hat{P}_j = (1/\zeta(S)) \sum_{n \in S} \ln(p_n^1/ p_n^0) \)

and the Dutot index as

---

\(^{14}\) The method of proof is to write the ratio of \( \hat{P}_{BW} \) to \( \hat{P}_C \) as a function \( f(\varepsilon) \) and expand this function as a Taylor series around 0. Notice thereby that \( \sum_{n \in S} \varepsilon_n = 0 \).
\[
\ln \hat{P}_D = \sum_{n=3} \left( \frac{L(p_n^0 / P^0, p_n^1 / \bar{P}^1)}{\sum_{n=5} L(p_n^0 / P^0, p_n^1 / \bar{P}^1)} \right) \ln(p_n^1 / p_n^0),
\]

where \( \bar{P}^t = (1/\zeta(S)) \sum_{n=5} p_n^t \) (\( t = 0, 1 \)) are the arithmetic average prices and \( L(,,,) \) denotes the logarithmic mean. This mean is, for any two positive numbers \( a \) and \( b \), defined by

\[
L(a,b) = (a - b) / \ln(a / b) \quad \text{and} \quad L(a,a) = a.
\]

It is a symmetric mean with the property that \( (ab)^{1/2} \leq L(a,b) \leq (a+b)/2 \), that is, it lies between the geometric and the arithmetic mean. \(^{15}\) Thus, \( L(p_n^0 / \bar{P}^0, p_n^1 / \bar{P}^1) \) can be interpreted as the mean relative price of element \( n \). Then

\[
\ln \hat{P}_D - \ln \hat{P}_I = \sum_{n=3} \left( \frac{L(p_n^0 / \bar{P}^0, p_n^1 / \bar{P}^1)}{\sum_{n=5} L(p_n^0 / \bar{P}^0, p_n^1 / \bar{P}^1)} - \frac{1}{\zeta(S)} \right) \ln(p_n^1 / p_n^0)
\]

\[
= \frac{1}{\zeta(S)} \sum_{n=3} \left( \frac{L(p_n^0 / \bar{P}^0, p_n^1 / \bar{P}^1)}{(1/\zeta(S)) \sum_{n=5} L(p_n^0 / \bar{P}^0, p_n^1 / \bar{P}^1)} - 1 \right) \ln(p_n^1 / p_n^0) - \ln \hat{P}_I,
\]

which means that the (sign of the) difference between the Dutot and the Jevons index depends on the (sign of the) covariance between relative prices and price relatives. Whether this difference is positive or negative, large or small, is an empirical matter.

11. Conclusion

The theoretical arguments advanced in the previous sections lead us to the following practical advice. The advice, to be practical, concerns simple random sampling (srs), sampling with probability proportional to base period quantity shares (in the case of a homogeneous aggregate), and sampling with probability proportional to base period or (price-updated) earlier period value shares (in the case of a heterogeneous aggregate) (pps). It is recognised that sampling in practice may take two stages: the sampling of respondents (outlets or firms) and of commodities. The discussion here was kept for simplicity in terms of single stage sampling. It is also recognised that purposive sampling and/or sampling with cut-off rules are often used at either stage. In such circumstances there are implicit sampling frames and selection rules and some judgement will be necessary as to which sample design outlined most closely corresponds to the method used, and the implications for choice of the sample index.

The following table presents the key results in the order of their appearance. In the first place, it is clear that respondents should be encouraged to provide timely data on comparison and base period values and prices (or quantities). Of course, in some areas this should be more feasible than others. In such cases sample indexes which mirror their population counterparts should be used and respondent-commodity pairs should be sampled using simple random sampling, since each sample index would then be an (approximately) unbiased estimator of the corresponding population one.

\(^{15}\) A simple proof was provided by Lorenzen (1990).
Sample
index  |  price   | Target
index  |  price   | Sampling
definition  | Expected value of
sample index  | Main
equation
---|---|---|---|---|---|---
Unit value | Unit value | Unit value | Unit value | (17)
Dutot | Unit value | Unit value | pps-q0 | Biased estimate of target index | (20)
Törnqvist | Törnqvist | Törnqvist | srs | Törnqvist | (24)
Jevons | Törnqvist | Törnqvist | pps-s0 | Geometric Laspeyres = Lloyd-Moulton (1) | (27)
Fisher | Fisher | Fisher | srs | Fisher | (34)
Carli | Fisher | Fisher | pps-s0 | Laspeyres = Lloyd-Moulton (0) | (36)
CSWD | Fisher | Fisher | pps-s0 | Lloyd-Moulton (1) | (45)
Walsh | Walsh | Walsh | srs | Walsh | (56)
Balk-Walsh | Walsh | Walsh | pps-s0 | Lloyd-Moulton (1) | (62) – (63)
Generalized Mean (σ) | Lloyd-Moulton (σ) | Lloyd-Moulton (σ) | pps-s0 | Lloyd-Moulton (σ) | (62) – (63)
Lowe | Lowe | Lowe | pps-sb | Lowe | (66)
Carli | Lowe | Lowe | pps-sb(0) | Lowe | (68)

When this approach is not feasible and the best one can obtain is a sample of (matched) prices, the sampling design should be such that important elements have a correspondingly higher probability of inclusion in the sample than unimportant elements. With respect to the sample price index to be used:

- For a homogeneous aggregate, that is an aggregate for which the quantities of the elements can be meaningfully added up, one should use the sample Dutot price index. Unfortunately, this index will exhibit bias, the magnitude whereof depends on the dispersion of the elementary quantity changes between the two periods compared.

- For a heterogeneous aggregate, except at the producers’ output side, one could use the sample Jevons price index. Its expected value will approximate the Geometric Laspeyres price index, which is identical to the Lloyd-Moulton price index with $\sigma = 1$.

- For a heterogeneous aggregate at the producers’ output side one could use a sample Generalized Mean price index with appropriately chosen parameter $\sigma \leq 0$, the limiting case being the sample Carli price index. The expected value of such a price index will approximate a Lloyd-Moulton price index.

- When the target is a Lowe price index, the sample Lowe and Carli price indexes exhibit appropriate behaviour.

In any case the time span between the two periods compared should not become too long, for the magnitude of the bias will in general grow with the length of the time span. That is, at regular time intervals one should undertake a base period change.
There remains the practical issue as to how to decide whether an aggregate is homogenous or not. The question posed in (2) above was:

Does it make (economic) sense to add up the quantities $q'_n$ of the elements $n=1,...,N$?

For example, if the aggregate consists of 14 inch television sets, the answer must be 'no'. Brand differences, additional facilities such as stereo, wide screens and much more account for significant variations in price. Tins of a specific brand and type of food of different sizes similarly lack homogeneity, since much of the price variation will be due to tin size. Homogeneity is lacking when the item itself varies according to identifiable price-determining characteristics. In principle the conditions of sale need to be taken into account, since an item sold by one manufacturer may command a price premium since it has better delivery, warranties or other such features. The price at initiation should be defined to have the same specified conditions of sale, but there may be elements of trust in the buyer-seller relationship that are difficult to identify. Nonetheless for practical purposes items sold by different establishments for the same product are practically treated as homogenous unless there are clearly identifiable differences in the terms and conditions surrounding the sale.
References


Why Price Index Number Formulae Differ: Economic Theory And Evidence On Price Dispersion

Mick Silver and Saeed Heravi
Cardiff Business School, Cardiff University, UK

Acknowledgements are due to the UK Office for National Statistics (ONS) for supplying the data for this study. This follows a well established research link with the ONS which has been of much benefit to the authors. Naturally the ONS are not responsible for any views, errors of omission or commission in this paper. We also acknowledge useful comments from Malcolm Galatin (City College New York).

Abstract: The results from different index number formulae can differ and can do so substantially. The main criteria for explaining such differences, and governing choice between them, are their ability to satisfy desirable test properties—the axiomatic approach—and their correspondence with plausible substitution behavior as predicted from economic theory. Yet the numerical differences between such formulae has been shown to be related to the extent of, and changes in, the dispersion of prices. However, within the index number literature there is, to the author’s knowledge, no formal attempt to explain differences between the results from individual formulae in terms of theories and evidence on price dispersion. Explaining differences between formulae in terms of changes in the dispersion of prices benefits from the existence of economic theoretical frameworks to explain such dispersion, and thus improve our understanding of why differences from formulae occur. Such frameworks include search cost and menu cost theories and signal extraction models. This paper outlines the nature of the relationships between formulae in term of price dispersion, then considers economic theories of price dispersion and uses them to model price variation both within months and over time using an extensive scanner data set on television sets amounting to over 70,000 observations over 51 months. It concludes by considering the implications for index number construction.

Keywords: Index Numbers; Relative Price Dispersion; Search Cost; Signal Extraction; Menu Cost; Hedonic Regression.

JEL classifications: C43, C81, D11, D12, D83, E31, L11, L15.

1. Introduction

Choice of formula for the measurement of inflation does matter. In January 1999 the formula principally used for aggregating price changes for the U.S. consumer price index (CPI) at the lower level of aggregation was changed from an arithmetic to a geometric mean. The effect of the change has been estimated by the Bureau of Labor Statistics (BLS) (2001) to have reduced the annual rate of increase by approximately 0.2 percentage points. Following
estimates from the Boskin Commission’s Report on the U.S. CPI (Boskin et al., 1996 and 1998), this implied a cumulative additional national debt from over-indexing the budget of more than $200 billion over a twelve year period up to the mid-1990s. This ‘lower’ level aggregation only applies to samples of prices from stores of finely-defined goods such as brands of washing machines, varieties of apples. These are the building blocks of a CPI and the choice of formula for their aggregation is an important practical matter. The difference between these formulae can be shown to be primarily determined by the changes in price dispersion.

The subsequent Schultze and Mackie (2002) report recommended the use of a trailing superlative index instead of the Laspeyres index since it would capture weighted ‘upper level’ substitution effects. One such superlative index which has much to commend it (Diewert, 1995), is the Fisher index, a geometric mean of Laspeyres and Paasche. Boskin estimated that upper level substitution accounted for 0.15 percentage points bias in the U.S. CPI. Again changes in price dispersion will be seen below to account for some of the difference between these formulae.

Although the Laspeyres formula is commonly thought to be the formula used for the U.S. and other CPIs at the upper level, the expenditure weights for a comparison, say between periods 0 and \( t \), relate to a previous time period \( b \), as opposed to period 0, since it takes time to compile the information from expenditure surveys for the weights. The resulting practically used index is a Young index which is shown below to be biased (Diewert, 2003). But the extent of the bias depends again on changes in price dispersion.

In spite of the importance of price dispersion in explaining the differences between these key formulae, it is axiomatic tests that are used to choose between such formulae along with the economic theory of consumer substitution. The contribution of this paper lies in its analysis of the differences between formula in terms of explanations of the factors governing price dispersion and changes in such dispersion over time. For explanation it draws on the very nature of product heterogeneity in terms of the characteristics, branding and outlet-types goods they are sold in. But it extends the analysis to theoretical frameworks not usually associated with index number work, including search cost and menu cost theories and signal extraction models. The empirical section is directly related to such theory and is based on detailed scanner data from retailers’ bar-code readers. More particularly, the empirical work on lower level indices provides a hitherto neglected focus on product heterogeneity to explain the bias in the Dutot index.

Section 2 considers elementary index number formulae. It outlines the three main formulae and their current justification from axiomatic considerations, economic and sampling theory. Section 3 shows how changes in price dispersion are important to any explanation of differences in upper level formulae, whose justification has again been in terms of consumer substitution theory and axiomatic tests. Section 4 and the Annexes provide the numerical relationships between the formulas in terms of how they differ in terms of changes in the variance of their prices. The paper brings to bear in section 5 a quite novel approach to the consideration of the difference between such formulae with a focus on search cost theory, but
including menu cost theory, signal extraction models, pass-through rates, price discrimination and consumer inventory models. Section 6 commences the empirical work, based on extensive retail bar-code scanner data for television sets (TVs) with an outline of data, variables and measures and section 7 provides the results. Section 8 concludes with implications for index number compilation. The analysis shows how economic theory rooted in the failure of the law of one price and the persistence of price dispersion can provide insights into differences between index number formula at upper and lower levels. This novel approach\(^3\) complements the still valuable analysis previously considered only in terms of axioms and consumer substitution theory.

2. Lower-Level Formulae and their Rationale: Axioms and Consumer Substitution Theory

2.1 The formulae

The main formulas used in practice (see Dalen (1992) and Diewert (1995 and 2003) for details of other such indices)\(^4\) are given, for \(m=1,..M\) items with prices and quantities in period \(t\), \(p_m^t\) and \(q_m^t\) respectively for \(t=0, t\), by:

The arithmetic mean of price relatives—the Carli price index \(P_C\):

\[
P_C(p^0,p^t) = \sum_{m=1}^{M} \left( \frac{1}{M} \right) \frac{p_m^t}{p_m^0} \]  

The relative of the arithmetic means of arithmetic averages—the Dutot price index, \(P_D\):

\[
P_D(p^0,p^t) = \frac{\sum_{m=1}^{M} (p_m^t/M)}{\sum_{m=1}^{M} (p_m^0/M)} = \frac{\sum_{m=1}^{M} (p_m^t)}{\sum_{m=1}^{M} (p_m^0)} = \frac{\sum_{m=1}^{M} (p_m^t/p_m^0)}{\sum_{m=1}^{M} (p_m^0)} \]  

which can be seen to be a base-period price share weighted Carli index.

The geometric mean of price relatives (which is also equal to the relative (ratio) of geometric means of the prices in periods 0 and \(t\))—the Jevons price index \(P_J\):

\[
P_J(p^0,p^t) = \prod_{m=1}^{M} \left[ \frac{p_m^t}{p_m^0} \right]^{1/M} \]  

The use of the geometric mean is not a novel idea. It was first proposed in 1922 on axiomatic grounds by Irving Fisher, though its adoption was prompted by the Boskin Commission’s Report in 1996 based on the economic theory of consumer substitution behavior. These three widely-used simple formulas are for calculating lower level, aggregate, unweighted price changes of matched items over time. These humble formulas are the building blocks of a CPI until weights are used at a higher level.

\(^3\) Balk (2001:2) comments that some insights have been obtained by looking at changes in variances, but only using an approximate “...more or less intuitive economic reasoning.” There has been, to the author’s knowledge, no formal examination of changes in dispersion in this context.

\(^4\) The main alternative formulae are the harmonic mean of price relatives—the harmonic version of equation (2); the relative of the harmonic means; and the geometric mean of the Carli arithmetic mean \(P_C\) of price relatives and harmonic mean of price relatives.
2.2 The axiomatic approach, price dispersion and commensurability

The axiomatic approach identifies which formulae are desirable on the basis of their satisfaction of reasonable test properties. The Carli index fails the time reversal test such that $P_c(p^0, p^1) \times P_c(p^1, p^0) \geq 1$; it is upwards-biased. The Jevons index satisfies all of the tests as does the Dutot index with the important exception of Commensurability Test, i.e., if we change the units of measurement for each commodity in each outlet, then the elementary index remains unchanged. There is an implication for quality variation here. In practice the quality differences—be they brands, technical specifications, or level of service in outlets—amount to a change in the (utility flow) unit the price is measured in. Thus while each of the matched models will have the same units over time, they may differ across units. This is the concern of the Commensurability Test. A heterogeneous collection of items gives rise to varying units of measurement. This can be seen by identifying the Dutot index in terms of an arithmetic average of weighted price changes the weights being the base period price shares as in (2). If, for example, the quality and prices of say washing machines in the basket are very diverse, then the Dutot index will give more emphasis to models with higher prices, for which there is no immediate justification. Diewert (2003) notes:

"...in actual practice, there will usually be thousands of individual items in each elementary aggregate and the hypothesis of item homogeneity is not warranted. Under these circumstances, it is important that the elementary index satisfy the commensurability test, since the units of measurement of the heterogeneous items in the elementary aggregate are arbitrary and hence the price statistician can change the index simply by changing the units of measurement for some of the items."[their emphasis].

He continues:

"If there are heterogeneous items in the elementary aggregate, this is a rather serious failure and hence price statisticians should be careful in using this index under these conditions."

There is thus a concern with the absolute level of price dispersion and why it arises as well as the changes in dispersion. This paper addresses two main issues in this context. First, that the axiomatic approach only supports the Jevons against the Dutot index because the heterogeneity of items/outlets bundled together in the aggregation leads to bias in the Dutot. We seek to explain such heterogeneity and, in the empirical section, examine heterogeneity-controlled prices for the Dutot index. The superiority of a Jevons index against a heterogeneity-controlled Dutot index may not be straightforward. Second, that the differences between results from such index formulae over time can be explained by changes in the dispersion of prices. Signal extraction search cost and menu cost theory are brought to bear to explain and model such differences in dispersion, both heterogeneity-controlled and otherwise.

---

5 Fisher (1922) famously commented: "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose." Irving Fisher (1922; 29-30).

6 The BLS in their CPI argued that only 60% of product areas should be changed to using the Jevons index as justified by a likely approximation to substitution behaviour characterised by an elasticity of substitution equations.
2.3 Consumer substitution theory

Consumer substitution theory holds that utility-maximizing consumers substitute away from items with relatively high prices. An index that ignores such effects in its weighting of price changes is open to substitution bias. However, lower level indices do not include information on weights so such theory may be argued to not be relevant. Yet Balk (2002) has shown that if the items are selected with probability proportionate to (quantity or value share) size (pps), then the sample unweighted estimator is of a population weighted target index, for which economic theory applies. For example, if sampling is with probability proportionate to base period value shares, then the expected value of a Carli index is a Laspeyres index. A Laspeyres index restricts consumer substitution to be zero and overstates inflation since items with above average price changes are not given less weight since any fall in quantity is not reflected in the weights. However, the expected value of a Jevons index under the same sampling scheme produces a base period weighted geometric mean which Balk (2002) has shown to correspond to consumer substitution behavior consistent with an elasticity of substitution of unity. The incorporation of such substitution effects was the main justification for the BLS switch to the Jevons index.


3.1 The formulae

Laspeyres and Paasche indices are fixed basket indices measuring the price change of a basket of goods whose quantities are either fixed in period 0, Laspeyres:

\[
P_L = \frac{\sum_{m=1}^{M} P_m^t q_m^0}{\sum_{m=1}^{M} P_m^0 q_m^0}
\]

or period \( t \), Paasche

\[
P_P = \frac{\sum_{m=1}^{M} P_m^t q_m^t}{\sum_{m=1}^{M} P_m^0 q_m^t}
\]

A geometric mean of the two is the Fisher index:

\[
P_F = (P_L P_P)^{1/2}
\]

Paasche and Fisher cannot be used in real time because it takes time to compile the quantity weights from expenditure surveys. While the CPI is considered to be a Laspeyres index, this is not the case in fact. The weights are taken from a survey of expenditure patterns in a weight reference period \( b \). It takes time to compile such results so that their final use is for a subsequent comparison between price reference period 0 and period \( t \). The resulting Young index is:

\[
P_Y = \sum_{m=1}^{M} s_m^b \left( \frac{P_m^t}{P_m^0} \right)
\]

where

\[
s_m^b = \frac{P_m^b q_m^b}{\sum_{m=1}^{M} P_m^b q_m^b}; \quad m = 1,\ldots,M.
\]
3.2 The axiomatic approach and consumer substitution theory

Fisher (1922) defined (6) as 'ideal' in terms of its satisfaction of desirable axioms (see also Diewert, 1996). In seminal work Konüs (1924) found Laspeyres and Paasche price indices to provide upper and lower bounds on theoretical cost-of-living indices and Diewert (1976 and 1978) defined a class of superlative formulae, one of which was the Fisher index, to incorporate forms of substitution effects corresponding to flexible functional forms. Such results render a Fisher price index as superior to Laspeyres price index with regard to its ability to incorporate substitution effects, which the Laspeyres fixed basket index cannot, and thus as a better approximation to a cost-of-living index.

4. Numerical Relationships between Frequently Used Index Number Formulae – the Importance of Dispersion

4.1 Lower level indices

The relationships have been developed by Marks and Stuart (1971), Carruthers et al., (1980), Dalen (1992), and Diewert (1995 and 2003). We borrow in this section primarily from Diewert for the exposition. Three things are apparent from the above section 2. First, the Carli index may be biased on axiomatic grounds. Second, the Dutot index is only advisable when there is limited price dispersion arising from quality differences in the item itself or the services provided by the outlet at the time of sale. Finally, that the Jevons index has much to commend it on axiomatic grounds and also, under pps, when consumer substitution behaviour approximates that characterised by a unitary elasticity of substitution. More innocuous is the incentive for governments to encourage a switch to the use of the Jevons index since it will lead to lower inflation than its arithmetic counterparts and thus, lower public (index-linked) debt (Hulten, 2002). We thus consider the numerical relationships between these three formulas.

4.1.1 The relationship between Dutot and Jevons indices

First, it can be shown that the Carli and Jevons satisfy the following inequality:

\[ P_J(p_0, p_1) \leq P_C(p_0, p_1) \]  

i.e., the Jevons index is always equal to or less than the Carli index. In fact, the strict inequality in (6) will hold provided that the period 0 vector of prices, \( p_0 \), is not proportional to the period 1 vector of prices, \( p_1 \).

The inequality (6) does not tell us by how much the Carli index will exceed the Jevons index. Carruthers, Sellwood and Ward, (1980:25) show an approximate relationship between Dutot \( P_D(p_0, p_1) \) and Jevons \( P_J(p_0, p_1) \) — see also Diewert (1995a:27-28) and Balk (2002: 23-4) and Annex 1 for more detail. Consider

---

8 As noted by Diewert (1995) each of the three indices \( P_H \), \( P_J \) and \( P_C \) is a mean of order \( r \) where \( r \) equals \(-1, 0 \) and \( 1 \) respectively and so the inequalities follow from Schlömilch’s inequality.
\[ P_m^t = \bar{p}^t (1 + \varepsilon_m^t) \quad \text{for } t = 0, 1, \ldots, t \]  
where \( \sum_{m=1}^{M} \varepsilon_m^t = 0 \) \hspace{1cm} (9)

and \( \bar{p}^t \) is the arithmetic mean of prices in period \( t \). To realize the individual prices the mean is multiplied by each item’s deviation from the mean, \( (1 + \varepsilon_m^t) \). Since the Dutot price index is defined as \( P_D(p^0, p^1) = \bar{p}^t / \bar{p}^0 \) it follows that the Jevons price index is given by:

\[ P_J(p^0, p^1) = \prod_{m=1}^{M} \left[ \left( 1 + \varepsilon_m^t \right) / \bar{p}^0 \right]^{1/M} = \prod_{m=1}^{M} \left[ \bar{p}^t / \bar{p}^0 \left( 1 + \varepsilon_m^t \right) / (1 + \varepsilon_m^0) \right]^{1/M} \]

\[ = P_D(p^t, p^0) f(\varepsilon^t, \varepsilon^0) \quad \text{where } f(\varepsilon^t, \varepsilon^0) = \prod_{m=1}^{M} \left( (1 + \varepsilon_m^t) / (1 + \varepsilon_m^0) \right)^{1/M} \] \hspace{1cm} (10)

Expanding \( f(\varepsilon^t, \varepsilon^0) \) using a second order Taylor series around \( \varepsilon^0 \) and \( \varepsilon^t \), the following second order approximation results:

\[ P_J(p^0, p^1) = P_D(p^1, p^0) \left( 1 + (1/2) \text{var}(\varepsilon^0) - (1/2) \text{var}(\varepsilon^t) \right) \] \hspace{1cm} (11)

4.1.2 The relationship between Carli and Jevons indices

In equation (2) the elementary index number took the form of a mean of price relatives. Alternative means of price relatives considered in section 2.1 were the arithmetic Carli index \( P_C \), geometric Jevons index \( P_J \), harmonic index \( P_H \), and the geometric mean of \( P_H \) and \( P_C \), the \( P_{HRC} \) index. Since these are all functions of price ratios it is quite straightforward to establish the mathematical relationships between them (see Dalen (1992) and Diewert (1995) for formal proofs). The interest here lies in the conditions under which these formulae approximate each other and the factors determining their differences. Let the price relatives for a comparison between 0 and \( t \) be given by:

\[ R_{m}^t = p_m^t / p_m^0 = \bar{r}^t (1 + \nu_m^t) \] \hspace{1cm} (12)

where \( \nu_m^t \) is the deviation from the mean \( r^t \) of the price relative for item \( m \) in period \( t \) and \( \sum_{m=1}^{M} \nu_m^t = 0 \). Now regard each formula as a function of the deviations \( \nu_m^t \) and take second order Taylor series approximations around the point \( \nu_m^t = 0 \) for \( m = 1, \ldots, M \) and we obtain the following approximations:

\[ P_C(p^0, p^t) = \sum_{m=1}^{M} \left( r_m^t (1/M) = \bar{r}^t \sum_{m=1}^{M} (1 + \nu_m^t) (1/M) = \bar{r}^t \prod_{m=1}^{M} \right)^{1/M} \] \hspace{1cm} (13)

\[ P_J(p^0, p^t) = \prod_{m=1}^{M} \left[ \left( 1 + \nu_m^t \right) \right]^{1/M} = \bar{r}^t \prod_{m=1}^{M} \left( 1 + \nu_m^t \right) \equiv \bar{r}^t (1 - (1/2) \text{var}(\nu_m^t)) \] \hspace{1cm} (14)

This formulation of differences in formulae in terms of the dispersion of price relatives is of use since, as will be shown in section 3, there are theoretical frameworks in economics concerned with explaining the existence and change in price dispersion. Annex 2 outlines the

9 Similar results can be found in relation to the Carli index and each of the Harmonic mean of price relatives, the Carruthers, Sellwood, Ward and Dalen index (geometric mean of Harmonic and Carli index), and the Balk-Walsh index, the approach having a wider application than to the more widely used Carli, Dutot and Jevons indices (Balk, 2002: 22).

International Working Group on Price Indices - Seventh Meeting 181
relationship in more detail. Differences between Dutot and Carli can also in part be explained by changes in such variances (Dievert, 1995) and Annex 3.

4.2 Upper level indices

4.2.1 The relationship between Laspeyres and Paasche

Annex 4 shows the extent of the divergence between these two formulae, and by extension between Laspeyres and Fisher, to depend in part on the extent of the dispersion in price relatives where dispersion is considered in terms of the coefficient of variation.

4.2.2 The bias in the Young index

It was noted in section 3 that in practice Laspeyres index is not used at the upper level for a price comparison between periods 0 and $t$ since expenditure weights are unavailable in period 0 due to the time taken to compile them. Instead the weight reference period refers to an earlier period $b$, the resulting index being the Young index defined by equation (7). Annex 5 follows Dievert (2003) and shows the Young index to be biased, the extent of the bias depending on the dispersion in price relatives.

4.2.3 A note concerning the Taylor expansion/approximation

The results from the above sections and annexes show the differences between formulae in terms of variances, usually arising from a Taylor expansion around zero. It is an approximation and Annex 6 considers the expansion in more detail.

5. Some economic theory of price dispersion and its change over time

The fact that the decomposition of the differences between index number formulae can be identified, at least partially, in terms of changes in variances allows recourse to economic theory concerned with such changes. A body of well-developed economic theory is available regarding the existence and persistence of changes in, the variances of prices and this theory. The theory is quite different from classical theory, for which price dispersion is an anathema. Jevon’s law of one price predicts that under perfect competition identical items will be sold at the same price. There is a burgeoning theoretical literature that explains price dispersion of homogeneous goods. There is a related literature that links changes in (relative) price dispersion to changes in the mean (unanticipated and otherwise) of such prices. Much of this theory which includes search cost, menu cost and signal extraction models relates to micro-economic behavior. Lach (2002), in his useful contribution using micro data, notes a dearth of related empirical studies blaming it on problems with access to micro-level data.

5.1 Search Costs and the law of one price: cross-sectional price dispersion

Stigler (1961) argued that optimizing consumers with imperfect information search for additional information such that their (rising) marginal search cost equals the (falling) marginal search benefits. Even in markets with symmetric firms selling homogeneous goods, product prices may differ in equilibrium if there is a positive, but uncertain, probability that a randomly chosen customer knows only one price.¹⁰ This would result in imperfect information that the firm can exploit by charging a higher price (Sorensen, 2002). It is in the

---

¹⁰ It is not even required that search costs vary across buyers. Heterogeneity of beliefs about the (cumulative distribution function of) prices for buyers with identical search costs is sufficient (Rauh, 2001).
interests of firms to adopt strategies which increase search costs, the effect of which is to increase price dispersion and thus the difference between formulae.

Electronic consumer durables (ECDs) have particular characteristics with regard to search costs. First, they are highly differentiated by brand and features. This may be argued to meet the needs of different segments of the market, but this also hinders search and increases price dispersion. There is competition within stores, between brands and within brands with different features. There is also competition between stores yet, as will be shown, stores do not often stock the same brands or models of a brand to further hinder price comparisons.

It may be argued that advertising serves to reduce search costs. There are intensive media advertising and door-to-door flyers for ECDs. Yet this is usually on a small selection of models. Koch and Cebular (2002) distinguish between advertising expenditure which reduces consumer search costs and decreases mean prices and advertising that focuses on branding to diminish price elasticity of demand thus increase mean prices and their dispersion. Cohen (2002) similarly argues that while greater brand selection increases rivalry and stimulates price competition, it also increases the value of information on prices and features providing scope for poorly informed customers which can dampen price competition.

Second, models are sold in a variety of outlets offering different levels of customer services. Price dispersion may be due to outlet heterogeneity as well as feature and brand heterogeneity. ECDs are sold in electrical multiples/chains (EM), mass merchandisers (department stores) (MM), independents (IND), and mail order catalogues (MAIL). The majority of sales are in EMs which specialize in electrical goods, each chain being made up of hundreds of branches spread across the country selling a similar and large selection of goods at the same prices. The different types of outlets provide different types of service: EMs are often large out-of-town (easy parking), specialist warehouses while MMs are usually in-town department stores selling a much wider range of goods. Sorensen (2000) found store effects to not be a source of price variation for pharmaceutical products. In contrast, Lach (2002) found outlet-type to explain some of price variability for chicken, flour, coffee and refrigerators. In this study outlet types are one feature considered to explain price dispersion.

Third, Sorensen (2000) found the prices of repeatedly purchased prescriptions to be lower and less dispersed than irregularly purchased ones. He argued that the search benefits from repeat prescriptions were higher since the savings could be repeatedly realized. Yet if search savings were accumulated, a store's ranking in the price distribution would be stable over time and Varian (1980) and Lach (2002) provide theory and evidence respectively that this need not be the case. While infrequently purchased items such as ECDs provide less incentive to accumulate information, their being higher priced provides more incentive to reap search benefits (Lach, 2002).

---

11 ECD online purchases are rare and while differences in prices, between on-line and regular stores have been studied for books by Brynjolfsson and Smith (1999) and Clay et al.(2002), the results of the studies differ as to which is the cheapest on average.

12 The prices were for a single identical model of refrigerator (size, brand type and so forth), size and type of chicken, coffee and flour, being collected from on average 38, 37, 14 and 15 stores respectively. The study by Lach (2000) controlled for product heterogeneity by looking at only one item. This study in this paper covers virtually the whole market, as required by theory, with the heterogeneity of the items being controlled for by the hedonic regressions.
Price dispersion can also be explained simply as a result of price discrimination. Yoskowitz (2002) found price dispersion for the purchase of raw water by companies and municipalities. Price discrimination arises from consumer heterogeneity, the Schultz Panel's (2002) Report on the U.S. CPI noting extensive heterogeneity within a stratum of goods: "...different people buy widely different qualities and brands of goods, often shop at different retail outlets and pay different prices for the same product." (Mackie and Schultz, 2002: 224). The Report attributes the product heterogeneity to consumer heterogeneity of tastes\textsuperscript{13}, differences in age, family composition, geographical location and income, the latter affecting the willingness of an individual to substitute in response to relative prices. Basic marketing teaches managers to segment their markets according to differences in price elasticity by offering products of different quality to different segments and, more particularly, targeting brands to different segments to exploit any consumer surplus. The analysis by quality characteristic and brand picks up such price dispersion.

5.2 The persistence of price dispersion

Over time marginal search costs and benefits will change and consumers will build up their stock of knowledge when costs fall, say on routine shopping trips. With stable prices (no depreciation in knowledge) price dispersion should diminish. Yet while it is in the interests of consumers to build up their knowledge to identify the lowest prices, it is in the interest of stores to prevent this. Varian (1980) explains the persistence of price dispersion by distinguishing between 'shoppers' who pay the lowest price and the remaining consumers with search costs who shop randomly. Price dispersion persists because outlets change their prices (randomly) so as to prevent consumers with search costs from becoming fully informed. Loch (2002) found an intensive process of re-positioning prices over time across stores consistent with Varian's (1980) random pricing model.

5.3 Price dispersion and its mean

As consumers purchase more, they learn more, and it can be argued that price dispersion and the differences between the results from different index number formulas should diminish. Yet the persistence of the failure of the law of one price can also be explained by search cost theory in terms of a relationship between (relative) price dispersion and its mean over time. Classical economic theory at the aggregate level argues that inflation is a monetary phenomenon and should have no effect on relative price distribution. Van Hoomissen (1988) argues that as inflation increases the value of existing information decreases requiring higher search costs to just return to the previous search equilibrium. With price increases, the consumer's understanding of the price distribution is eroded and with higher price increases it is eroded faster. Price dispersion thus persists and varies directly with inflation (Stigler and Kindahl, 1970). However, for infrequently purchased items the store of information should be minimal\textsuperscript{14} and any relationship between the dispersion and mean of prices (or relative price changes) requires an alternative theoretical framework, of which there are several.

Signal extraction models hold that relative price variability will increase with inflation as consumers become less able to distinguish between unanticipated inflationary price variation and relative price changes (Barro (1978), Lucas (1973) and Friedman (1977), extensive empirical work including Vining and Elwertowsky (1976), Parks (1978), Balk (1983),

\textsuperscript{13} Rauh (2002) has shown that with heterogeneity of taste and search costs alone there will be price dispersion.

\textsuperscript{14} Since stores sell a range of infrequently purchased items including fridges, washing machines, dishwashers, stereos, television sets and the like, it might be argued that search information is accumulated on the store, as opposed to the item, giving some credence to the theory.
Domberger (1987), Debelle and Lamont (1993), Reinsdorf (1994) and Silver and Ioannidis (2001)). At higher rates of (unanticipated) mean prices (or relative prices) higher rates of (relative) price dispersion are expected.

Menu cost models find price dispersion occurs when firm’s nominal prices are held constant since there are costs to undertaking price changes.¹⁵ Yet there will come a point (lower bound) when the extent of the change in its real price demands a nominal price adjustment to its upper bound. The resulting staggered price changes give rise to a positive relationship between price dispersion and inflation (Sheshinski and Weiss (1977), Bénabou and Gertner (1993), Ball and Mankiw (1994 and 1995) and Levy and Bergen (1997)).¹⁶

Many ECDs are imported or assembled from imported components. If prices are set in the consumer’s local currency then changes in nominal exchange rates do not affect prices; there is zero pass-through of exchange rate changes. Feenstra and Kendall (1997) and Engel and Rogers (2001) found a significant proportion of price dispersion to be due to incomplete exchange rate fluctuations. If nominal exchange rates fluctuate with inflation, so too might price dispersion.

Serial correlation in price dispersion may arise from sales, in which prices are marked down for a short period only to return to their preceding levels. Hong et al. (2002) argued that serial correlation will be induced for fmgs such as paper towels as consumers build up inventories at sale prices. While this is not applicable to ECDs, prices are reduced at well known sale times and consumers may delay their purchase until such times. This can be modeled as a seasonal effect.

In conclusion, there are reasons to expect price dispersion for ECDs. First, we explain such cross-sectional price dispersion to model and outlet heterogeneity which has been argued to induce search costs. Use is made of hedonic regressions which relate prices to the quality characteristics of these differentiated models, their brand and outlet-type. The residuals are then related to a more direct search cost variable, the proportion of stores in which the model is sold. Second, our concern is whether the dispersion persists or converges and the distributions of the resulting residuals from each month are compared. Third, time series properties and the relationship between price dispersion and (unanticipated) mean prices are investigated.

5.4 Price skewness and its mean

The relationship between the skewness of relative price changes and inflation (Ball and Mankiw, 1995) has also been the subject of study. The aforementioned lower and upper bounds on prices, within which it is not optimal for price changes to be made, come into play here, but with an asymmetry. If, for example, the distribution of desired relative price changes (shocks) is skewed to the right, the average price level will rise since in the short run firms will respond to the large relative price changes (shocks), the smaller ones falling in the range of inactivity. Similarly the price level would fall for negatively skewed desired relative price changes.

¹⁵ Levy and Bergen (1997) show such costs can be substantial.
¹⁶ More recently the focus of such work has been on the relationship between the skewness of relative price changes and inflation (Ball and Mankiw, 1995). While Balke and Wynne (1996) have argued for a similar relationship using a multisector real-business-cycle model, Bryan and Cecchetti (1999) have dismissed the relationship as a statistical artefact due to small sample bias. Our concern with dispersion only has potential statistical small-sample bias if the data are drawn from a skewed distribution. Our use of the population of observations in any event argues against any such bias induced, spurious relationship.
changes. This holds even if the mean of desired prices remains unchanged. If the distribution of desired prices is skewed, a larger variance magnifies the asymmetry in the tails and thus increased the change in the average price level. Ball and Mankiw's (1995) model thus postulates no independent effect for the variance, but a positive interaction effect for the variance with skewness. We thus have a model whereby the mean price is related to its higher moments for goods where frequent and potentially sizable adjustments are made to desired prices.

It is finally worth noting that other explanations have been given for relationships between the mean and its higher moments of relative price changes. Balke and Wynne (1996) have shown how spillover effects from large shocks to a few volatile sectors might generate a positive association between inflation and its higher moments. Bryan and Cecchetti (1996) have argued that the two theories can be distinguished if the periodicity of the data is varied, menu costs being a theory concerned with the short run (see also De belle and Lamont, 1997). While Balke and Wynne (1996) have argued for a similar relationship using a multisector real-business-cycle model, Bryan and Cecchetti (1999) have dismissed the relationship as a statistical artefact due to small sample size bias. Our concern is mainly with dispersion and this only has potential statistical small-sample bias if the data are drawn from a skewed distribution. Our use of the population of observations in any event argues against any such 'small-sample' bias-induced, spurious relationship (see also Ball and Mankiw (1999) and Verbrugge (1998) and for empirical work on skewness see Rodger (2000) and Silver and Ioannidis (1996).

6. Empirical Work: Data and Measures

6.1 Data

The empirical work utilizes monthly scanner data for television sets from January 1998 to March 2002. The scanner data was supplemented by data from price collectors from stores without bar-code readers, though this was a negligible. The observations are for a model of the product, for which there was a transaction, in a given month in one of four different outlet types: multiples, mass merchandisers, independent and catalogue. For example, an observation in the data set for January 1998 includes the unit value (£275.80), volume (5,410 transactions) and quality characteristics (including possession of Nicam stereo and fastext text retrieval facilities) of the Toshiba 2173DB 21 inch television set sold in multiples only. For the 51 months of January 1998 to March 2002 there were 73,020 observations which covered 10.8 million transactions worth £3.9 billion.

6.2 Variables

The variable set on each observation included: Price, the unit value of a model in a month/outlet-type across all transactions. For example, there were 22,485 basic 14" TVs sold in 'mass merchandisers' outlets in December 1998, a seasonal ‘blip’ to meet the demand for Christmas presents. The 22,485 transaction prices are simply summed and divided by the number of transactions to yield the single observation: the price of this model in this store this month - £97.50 (see Balk, 1996 for the statistical properties of unit values). Volume is the sum of the transactions during the period. Many of the models sold in any month had relatively low sales. There were 38 brands—37 dummy variables benchmarked on Sony; the characteristics included (i) size of screen—dummy variables for about 19 screen sizes; possession of (ii) Nicam stereo; (iii) wide screen; (iv) on-screen text retrieval news and information panels from broadcasting companies, in order of sophistication: teletext, fastext
and top fasttext – 3 dummy variables; (v) reception system—6 types; (vi) monitor style; (vii) with Dolby Pro, Dolby SUR./DPL., Dolby Digital sound—3 dummy variables; (viii) Flat & Square, Super-Planar tube—2 dummy; (ix) s-vhs socket; (x) with satellite tuner, analogue/digital—2 dummy variables; (xi) digital; (xii) with DVD playback or DVD recording—2 dummy variables; (xiii) with rear speakers; (xiv) without PC-internet/PC+internet; (xv) with real flat tube; (xvi) 100 hertz, doubles refresh rate of picture image; (xvii) vintage and DIST—the percentage of stores in which the model was sold.

Outlet types are multiples, mass merchandisers, independents and catalogue. The hedonic regressions were based on about 100 variables.

6.3 Measures of dispersion and average prices

6.3.1 Weights

Our concern with differences between (weighted) upper-level indices gives rise to an empirical concern with difference between weighted variances. The empirical micro literature on price dispersion and the law-of-one price is dominated by the use of unweighted measures and include Clay et al. (2002), Lach (2002), Hong et al. (2002), Engel and Rodgers (2001), Cohen (2000), Sorensen (2000) and Beaulieu and Mattey (1999). But this use is only because weighted measures are often unavailable and this is a serious shortcoming of existing studies. An increase in price, for example, in one store will lead, ceteris paribus, to a fall in quantity and weight, unweighted measures exaggerating price dispersion. Since Annexes 4 and 5 have shown that the extent of any difference between Laspeyres and Paasche and the bias in the Young index to be at least in part dictated by differences in the weighted dispersion in prices over time, results for weighted dispersion measures are calculated and analyzed in the empirical section.

Our concern with differences between (unweighted) lower-level indices gives rise to an empirical concern with difference between unweighted variances. While the empirical work can be rightly undertaken in this manner, three points are worth noting. First, that prices are often collected by statistical agencies (at least initially) for major selling/typically purchased items for the CPI sample. If, as for the U.S. CPI, sampling is with probability proportionate to value share in the base period, our concern is with weighted indices. Under such sampling the expected value of a Carli index is a Laspeyres index (Balk, 2002) and the results for weighted indices apply. Second, and related to this, other sampling systems might be replicated using the data. In many countries ‘typically purchased’ items (at least in the price reference period) are sampled and unweighted variances based on cut-off sampling may be more appropriate for the empirical work.

Third, the analysis of dispersion at the weighted level is motivated by the economic theories of search cost, menu costs and signal extraction models. The differences between lower-level indices were shown in section 4 to be concerned with differences in unweighted price dispersion. Thus while the empirical work can ascertain patterns of unweighted price dispersion to explain the differences between these formulae, it may be argued that it should not draw on theory which relates to market behavior which includes prices and quantities. Against this such theories can be seen as theories of market failure in price setting: menu cost

17 There is some variability in this over time with DVD, rear speakers, top fasttext, Dolby digital and SUR/>DPL/> (as opposed to just Dolby Pro sound), 100 hertz and integrated PC not being used until January 2000; 11 variables excluded as not being relevant.

18 Spurious correlations between dispersion and its mean have been argued by Bryan and Cecchetti (1999) to arise when unweighted measures are used.
theory predicts that retailers will have costs of price adjustment and not undertake such adjustments unless the price change is outside of some bounds, thus leading to price dispersion—a case that can be argued for all models of a good irrespective of their sales quantities. Similarly if a proportion of the population has search costs some retailers can enjoy a surplus on some/all of their models which will again lead to price dispersion irrespective of sales quantities. And again mistakes in anticipating inflation will lead to erroneous decisions by economic agents, an argument that will lead to price dispersion. In all cases the welfare effects require quantities to be taken into account. But all that we require here is that the analysis of unweighted changes in the dispersion of prices be motivated by the aforementioned theoretical frameworks.

Finally, we have a genuine interest in the behavior of a heterogeneity-controlled Dutot index since the Dutot index has fine axiomatic only being dismissed from the analysis because it fares badly when the items are relatively heterogeneous. The analysis of price dispersion as a means to minimize such heterogeneity by statistical mechanisms as opposed to the selection of a limited matched sample is of interest.

6.3.2 Parametric measures of absolute and relative dispersion

The weighted standard deviation and coefficient of variation are given as absolute and relative parametric measures respectively by:

\[
SD_w = \left( \sum_{m=1}^{M} w_m (p_m - \bar{p}_w)^2 \right)^{1/2} \quad \text{and} \quad CV_w = \frac{SD_w}{\bar{p}_w} \tag{15}
\]

where the aggregation is over \( m = 1, \ldots, M \) models in a given month and outlet-type, \( w_i \) is their sales (value) share, \( p_m \) is the price and \( \bar{p}_w \) the mean price = \( \sum_{m=1}^{M} w_m p_m \). Unweighted measures are similarly defined. Non parametric/robust measures are not used since the index number formulae relationships are based on parametric measures.

7. Empirical Work: Results

For the empirical work two things are of interest. First, the extent of price dispersion in any period. Price collectors are required to collect prices of similar items for defined ‘representative items.’ The more similar the items, the less the dispersion in prices and the closer together the results of the different indices. Search cost theory tells us that the product heterogeneity may itself be a device to increase search costs and allow further ‘real’ dispersion. An understanding of the differences between formulae thus requires an understanding of the heterogeneity of prices and the derivation of estimates of heterogeneity-controlled prices and their dispersion. Second is the explanation of changes in dispersion over time. This was shown in sections 2 and 3 to be at the core of explaining differences between index number formulae and search cost, menu cost and signal extraction models will be used to underlie this empirical work. It was also shown in section 2 that the Dutot index is particularly sensitive to product heterogeneity, since it fails the commensurability test. This work will also compare Dutot, Carli and Jevons (matched) indices for heterogeneity-controlled prices as against uncontrolled, raw prices.

In both instances highly detailed scanner data from the bar-code readers of retailers will be used. Such data cover the market of transactions and provide information on the price-determining quality characteristics responsible for much of the price dispersion.
7.1 Explaining price variation

Table 1 provides summary statistics on raw price dispersion. First, the extent of dispersion in raw prices is substantial. The coefficients of variation (CV = standard deviation (SD)/\( \bar{x} \)) averaged 0.85. Second, the standard deviation can be seen to have increased substantially: by just over 20 percent for the unweighted measure. Yet the unweighted CV is relatively stable and shows much of this increase in price dispersion to be accounted for by inflation. Table 1 shows weighted dispersion increased by 40 percent compared with the 20 percent for the unweighted dispersion. Commonly purchased models have quite disparate price movements. Yet the weighted CV also increased substantially over the first three years, inflation not being an immediately obvious explanation for the increased dispersion in raw prices, but subsequently fell.

Measuring price dispersion under product differentiation requires controls for the brand and technical characteristics of the model and, since different outlet-types provide different services, the outlet-types in which the model is sold. Table 2 is based on a hedonic regression for observation \( m \) in period \( t \):

\[
\ln \text{Price}_{m}^{t} = \beta_0 + \sum_{i=1}^{11} \beta_i \text{Month}_{m}^{t} + \gamma \text{Year}_{m}^{t} + \sum_{j=1}^{48} \delta_j \text{Charac}_{m}^{t} + \\
\sum_{k=1}^{37} \alpha_k \text{Brand}_{m}^{t} + \sum_{l=1}^{3} \theta_l \text{Outlettype}_{m}^{t} + \varepsilon_{m}^{t}
\]  

(16)

Table 2 provides a nested decomposition of price variation to explain some of the existing price dispersion and identify the remaining dispersion—the heterogeneity-controlled prices—in terms of the residuals of the regression. The observations for such a pooled regression were the prices, characteristics and brands of individual models of TVs in a specific outlet-type over the 51 months of January 1998 to March 2003—just over 73,000 observations on 37 brand dummies, 3 outlet-type dummies, 19 screen size dummies, 6 tube-type dummies and 23 further characteristics as outlined in sections 6.1 and 6.2 above. The coefficients were almost invariably statistically significant and their signs accorded with a priori expectations.\(^{19}\) The \( R^2 \) for the estimated equation (3) shows that over 90 percent of variation in price was explained by the model. Month and time provided little explanatory power; it was product heterogeneity vis-à-vis product characteristics, brand and outlet-type variation that accounted for most of the price variation. Multicollinearity precludes our assigning variation separately to brands, characteristics or outlet-types, however, characteristics do most of the work: a regression on month, trend and brands only accounted for 0.35 percent of variation (\( R^2 = 0.0035 \)) and similarly low for month, trend and outlets. The regression successfully controlled prices for the heterogeneity of their features. The residuals, \( \varepsilon_{m}^{t} \), are estimates of heterogeneity-controlled prices. Mean variation in prices was reduced by over 50 percent by the regression, and the standard deviation of the heterogeneity-controlled prices was about one-fifth that of the actual price dispersion (Table 2). Bear in mind it is not just the variation in technical characteristics, brands and the services from outlet-types that explain the price dispersion. If this were the case the law-of-one-price should hold for the residuals. The sheer multiplicity of models hinders search giving rise to price dispersion, though we return to this later.

\(^{19}\) Details available from authors.
While characteristics, brands and outlet-types were found to be one source of price dispersion, we now consider a more explicit modeling of search to explain price dispersion within months. In section 6.2 DIST was defined for each model as the percentage of (television) shops stocking or selling that model. Unfortunately such data were only available for the first 24 months and we confine this analysis to this period. A model sold in fewer shops is less likely to have a comparable model available in any search carried out, thus precluding direct comparisons and allowing some premium margin to be charged, an expected negative sign to help explain the remaining variation. To establish whether heterogeneity–controlled price variation can in part be explained by this search cost variable we first estimated separate hedonic regressions of the form in equation (16) for each month. The mean $R^2$ for the 24 regressions was 0.89 with a minimum of 0.85 (and for weighted regressions 0.96 and 0.95 respectively), normality of residuals (Jacques-Bera ) was by and large rejected for the OLS and WLS estimators, though the null of homoskedasticity was generally not rejected for both estimators (Breusch-Pagan). We second, regressed heterogeneity-controlled prices on DIST:

$$\ln z_m^t = \alpha_0^t + \alpha_1^t \text{DIST}_m + \nu_m^t$$

where $\tau_m^t$ are the (exponents of the) residuals from hedonic regressions akin to equation (16) but run each month. The results are in Table 3, though only for the first 24 months such data not subsequently being part of routine data provision by the supplier. For an OLS estimator Table 3 shows a consistently positive, statistically significant relationship. The more stores a model is sold in the higher its heterogeneity–controlled price. This runs contrary to our expectation from search cost theory. One explanation for the positive sign is that the OLS results give equal weight to models with very few sales, that is to models near the end of their life cycle. Such models have been shown to have (Silver and Heravi, 2002a) low prices relative to their specifications and are no longer sold in all the stores they were previously sold in. Similarly Silver and Heravi found new more widely distributed models to have a premium snob price relative to their specifications. The WLS estimator for equation (17) virtually ignores the low selling old models with their low quality-adjusted prices to give more ambiguous results for most of the period, except for towards the end when consistently

---

20 Even for this period data on DIST were more limited than the rest of the data, comprising nearly 22,000 observations (as opposed to 29,000).

21 Though for the WLS estimator the null of the (component) test for skewness (Davidson and MacKinnon, 1993) was not rejected in one-third of the months—details available from authors.

22 The Breusch-Pagan test failed to reject the null of homoskedascity for all but three months using the OLS estimator, though rejected in one-third of the months using WLS (by value)—details available from authors.

23 This deviation from the normality assumption and some heteroskedasticity may not permit correct inferences to be drawn on the coefficients. However, a heteroskedasticity-consistent covariance matrix estimator (HCCME) was used following White (1980) to allow asymptotically correct tests to be undertaken. A wild bootstrap estimator is commonly applied to models with heteroskedastic and skewed residuals due to small-sample bias in the HCCME. Davidson and Flachaire (2001) show that the wild bootstrap is only necessary to alleviate small-sample bias; the HCCME estimator is appropriate for the large sample tests in this study.

24 DIST might well have been included in the hedonic regression equation except for a priori expectations that the residuals, heterogeneity–controlled prices, may be correlated with it, as indeed we found. Modern stores hold relatively little stock with advanced replacment ordering systems unrelated to price. The level of stocks was used each month as instruments for instrumental variable hedonic regressions that included DIST. The results from such regressions found DIST significant at a 5% level in 6 of the 24 months and at a 10% level in 8 months. In all such regressions the coefficients on DIST had a positive sign. A series of Hausman test for each month found the null of no relationship between the errors and DIST was rejected in 50% of the months using OLS but in only 20% of the months using WLS.
positive results are apparent from Table 3. When we look at better selling models there may be economies of scale in distribution lowering price dispersion.25 Thus for the difference between unweighted index number formulae, for which our concern is unweighted OLS results, heterogeneity-controlled price dispersion can be further explained by DIST.

Thus three things have been found to explain price variation and thus have the potential to control price dispersion within months: (i) the technical features, brand and outlet type of the models, (ii) for unweighted measures, the extent to which models are sold in different stores, the fewer the stores the lower the (quality-adjusted) prices and (iii) the use of weights.

### 7.2 Changes in price variation

Section 4 showed that it is the changes in dispersion that underlies the differences between the results from index number formulae. Table 4 and Figure 1 shows the unweighted and weighted standard deviation of residuals over time. Note that the residuals are estimated from separately estimated regressions each month and thus are normalized by their respective semi-log hedonic regression in each month to have a mean of unity. As such the standard deviation of the residuals is the de facto coefficient of variation; the standard deviation normalized by the mean. This is our measure of dispersion. The derivations of the differences between formulae in the annexes are based on normalized dispersion and the measurement and modeling naturally follows from this. Table 4 and Figure 1 shows an increase in such dispersion of nearly 100 percent over the 51 months, and this is after being controlled for heterogeneity and mean (quality-adjusted) inflation. Note that quality-adjusted mean prices were falling over this period,26 and that this should contribute to a fall in the standard deviation. So the adjustment by the mean implicit in the measure increases dispersion over time to account for this. But the resulting series can be seen to trend upwards; this relative concept of dispersion is increasing even after accounting for the fall in the mean. Also shown in Table 4 and Figure 1 is substantial volatility in the series. So dispersion is increasing accompanied by volatility, but so too must the difference between the formulae and this is after we control for heterogeneity and average price changes. So can we explain such changes?

The residuals from WLS hedonic regressions have been weighted by their relative expenditure shares to reflect their economic importance. Table 4 shows this weighted heterogeneity-controlled price dispersion to be about two-thirds that of its unweighted counterpart, and it too shows a striking increase, over 75 percent. Figure 1 shows much more volatility in the unweighted (OLS-based) measure, especially towards the end of the series.27 Thus weighting reduces dispersion, the change in dispersion, and thus the differences between the results from formulae, and volatility of dispersion. Though for weighted and unweighted

---

25 To corroborate the influence of weights on the relationship the OLS estimates were run for a sample of models selling 200 or more in an outlet-type in any month. There was no relationship for dispersion on DIST for these larger selling models in the 24 months.

26 The measurement of the change in mean heterogeneity-corrected prices is problematic since the expected values of the residuals, or their logs, will be zero or unity by construction and not vary over time. A chained index has been calculated using hedonic regressions with the same specification as (3) above, but the regressions are based on sets of two successive monthly stacked data where Month is a dummy variable which takes the value of one if it is the second month and zero otherwise, and the estimated coefficients on Month are linked by successive multiplication to form a chained index of the heterogeneity-controlled mean price. The index (results available from authors) fell by about 30 percent over the period.

27 There is also a shift at January 2000 but this may be due to a change in the format of the data and slightly more detailed variable definitions available from this month onwards.
measures the evidence is of a substantial increase in dispersion even after we have explained technical, brand and outlet-type variation.

7.3 Explaining changes in residual price dispersion

In this section we seek to explain changes in residual price dispersion over time both in relation to its time series properties, search cost related theories, signal extraction models (its anticipated and unanticipated mean) and menu cost theory.

Search cost theory argues that as inflation increases, the value of existing information decreases requiring higher search costs just to return to the previous search equilibrium (Van Hoomissen, 1988). It would predict increased dispersion from an increasing mean, but also increasing dispersion from a falling mean since in both cases the stock of knowledge would depreciate as average prices change. Bearing in mind average (heterogeneity-adjusted) prices were falling (ff.26) it would predict that dispersion normalized on the mean would rise. We test for trend stationarity in our mean-adjusted measure of dispersion and take non-rejection of a unit root as evidence to reject stationarity. Unit root test results are given in Table 5 and although not conclusive, the weight of evidence is towards accepting I (1) as expected. The drift from Figure 1 is upwards in both cases confirming a positive trend in dispersion over the mean. If search was frequent then consumers would become more knowledgable and dispersion would decrease. However, given the frequency of search for consumer durables is more limited than for frequently purchased goods the finding is not surprising.

Search cost theory would also predict that as the total number of models of television sets on sale increases, there would be higher search costs and thus dispersion, and thus a negative sign on the estimated coefficient for ‘number of models’ in a regression of dispersion on the latter. This would hold in spite of their infrequent purchase.

Signal extraction models require an indicator of unanticipated inflation. Anticipated inflation was predicted from ARIMA models for unweighted and weighted H-C means. AR(1) processes fitted the series best in both instances. Unanticipated inflation (H-C UNANTICIPweighted and H-C UNANTICIP.unweighted) was generated as the residuals. To test signal extraction models heterogeneity-controlled price dispersion was regressed on unanticipated inflation. The estimated coefficient on unanticipated inflation should have a positive sign following signal extraction theory that increased dispersion arises from the inability of economic agents to properly anticipate inflation, such inability increasing with inflation. However, the evidence of a positive relationship is far from conclusive. For example, Hesselman (1983), and Silver (1988) found negative relationships for the UK. Silver (1988) argued that the coefficient may have a negative sign as economic agents become more cautious in their price-setting and price-taking under uncertainty. Buck (1990), found negative association for Germany (using 19th century data), but positive for the US using data for the same period. Reinsdorf (1994) found a negative relationship for 65 categories of goods in 9 US cities, though this was for price levels. Reinsdorf (1994) found his explanation for a negative relationship from consumer search cost theories with unexpected inflation inducing more search due to consumers' incomplete information about price distributions.

28 We also differenced and subsequently repeated the unit root tests to ensure I(2) was rejected in favour of I(1). The p-values for (augmented) weighted asymmetric tau tests on weighted and unweighted differenced dispersion were 0.00578 and 0.00249 respectively rejecting I(2).
Silver and Ioannidis (2001) found negative relationships for a range of European countries using a consistent methodology.

*Menu cost theories* can be tested by examining the relationship between the dispersion and the mean for data with a bimonthly frequency and that with a monthly frequency. Larger price changes and dispersion should materialize in the latter, though the limited time series here precludes this study for the time being. More straightforward is to include a variable on the proportion of *models* (for unweighted dispersion) or *expenditure* (for weighted dispersion) in catalog stores (*CATALOG*... for unweighted dispersion) and *CATALOG*...respectively), the latter having higher menu cost than other stores. A positive sign would be expected on the estimated coefficient.

Table 6 shows the regression results to explain variation in changes in weighted and unweighted heterogeneity-controlled residuals normalized on their means. That Table 5 found some series to be I(1) held out the possibility of long-run cointegrating relationships. Engle-Granger (tau) cointegration tests were undertaken finding non-rejection of the null whichever variable was chosen as the dependent variable. The results for using the unweighted and weighted dispersion as the dependent variables respectively were:

<table>
<thead>
<tr>
<th>Dep.Var.</th>
<th>TestStat</th>
<th>P-value</th>
<th>Num.lags</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDunweighted</td>
<td>-3.08894</td>
<td>0.38296</td>
<td>6.00000</td>
</tr>
<tr>
<td>SDweighted</td>
<td>-3.65520</td>
<td>0.14725</td>
<td>4.00000</td>
</tr>
</tbody>
</table>

Non-rejection of null that the residuals of the cointegrating relationship has a unit root is evidence of noncointegration precluding estimation of an error correction model. However, where the null of unit roots was rejected first differences were used in the regression. For unweighted dispersion the coefficient on the ‘number of models’ sold is positive and statistically significant as predicted from search cost theory. No corresponding result is found for weighted dispersion. Unanticipated inflation has the expected negative coefficient for unweighted and weighted dispersion the relationship being more consistent for weighted dispersion there being evidence of it being affected by multicollinearity. There is also evidence of a negative relationship for both weighted and unweighted dispersion for the incidence of models/sales in catalog stores. We argued for a positive relationship from a menu cost stance on the grounds that the costs and resulting delays in making adjustments would lead to relatively large changes when they took place. Catalog stores are more prone to such delays. Yet during the print run of the catalog there should be less fluctuations and it may be that this overshadowed the price spikes from the adjustments. More generally while there is evidence of multicollinearity, Table 6 shows some explanatory power of economic variables in explaining price dispersion.

8. Index number implications

Since differences arising from index number formulae can be significant and since they are to varying extents determined by the variance in prices and its changes, the motivation to explain such variation is appropriate, especially since there is an economic theory of price dispersion to ground the work in. A number of implications arise from the study.

First, the discrepancy between elementary indices increases as the dispersion in prices increases. Since much of the variation can be explained by the heterogeneity of the brand, technical characteristics of the good, the requirement is for well-defined, item specifications (albeit at the cost of coverage - Silver and Heravi, 2002b). Bias in the Dutot index arises from
such heterogeneity, in spite of its otherwise good axiomatic properties. We calculated a heterogeneity-controlled (matched) Dutot index using the residuals from the hedonic regression and compared it with a Dutot index without such controls, as well as Carli and Jevons matched indices. The results are in Table 7. First, formula does matter; there is 15 per cent fall according to Carli and Dutot, but 20 percent fall given by the Jevons index. Second, the Dutot index, like Jevons, performs well from the test approach, but suffers from its failure of the commensurability test particularly with regard to product heterogeneity. The bias (as measured here) of the Dutot index against the heterogeneity-corrected Dutot index is about 1.5 percentage points upwards over the period.

Second, there was found to be considerable price dispersion in this product area which is not unusual for highly differentiated consumer durables (Table 1). Brand, characteristics and outlet-type together explained much of such variation (Table 2). Minimizing price variance for price index number compilation requires either the use of hedonic indices or detailed specifications for a selection of 'representative' items and care in judging replacements to be 'comparable' when an item goes missing if its characteristics are different. When a model is missing the price collector may judge another model to be of comparable quality and compare its prices with those of the old model. There is too much price variation associated with product heterogeneity to be lax about any leniency in such selections.

Third, as soon as weights were applied the dispersion was reduced (Table 2) and thus the difference between formulae. Selections of more popular models serve to not only make the index more representative, but also to reduce the disparity between the results from different formulae.

Fourth, for unweighted indices, product heterogeneity aside, models sold in more stores (DIST) have higher prices (Table 3). Selection of items should take into account a model's coverage of stores for the sample to be representative.

Fifth, we found an increase in dispersion of nearly 100 percent over the 51 months, and this was after being controlled for heterogeneity and mean (quality-adjusted) inflation. Such differences lead to formulae differences and require explanation (Table 4). Some of the explanation could be identified via the trend, the upwards drift in the series. The drift was more volatile and accentuated for weighted dispersion than unweighted dispersion (Figure 1).

Sixth, differences in dispersion over time accord with aspects of search cost theory, menu cost theory and signal extraction models. Such frameworks were shown to explain some of the variation in dispersion over time (Table 6) and thus the increasing differences between the results of index number formulae. This applied both to weighted dispersion (formulae) and unweighted dispersion (formulae), though more successfully to the former.
Annex 1: The relationship between Dutot and Jevons indices

The first approximate relationship that will be derived is between the Carli index $P_C$ and the Dutot index $P_D$. For each period $t$, define the arithmetic mean of the $M$ prices pertaining to that period as follows:

\[(A1.1)\]

\[P^*_t = \frac{1}{M} \sum_{m=1}^{M} p^*_m \quad t = 0, 1.\]

Defining the multiplicative deviation of the $m^{th}$ price in period $t$ relative to the mean price in that period, $e'_m$, as:

\[(A1.2)\]

\[p'_m = P^*_t \left(1 + e'_m\right); \quad m = 1, \ldots, M; \quad t = 0, 1.\]

Note that (A1.1) and (A1.2) imply that the deviations $e'_m$ sum to zero in each period; i.e., we have:

\[(A1.3)\]

\[\sum_{m=1}^{M} e'_m = 0; \quad t = 0, 1.\]

The Dutot index can be written as the ratio of the mean prices, $p^*/p^0*$; i.e., we have:

\[(A1.4)\]

\[P_D(p^0, p^t) = \frac{p^*_t}{p^0_*}.\]

Substitute equations (A1.2) into the definition of the Jevons index, (3) using (A1.4):

\[(A1.5)\]

\[P_J(p^0, p^t) = \prod_{m=1}^{M} p_m^0 \left(1 + e'_m\right) = \prod_{m=1}^{M} p_m^t \left(1 + e'_m\right) = P_D(p^0, p^t) f(e^0, e').\]

where $e^t = [e^t_1, \ldots, e^t_M]$ for $t = 0$ and $1$ and the function $f$ is defined as follows:

\[(A1.6)\]

\[f(e^0, e') = \prod_{m=1}^{M} \left(\frac{1 + e'_m}{1 + e^0_m}\right).\]

Expand $f(e^0, e')$ by a second order Taylor series approximation around $e^0 = 0_M$ and $e' = 0_M$. Using (A1.3), it can be verified\(^{29}\) that we obtain the following second order approximate relationship between $P_J$ and $P_D$:

\[(A1.7)\]

\[P_J(p^0, p^t) \approx P_D(p^0, p^t) [1 + (1/2M)e^0 \cdot e^0 - (1/2M)e^0 \cdot e'] + P_D(p^0, p^t) [1 + (1/2)\text{var}(e^0) - (1/2)\text{var}(e')].\]

\(^{29}\) This approximate relationship was first obtained by Carruthers, Sellwood and Ward (1980; 25).
where \( \text{var}(e^t) \) is the variance of the period \( t \) multiplicative deviations; i.e., for \( t = 0,1 \). Since \( e^{t*} = 0 \) using (A1.3):

\[
(A1.8) \quad \text{var} (e^t) = \frac{1}{M} \sum_{m=1}^{M} (e^t_m - e^{t*}_m)^2 = \frac{1}{M} \sum_{m=1}^{M} (e^{t*}_m)^2 = \frac{1}{M} e^t e^t
\]

Diewert notes that "Under normal conditions, the variance of the deviations of the prices from their means in each period is likely to be approximately constant and so under these conditions, the Jevons price index will approximate the Dutot price index to the second order." He footnotes 'normal conditions' with the caveat that: "If there are significant changes in the overall inflation rate, some studies indicate that the variance of deviations of prices from their means can also change. Also if \( M \) is small, then there will be sampling fluctuations in the variances of the prices from period to period." Our concern is with former.
Annex 2: The relationship between Jevons and Carli indices

Both of these indices are functions of the relative prices of the $M$ items being aggregated. This fact is used in order to derive some approximate relationships between these indices.\textsuperscript{30} Define the $m$\textsuperscript{th} price relative as $r_m = \frac{p_m^i}{p_m^0}$ for $m = 1, \ldots, M$ and the arithmetic mean of the $m$ price relatives as

\begin{equation}
A2.1 \quad r^* = \frac{1}{M} \sum_{m=1}^{M} r_m = P_C(p^0, p')
\end{equation}

where the last equality follows from the definition (3) of the Carli index. Finally, define the deviation $e_m$ of the $m$\textsuperscript{th} price relative $r_m$ from the arithmetic average of the $M$ price relatives $r^*$ as follows:

\begin{equation}
A2.2 \quad r_m = r^* (1 + e_m); \quad m = 1, \ldots, M.
\end{equation}

Note that (A2.1) and (A2.2) imply that the deviations $e_m$ sum to zero; i.e., we have:

\begin{equation}
A2.3 \quad \sum_{m=1}^{M} e_m = 0
\end{equation}

Now substitute equations (A2.2) into the definitions of $P_J$, above, in order to obtain the following representations for these indices in terms of the vector of deviations, $e = [e_1, \ldots, e_M]$:

\begin{equation}
A2.6 \quad P_J(p^0, p') = \prod_{m=1}^{M} q_m r_m = r^* \prod_{m=1}^{M} q_m [1 + e_m] = P_C(p^0, p') \prod_{m=1}^{M} q_m [1 + e_m]
\end{equation}

The multiplicative factor which defines the difference between the Jevons and Carli indices is approximated using a second order Taylor series expansion around the point $e = 0$ by:

\begin{equation}
A2.7 \quad \prod_{m=1}^{M} q_m [1 + e_m] \approx 1 - (1/2M) e \cdot e = 1 - (1/2)\text{var}(e);
\end{equation}

Thus, since var($e$) must be positive, to the second order the Carli index $P_C$ is argued to exceed the Jevons index, by $(1/2)r^*\text{var}(e)$, which is one half the variance of the $M$ price relatives $p_m^i/p_m^0$.

\textsuperscript{30} It is very straightforward to do the same with a harmonic mean of price relatives.
Annex 3: The relationship between Dutot and Carli indices
Following Diewert (1995a, 27)

\[(A3.1)\]

\[
P_D(p^0, p^t) = \sum_{m=1}^{M} (r_m^t) \cdot p_m^0 / \sum_{m=1}^{M} p_m^0 = \sum_{m=1}^{M} r_m^t / m + \sum_{m=1}^{M} \left[ p_m^0 / \sum_{m=1}^{M} p_m^0 - 1 / M \right]
\]

\[= P_C(p^0, p^t) + \sum_{m=1}^{M} r_m^t \left[ p_m^0 / \sum_{m=1}^{M} p_m^0 - \left( \sum_{m=1}^{M} p_m^0 / M \right) / \sum_{m=1}^{M} p_m^0 \right]
\]

which is \(P_C(p^0, p^t)\) plus the covariance of normalized \(r_m^t\) and \(p_m^0\). The correlation coefficient between price relatives and base period prices is defined as \(\rho(r_m^t, p_m^0)\) which is equal to the covariance of \((r_m^t, p_m^0)\) divided by the product of the variances of the individual variables. Therefore:

\[(A3.2)\]

\[
P_D(p^0, p^t) = P_C(p^0, p^t) + m[\text{var}(r_m^t) \cdot \text{var}(p_m^0)]^{1/2} \rho(r_m^t, p_m^0)
\]

Since the variances must be positive, the sign of \(\rho\) determines which of these formulae will give results with higher values. The correlation would be expected to be negative as higher base period prices for similar items should have lower price increases. Thus \(P_C(p^0, p^t)\) is expected to exceed \(P_D(p^0, p^t)\). The two formulae will give the same results if the \(\text{var}(r_m^t) = 0\), that is, all price relatives are the same, or the \(\text{var}(p_m^0) = 0\), all (normalized) base period prices are the same, or if \(\rho(r_m^t, p_m^0) = 0\), there is no correlation between price relatives and base period prices. As either of these depart from zero, the difference between the results from the two formulae will increase. Any difference due to the above factors can be seen to be magnified as \(m\), the number of prices increases.

Note that the relationships in this section have been phrased as long-run ones, between periods 0 and \(t\). As time progresses it might be expected that the correlation \(\rho(r_m^t, p_m^0)\) weakens though the variance of price relatives \(\text{var}(r_m^t)\) may increase.
Annex 4: The relationship between Laspeyres and Paasche

The results are due to Bortkiewicz (1922, 1924) reproduced in Allen (1975, 62-64).

The weighted means of price and quantity relatives are Laspeyres price and quantity index numbers:

\[(A4.1) \quad P_L = \sum_{m=1}^{M} w_m^0 \frac{P_m}{P_m^0} \quad \text{and} \quad Q_L = \sum_{m=1}^{M} w_m^0 \frac{q_m}{q_m^0} \quad \text{where} \quad w_m^0 = p_m^0 q_m^0 \]

Thus the notation defines Laspeyres price and quantity indices to respectively be \(P_L\) and \(Q_L\). Paasche price and quantity indices can be similarly defined and denoted as \(P_P\) and \(Q_P\), and the value index as \(V^{0t}\). It is easily demonstrated that \(P_L \times Q_P = V^{0t}\) and \(P_P \times Q_L = V^{0t}\) so that:

\[(A4.2) \quad \frac{P_P}{P_L} = \frac{Q_P}{Q_L} \]

This is the ratio of Paasche to Laspeyres formulae we seek to explain. The weighted variances for Laspeyres price and quantity indices defined in (1) are:

\[(A4.3) \quad \sigma_p^2 = \sum_{m=1}^{M} w_m^0 \left\{ \frac{P_m}{P_m^0} - P_L \right\}^2 / \sum_{m=1}^{M} w_m^0 \quad \text{and} \quad \sigma_q^2 = \sum_{m=1}^{M} w_m^0 \left\{ \frac{q_m}{q_m^0} - Q_L \right\}^2 / \sum_{m=1}^{M} w_m^0 \]

The weighted covariance times \(\sum w_m^0\) is:

\[
\sum_{m=1}^{M} w_m^0 \left\{ \frac{P_m}{P_m^0} - P_L \right\} \left\{ \frac{q_m}{q_m^0} - Q_L \right\} \\
= \sum_{m=1}^{M} w_m^0 \frac{P_m q_m}{P_m^0 q_m^0} - P_L \sum_{m=1}^{M} w_m^0 \frac{q_m}{q_m^0} - Q_L \sum_{m=1}^{M} w_m^0 \frac{P_m}{P_m^0} + P_L Q_L \sum_{m=1}^{M} w_m^0 \\
= \sum_{m=1}^{M} w_m^0 \frac{P_m q_m}{P_m^0 q_m^0} - P_L Q_L \sum_{m=1}^{M} w_m^0 
\]

We divide by \(\sigma_p \sigma_q \sum w_m^0\) to get the weighted correlation coefficient \(r\) between price and quantities:

\[(A4.4) \quad r = \frac{\sum_{m=1}^{M} w_m^0 \frac{P_m}{P_m^0} \frac{q_m}{q_m^0}}{\sigma_p \sigma_q \sum_{m=1}^{M} w_m^0} \frac{P_L}{Q_L} \frac{P_L}{P_L Q_L} \frac{Q_L}{Q_L} \]

Using \(w_m^0 = p_m^0 q_m^0\) and (A2.1)

\[(A4.5) \quad \sum_{m=1}^{M} w_m^0 = \sum_{m=1}^{M} \frac{P_m q_m}{P_m^0 q_m^0} = V^{0t} = P_L Q_L \]
This brings in the Paasche price index. Substituting in (A4.4):

\[
(A4.6) \quad r = \frac{P_P Q_L - P_L Q_L}{\sigma_p \sigma_q} = \frac{P_L Q_L}{\sigma_p \sigma_q} \left( \frac{P_P}{P_L} - 1 \right)
\]

Rearrangement gives the required common ratio (A2.1) of the Paasche to the Laspeyres index numbers:

\[
(A4.7) \quad \frac{P_P}{P_L} = \frac{Q_P}{Q_L} = 1 + r \frac{\sigma_p \sigma_q}{P_L Q_L}
\]

To interpret (A4.5), note that the operative terms are the coefficient of correlation \( r \) between price and quantity relatives, multiplied by two coefficients of variation, i.e. the standard deviations from (A4.3) as ratios of the means (A1.2). The coefficients of variation are positive so that the sign of \( r \) to fix the direction of the divergence of the Paasche from the Laspeyres index. The Paasche index is the greater if \( r > 0 \) and the Laspeyres index if \( r < 0 \). From (2) it follows that the direction of the divergence of the quantity index numbers is the same as that of the price index numbers.

The extent of the divergence, in whichever direction it is, depends partly on the strength of the correlation \( r \) and partly on the dispersion of the price and quantity relatives as shown up in the coefficients of variation.
Annex 5: The relationship between Young and rectified Young

While the Laspeyres index is well-known, it is not used in index number compilation by statistical agencies. This is because expenditure weights for the price reference period 0, for a comparison between periods 0 and t, take time to be compiled and relate to an earlier weight reference period b. The resulting Young index is:

(A5.1) \[ P_Y = \sum_{m=1}^{M} s_m^b \left( \frac{p'_m}{p_m^0} \right) \]

where \[ s_m^b = \frac{p_m^b q_m^b}{\sum_{m=1}^{M} p_m^b q_m^b} \]; \[ m = 1, ..., M \].

A problem with this index is that it fails the time reversal test: the index between periods 0 and t exceeds its time antithesis—it has an upwards bias. Diewert (2003) compares the Young index with its reciprocal to ascertain the bias and finds that to the accuracy of a certain second order Taylor series approximation, the following relationship holds between the direct Young index, \( P_{Y(0,t)} \) and its time antithesis, \( P_{Y(0,t)} \):

(A5.2) \[ P_{Y(0,t)} \approx P_{Y(0,t)} + P_{Y(0,t)} \text{ var (e) } \]

where \[ \text{ var(e) } = \sum_{m=1}^{M} s_m^b \left( e_m - e^* \right)^2 \]

The deviations \( e_m \) are defined by \( 1+e_m = r_m/r^* \) for \( m = 1, ..., M \) where the \( r_m \) and their weighted mean \( r^* \) are defined by:

(A5.3) \[ r_m = \frac{p'_m}{p_m^0} \]; \[ m = 1, ..., M \];

(A5.4) \[ r^* = \sum_{m=1}^{M} s_m^b r_m \]

which turns out to equal the direct Young index, \( P_Y \). The weighted mean of the \( e_m \) is defined as

(A5.5) \[ e^* = \sum_{m=1}^{M} s_m^b e_m \]

which turns out to equal 0. Hence the more dispersion there is in the price relatives \( p'_m/p_m^0 \), to the accuracy of a second order approximation, the more the direct Young index will exceed its counterpart that uses month t as the initial base period rather than month 0.
Annex 6: A note concerning the Taylor expansion/approximation

The results from the above sections and annexes show the differences between formulae in terms of variances, usually arising from a Taylor expansion around zero. It is an approximation and Annex 6 considers the expansion in more detail. The variances in equation (10) contain squares and cross-products and these cross-products are over time across products as well as for a given time period across items.

Say there are only \( m = 4 \) items over two periods, 0 and \( t \), as depicted in Figure A6.1 below. The \( e^t_j \) and \( e^0_j \) are the normalized errors so that in any period -0.2 for example, is 20% below the unitary mean and +0.2 is 20% above it.

\[
(A6.1) \quad \prod_{m=1}^{M} \left[ (1 + e^t_m)/ (1 + e^0_m) \right]^{1/4} \approx 1 + \frac{1}{4} \left[ \sum_{m=1}^{4} e^t_m - \sum_{m=1}^{4} e^0_m \right] - \frac{1}{4} \left[ 6 \sum_{m=1}^{4} (e^t_m)^2 - 10 \sum_{m=1}^{4} (e^0_m)^2 \right] \\
- \frac{1}{4^2} \left[ \sum_{m=n} \sum_{n} e^0_m e^t_n + \sum_{m=n} \sum_{n} e^0_m e^t_n - \sum_{m=n} \sum_{n} e^0_m e^0_n - \sum_{m=n} \sum_{n} e^t_m e^t_n \right] + \frac{1}{4^4} \left[ 14 \sum_{m=1}^{4} (e^t_m)^4 - 30 \sum_{m=1}^{4} (e^0_m)^4 \right] + \ldots \]

Note that the terms in the first square brackets sum to zero by definition in (9). The next two terms denote the difference between the normalized variances. Consider the cross-product term \( e^0_j e^t_k \) for \( m \neq n \) and Figure A6.1 below. There should be \( 12 \) cross-products of which 12 are for \( m \neq n \) and \( t \neq 0 \) – these are comparisons over time between different items. Say prices fluctuate around a normalized mean, so that in period 0, items 1 and 3 are above average (+ve) and items 2 and 4 below average (-ve), and the positions are in reversed period 1. For large \( M \) these cross-products will cancel. However, the next term are the 4 changes over time, \( e^0_1 e^t_1, e^0_2 e^t_2, e^0_3 e^t_3 \) and \( e^0_4 e^t_4 \) for which will all be negative, followed by the 6 cross-products across items in each of period 0 and 1 respectively \( m \neq n \cap t \neq 0 \), which should cancel to zero from Figure A6.1 for large \( m \). Finally there are 8 cubic terms which will naturally be followed by their cross-products.

The difference between the formulae also depends upon changes in the skewness of the price deviations.

**Figure A6.1**

<table>
<thead>
<tr>
<th>m</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+ve)</td>
<td>( e^0_1 )</td>
<td>(-ve)</td>
</tr>
<tr>
<td>(-ve)</td>
<td>( e^0_2 )</td>
<td>(+ve)</td>
</tr>
<tr>
<td>(+ve)</td>
<td>( e^0_3 )</td>
<td>(-ve)</td>
</tr>
<tr>
<td>(-ve)</td>
<td>( e^0_4 )</td>
<td>(+ve)</td>
</tr>
</tbody>
</table>

\(^{31}\) Balk (2002) shows an alternative derivation where the difference is identified as being dependent on the covariance between (the log mean of each period's) relative prices and price relatives. Such a covariance can be decomposed into the variances of the items on the Taylor expansion below.
The focus remains on changes in the variance to explain differences in the formulae, though fluctuations in the levels for individual items in the manner depicted in Figure 1 will also have an effect, as may changes in the cubic (skewness) terms.

It is not immediately obvious as to how to reconcile (A6.1) with (10). However, if a common denominator of $4^2=64$ is used the variances have weights of $-6/64$ and $+10/64$ in periods $t$ and 0 respectively and the cross products of $(-12/16, -4/16, +6/16$ and $+6/16)$ respectively which sums to $-16/64$. All items at the second order thus have weights summing to $-12/64=1/4$, i.e., the $1/M$ in (10).

Note the asymmetry in the weights for the variances in (A6.1). If the variances were the same the higher positive weight given to period 0’s variance would lead to $P_D(p_0, p_1) > P_D(p_1, p_0)$, though the cross-products might ameliorate the situation, especially the negative influence of price cumulated price changes under the scenario in Figure 1 outlined above. Any increase in the variances over time might tip the expression for the difference between the variances in (A6.1) to be negative, and the expansion to be less than unity so that $P_D(p_0, p_1) < P_D(p_1, p_0)$ as is apparent from (10).


Boskin MS (Chair) Advisory Commission to Study the Consumer Price Index (1996), Towards a More Accurate Measure of the Cost of Living, Interim Report to the Senate Finance Committee, Washington DC.


Table 1: Descriptive statistics on average monthly price dispersion*

<table>
<thead>
<tr>
<th>Measure of dispersion</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD unweighted</td>
<td>364.05</td>
<td>389.40</td>
<td>419.10</td>
<td>436.58</td>
<td>443.14</td>
</tr>
<tr>
<td>SDw weighted</td>
<td>269.92</td>
<td>316.72</td>
<td>343.79</td>
<td>375.56</td>
<td>381.84</td>
</tr>
<tr>
<td>Relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CV unweighted</td>
<td>0.83</td>
<td>0.86</td>
<td>0.85</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>CVw weighted</td>
<td>0.78</td>
<td>0.86</td>
<td>0.90</td>
<td>0.86</td>
<td>0.86</td>
</tr>
</tbody>
</table>

*The figures for the standard deviations (SD) are calculated for each month, the annual figures being simple averages of the 12 monthly SD measures. The CV annual averages follow accordingly. In 2002 there were only 3 months data the averages being for the months to March 2002.

Table 2: Decomposition of price variation

|                     |  $\bar{R}^2$ | (p-value) | Absolute value of residuals $|\hat{e}_m|$ | mean | standard deviation |
|---------------------|--------------|-----------|--------------------------|------|-------------------|
| **Ordinary least squares** |              |           |                          |      |                   |
| constant             | 2.11         | 1.00      |                          |      |                   |
| Month and Trend      | 0.004        | 0.0000    | 2.11                     | 1.00 |                   |
| Month, Trend and Charac | 0.88      | 0.0000    | 1.24                     | 0.36 |                   |
| Month, Trend, Charac and Brands | 0.91 | 0.0000 | 1.19                     | 0.278|                   |
| Month, Trend, Charac, Brands and Outlet-type | 0.92 | 0.0000 | 1.18 | 0.279 |
| **Weighted least squares** |              |           |                          |      |                   |
| constant             | 2.35         | 1.28      |                          |      |                   |
| Month and Trend      | 0.004        | 0.0000    | 2.34                     | 1.33 |                   |
| Month, Trend and Charac | 0.87       | 0.0000    | 1.25                     | 0.37 |                   |
| Month, Trend, Charac and Brands | 0.91 | 0.0000 | 1.20 | 0.29 |
| Month, Trend, Charac, Brand and Outlet-type | 0.92 | 0.0000 | 1.19 | 0.28 |

* $\bar{R}^2$ are based on untransformed variables.
Table 3: Results for regression of H-C prices on DIST

<table>
<thead>
<tr>
<th></th>
<th>OLS estimator</th>
<th></th>
<th>WLS estimator (values)</th>
<th></th>
<th>OLS estimator</th>
<th></th>
<th>WLS estimator (values)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>t-statistic</td>
<td>Coefficient</td>
<td>t-statistic</td>
<td>Coefficient</td>
<td>t-statistic</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Jan-98</td>
<td>0.0026</td>
<td>5.37</td>
<td>-0.000393</td>
<td>-0.98</td>
<td>Jan-99</td>
<td>0.0017</td>
<td>3.10</td>
</tr>
<tr>
<td>Feb-98</td>
<td>0.0022</td>
<td>3.99</td>
<td>-0.000334</td>
<td>-0.86</td>
<td>Feb-99</td>
<td>0.0018</td>
<td>3.06</td>
</tr>
<tr>
<td>Mar-98</td>
<td>0.0034</td>
<td>6.95</td>
<td>-0.000303</td>
<td>-0.84</td>
<td>Mar-99</td>
<td>0.0012</td>
<td>2.13</td>
</tr>
<tr>
<td>Apr-98</td>
<td>0.0030</td>
<td>6.25</td>
<td>-0.000359</td>
<td>-0.86</td>
<td>Apr-99</td>
<td>0.0023</td>
<td>3.81</td>
</tr>
<tr>
<td>May-98</td>
<td>0.0032</td>
<td>6.13</td>
<td>-0.000405</td>
<td>-0.85</td>
<td>May-99</td>
<td>0.0032</td>
<td>5.17</td>
</tr>
<tr>
<td>Jun-98</td>
<td>0.0037</td>
<td>6.72</td>
<td>-0.00021</td>
<td>-0.51</td>
<td>Jun-99</td>
<td>-0.1211</td>
<td>-1.25</td>
</tr>
<tr>
<td>Jul-98</td>
<td>0.0029</td>
<td>4.89</td>
<td>-0.000338</td>
<td>-0.69</td>
<td>Jul-99</td>
<td>0.0020</td>
<td>3.62</td>
</tr>
<tr>
<td>Aug-98</td>
<td>0.0027</td>
<td>4.86</td>
<td>-0.00053</td>
<td>-1.05</td>
<td>Aug-99</td>
<td>0.0023</td>
<td>3.66</td>
</tr>
<tr>
<td>Sep-98</td>
<td>0.0025</td>
<td>4.52</td>
<td>-0.00055</td>
<td>-1.22</td>
<td>Sep-99</td>
<td>0.0030</td>
<td>5.16</td>
</tr>
<tr>
<td>Oct-98</td>
<td>0.0027</td>
<td>4.62</td>
<td>-0.000664</td>
<td>-1.51</td>
<td>Oct-99</td>
<td>0.0022</td>
<td>4.14</td>
</tr>
<tr>
<td>Nov-98</td>
<td>0.0022</td>
<td>4.08</td>
<td>-0.000437</td>
<td>-0.73</td>
<td>Nov-99</td>
<td>0.0016</td>
<td>3.07</td>
</tr>
<tr>
<td>Dec-98</td>
<td>0.0022</td>
<td>3.98</td>
<td>-0.000899</td>
<td>-1.54</td>
<td>Dec-99</td>
<td>0.0021</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Table 4: Heterogeneity-controlled dispersion

<table>
<thead>
<tr>
<th>Period</th>
<th>Unweighted</th>
<th>Weighted</th>
<th>Unweighted</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-98</td>
<td>0.141</td>
<td>0.097</td>
<td>Jan-00</td>
<td>0.170</td>
</tr>
<tr>
<td>Feb-98</td>
<td>0.137</td>
<td>0.096</td>
<td>Feb-00</td>
<td>0.169</td>
</tr>
<tr>
<td>Mar-98</td>
<td>0.154</td>
<td>0.104</td>
<td>Mar-00</td>
<td>0.193</td>
</tr>
<tr>
<td>Apr-98</td>
<td>0.163</td>
<td>0.108</td>
<td>Apr-00</td>
<td>0.211</td>
</tr>
<tr>
<td>May-98</td>
<td>0.181</td>
<td>0.116</td>
<td>May-00</td>
<td>0.169</td>
</tr>
<tr>
<td>Jun-98</td>
<td>0.163</td>
<td>0.110</td>
<td>Jun-00</td>
<td>0.223</td>
</tr>
<tr>
<td>Jul-98</td>
<td>0.152</td>
<td>0.114</td>
<td>Jul-00</td>
<td>0.209</td>
</tr>
<tr>
<td>Aug-98</td>
<td>0.166</td>
<td>0.134</td>
<td>Aug-00</td>
<td>0.219</td>
</tr>
<tr>
<td>Sep-98</td>
<td>0.166</td>
<td>0.123</td>
<td>Sep-00</td>
<td>0.233</td>
</tr>
<tr>
<td>Oct-98</td>
<td>0.149</td>
<td>0.112</td>
<td>Oct-00</td>
<td>0.228</td>
</tr>
<tr>
<td>Nov-98</td>
<td>0.160</td>
<td>0.116</td>
<td>Nov-00</td>
<td>0.248</td>
</tr>
<tr>
<td>Dec-98</td>
<td>0.178</td>
<td>0.116</td>
<td>Dec-00</td>
<td>0.250</td>
</tr>
<tr>
<td>Jan-99</td>
<td>0.186</td>
<td>0.128</td>
<td>Jan-01</td>
<td>0.212</td>
</tr>
<tr>
<td>Feb-99</td>
<td>0.179</td>
<td>0.124</td>
<td>Feb-01</td>
<td>0.209</td>
</tr>
<tr>
<td>Mar-99</td>
<td>0.167</td>
<td>0.123</td>
<td>Mar-01</td>
<td>0.212</td>
</tr>
<tr>
<td>Apr-99</td>
<td>0.174</td>
<td>0.123</td>
<td>Apr-01</td>
<td>0.215</td>
</tr>
<tr>
<td>May-99</td>
<td>0.167</td>
<td>0.123</td>
<td>May-01</td>
<td>0.201</td>
</tr>
<tr>
<td>Jun-99</td>
<td>0.180</td>
<td>0.123</td>
<td>Jun-01</td>
<td>0.274</td>
</tr>
<tr>
<td>Jul-99</td>
<td>0.186</td>
<td>0.136</td>
<td>Jul-01</td>
<td>0.246</td>
</tr>
<tr>
<td>Aug-99</td>
<td>0.239</td>
<td>0.162</td>
<td>Aug-01</td>
<td>0.262</td>
</tr>
<tr>
<td>Sep-99</td>
<td>0.250</td>
<td>0.162</td>
<td>Sep-01</td>
<td>0.289</td>
</tr>
<tr>
<td>Oct-99</td>
<td>0.227</td>
<td>0.161</td>
<td>Oct-01</td>
<td>0.357</td>
</tr>
<tr>
<td>Nov-99</td>
<td>0.242</td>
<td>0.171</td>
<td>Nov-01</td>
<td>0.284</td>
</tr>
<tr>
<td>Dec-99</td>
<td>0.334</td>
<td>0.169</td>
<td>Dec-01</td>
<td>0.359</td>
</tr>
</tbody>
</table>

210 International Working Group on Price Indices - Seventh Meeting
### Table 5: Unit root tests

<table>
<thead>
<tr>
<th></th>
<th>Augmented weighted</th>
<th></th>
<th>Augmented Dickey-Fuller</th>
<th></th>
<th>Phillips-Perron</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symmetric tau</td>
<td></td>
<td>Dickey-Fuller</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>p-value</td>
<td>Coefficient</td>
<td>p-value</td>
<td>Coefficient</td>
<td>p-value</td>
</tr>
<tr>
<td>H-C Normalised SD&lt;sub&gt;weighted&lt;/sub&gt;</td>
<td>-3.96</td>
<td>0.01</td>
<td>-2.79</td>
<td>0.20</td>
<td>-7.64</td>
<td>0.61</td>
</tr>
<tr>
<td>H-C Normalised SD&lt;sub&gt;unweighted&lt;/sub&gt;</td>
<td>-2.61</td>
<td>0.23</td>
<td>-2.54</td>
<td>0.31</td>
<td>-26.54</td>
<td>0.02</td>
</tr>
<tr>
<td>H-C UNANTICIP&lt;sub&gt;weighted&lt;/sub&gt;</td>
<td>-3.60</td>
<td>0.02</td>
<td>-3.53</td>
<td>0.04</td>
<td>-53.30</td>
<td>0.00</td>
</tr>
<tr>
<td>H-C UNANTICIP&lt;sub&gt;unweighted&lt;/sub&gt;</td>
<td>-4.30</td>
<td>0.00</td>
<td>-4.15</td>
<td>0.01</td>
<td>-55.59</td>
<td>0.00</td>
</tr>
<tr>
<td>Number of models</td>
<td>-2.75</td>
<td>0.17</td>
<td>-2.47</td>
<td>0.34</td>
<td>-23.72</td>
<td>0.03</td>
</tr>
<tr>
<td>CATALOG&lt;sub&gt;weighted&lt;/sub&gt;</td>
<td>-2.03</td>
<td>0.63</td>
<td>-1.83</td>
<td>0.69</td>
<td>-8.57</td>
<td>0.54</td>
</tr>
<tr>
<td>CATALOG&lt;sub&gt;unweighted&lt;/sub&gt;</td>
<td>-1.66</td>
<td>0.84</td>
<td>-1.35</td>
<td>0.88</td>
<td>-6.20</td>
<td>0.73</td>
</tr>
</tbody>
</table>

### Table 6: Regression results

<table>
<thead>
<tr>
<th>Regression of:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef</td>
<td>t-statistic</td>
<td>Coef</td>
<td>t-statistic</td>
<td>Coef</td>
</tr>
<tr>
<td>ΔH-C normalised SD&lt;sub&gt;unweighted&lt;/sub&gt; on:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>0.003</td>
<td>0.54</td>
<td>0.001</td>
<td>0.03</td>
<td>0.004</td>
</tr>
<tr>
<td>ΔNumber of models</td>
<td>0.0001</td>
<td>2.06**</td>
<td>0.001*</td>
<td>1.83</td>
<td>0.000**</td>
</tr>
<tr>
<td>H-C UNANTICIP&lt;sub&gt;unweighted&lt;/sub&gt;</td>
<td>-0.200</td>
<td>0.56</td>
<td>-0.510</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td>ΔCATALOG&lt;sub&gt;unweighted&lt;/sub&gt;</td>
<td>-0.235</td>
<td>1.59</td>
<td></td>
<td></td>
<td>-0.279**</td>
</tr>
<tr>
<td>R²</td>
<td>0.14</td>
<td>0.11</td>
<td>0.15</td>
<td></td>
<td>0.06</td>
</tr>
</tbody>
</table>

| ΔH-C normalised SD<sub>weighted</sub> on: |          |           |          |           |          |           |          |           |          |           |
| constant                            | 0.002    | 1.36      | 0.0014   | 0.85      | 0.002    | 1.39       | -0.0013  | 0.83       |          |           |
| ΔNumber of models                   | -0.0000  | 0.09      | 0.137    | 0.65      | -0.0000  | 1.01       |          |           |          |           |
| H-C UNANTICIP<sub>unweighted</sub>  | -0.012   | 0.08      | -0.290   | 2.33**    | -        |            | 0.295**  | 2.40       |          |           |
| ΔCATALOG<sub>weighted</sub>         | -0.110   | 2.82***   | -0.112*  | 3.81      |          |            |          |           |          |           |
| R²                                  | 0.20     | 0.08      | 0.21     | 0.21      |          |            |          |           |          |           |

Tests are two-tailed and *, **, *** denotes statistically significant at a 5%, 1% and 0.1% level respectively.
Table 7: Index number formulae results

**Heterogeneity-controlled**

<table>
<thead>
<tr>
<th>Period</th>
<th>Dutot</th>
<th>Dutot Carli</th>
<th>Jevons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-98</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Feb-98</td>
<td>99.33</td>
<td>100.15</td>
<td>99.40</td>
</tr>
<tr>
<td>Mar-98</td>
<td>97.90</td>
<td>101.72</td>
<td>97.86</td>
</tr>
<tr>
<td>Apr-98</td>
<td>96.86</td>
<td>100.39</td>
<td>96.92</td>
</tr>
<tr>
<td>May-98</td>
<td>95.96</td>
<td>96.52</td>
<td>95.38</td>
</tr>
<tr>
<td>Jun-98</td>
<td>94.01</td>
<td>95.01</td>
<td>92.91</td>
</tr>
<tr>
<td>Jul-98</td>
<td>93.23</td>
<td>93.99</td>
<td>92.09</td>
</tr>
<tr>
<td>Aug-98</td>
<td>91.97</td>
<td>96.74</td>
<td>91.33</td>
</tr>
<tr>
<td>Sep-98</td>
<td>89.28</td>
<td>90.55</td>
<td>88.51</td>
</tr>
<tr>
<td>Oct-98</td>
<td>88.50</td>
<td>90.03</td>
<td>87.49</td>
</tr>
<tr>
<td>Nov-98</td>
<td>86.59</td>
<td>87.91</td>
<td>85.36</td>
</tr>
<tr>
<td>Dec-98</td>
<td>85.16</td>
<td>89.52</td>
<td>84.30</td>
</tr>
<tr>
<td>Jan-99</td>
<td>84.67</td>
<td>86.30</td>
<td>83.44</td>
</tr>
<tr>
<td>Feb-99</td>
<td>98.94</td>
<td>99.77</td>
<td>98.73</td>
</tr>
<tr>
<td>Mar-99</td>
<td>96.73</td>
<td>97.39</td>
<td>96.10</td>
</tr>
<tr>
<td>Apr-99</td>
<td>95.86</td>
<td>96.45</td>
<td>95.20</td>
</tr>
<tr>
<td>May-99</td>
<td>94.25</td>
<td>94.85</td>
<td>93.29</td>
</tr>
<tr>
<td>Jun-99</td>
<td>92.33</td>
<td>93.13</td>
<td>91.32</td>
</tr>
<tr>
<td>Jul-99</td>
<td>91.79</td>
<td>93.08</td>
<td>91.19</td>
</tr>
<tr>
<td>Aug-99</td>
<td>89.76</td>
<td>91.22</td>
<td>88.72</td>
</tr>
<tr>
<td>Sep-99</td>
<td>88.63</td>
<td>88.65</td>
<td>87.33</td>
</tr>
<tr>
<td>Oct-99</td>
<td>86.45</td>
<td>88.16</td>
<td>83.91</td>
</tr>
<tr>
<td>Nov-99</td>
<td>89.18</td>
<td>90.85</td>
<td>87.52</td>
</tr>
<tr>
<td>Dec-99</td>
<td>84.21</td>
<td>85.79</td>
<td>80.94</td>
</tr>
</tbody>
</table>

Figure 1: Heterogeneity-controlled dispersion
Introducing the Chained Consumer Price Index

Robert Cage, John Greenlees, and Patrick Jackman

U.S. Bureau of Labor Statistics

The authors would like to thank Ralph Bradley, David Scott Johnson, Joshua Klick, and Cassandra Wirth for helpful comments and suggestions on this paper. They also thank the organizers of and participants in seminars on the Chained CPI at the Federal Reserve Banks of New York, Philadelphia, Dallas, Atlanta, San Francisco, Chicago, St. Louis, and Richmond. Particular thanks are also due to Erwin Dievwert for his insights on superlative index design problems, without assigning him any responsibility for the choices ultimately made for the C-CPI-U. Finally, credit for many of the analyses reported in the paper goes to the Superlative Design Team of the Bureau of Labor Statistics: Ralph Bradley, Robert Cage, Alan Dorfman, Dennis Fixler, Richard Kerr, Janice Lent, Robert McClelland, Gary McMullin, and Janet Williams.

Abstract: In August 2002, the U.S. Bureau of Labor Statistics began publishing a consumer price index (CPI) called the Chained Consumer Price Index for All Urban Consumers. Designated the C-CPI-U, the index employs a superlative Tornqvist formula and utilizes expenditure data in adjacent time periods in order to reflect the effect of any substitution that consumers make across item categories in response to changes in relative prices. The new measure is designed to be a closer approximation to a "cost-of-living" index than the existing BLS measures.

Expenditure data required for the calculation of the C-CPI-U are available only with a time lag. Thus, monthly values of the C-CPI-U are issued first in preliminary form using the latest available expenditure data and are subject to two subsequent revisions. Accordingly, at the time of its introduction in August 2002, "Final" values of the C-CPI-U were issued for the 12 months of 2000, "Interim" values were issued for the 12 months of 2001, and "Initial" values were issued for January-July of 2002. In February 2003, with release of the January 2003 index, revised Interim indexes for the 12 months of 2002 were published, and the index values for 2001 were revised and became Final.

This paper details the calculation of the C-CPI-U and discusses the issues that were addressed in its design. The paper also describes the differences between the new index and the existing Laspeyres-formula CPI-U, and the February 2003 data revisions.

1. Introduction

In August 2002, the U.S. Bureau of Labor Statistics (BLS) began publishing a new index of consumer price change called the Chained Consumer Price Index for All Urban Consumers. Designated the C-CPI-U, the index supplements the existing indexes already produced by the BLS: the CPI for All Urban Consumers (CPI-U) and the CPI for Urban Wage Earners and Clerical Workers (CPI-W). The BLS is producing the C-CPI-U in order to address a
perceived weakness of the CPI-U: upper-level substitution bias. By utilizing a superlative price index aggregator across items, the C-CPI-U is designed to be a closer approximation to a "cost-of-living" index (COLI) than the CPI-U and the CPI-W.

This paper provides comprehensive information on concepts, definitions, statistical procedures, and estimation methods used by BLS to compile the C-CPI-U. Because the primary motivation behind publishing the C-CPI-U is to provide data users an index that more closely approximates a COLI, the paper begins with a brief summary of cost-of-living and consumer substitution theory. A cursory review of prior research on the substitution bias inherent in the CPI-U is also provided. Next, a detailed explanation of how BLS computes the C-CPI-U is outlined in a step-by-step fashion. Estimation methodologies for the C-CPI-U and CPI-U are then compared and contrasted, and differences between the two indexes are evaluated. Finally, because price indexes can be used for many purposes and indexes well-suited for one purpose may be ill-suited for another, the limitations of the C-CPI-U are identified so that users may evaluate the suitability of the index for their needs.

2. Conceptual Framework

The Consumer Price Index is a measure of the average price change of a fixed market basket of goods and services purchased by the average urban household in the United States. The market basket consists of a sample of items - food, clothing, shelter, fuels, and other goods and services—that consumers buy for day-to-day living. Price change is measured by repricing essentially the same market basket of goods and services at regular intervals and comparing aggregate costs with the costs of the same market basket in an arbitrarily selected base period.

BLS calculates a CPI for two population groups: (1) urban wage-earners and clerical workers and (2) all urban consumers. The Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W) is a continuation of the historical index introduced by BLS in the early 1900's for use in wage negotiations. The Consumer Price Index for All Urban Consumers (CPI-U) was introduced in 1978 as a broader and more representative index of the urban, non-institutional population of the United States. Because the same basic methodology is used for calculating both the CPI-W and CPI—U, the discussion in this paper focuses on the relationship between the CPI-U and the new superlative index.

Over the past 80 years, the methodology for producing both CPI’s has been refined by way of improvements in price data collection techniques, adjustments in estimation methods, and continuous surveys of consumer spending behavior. Comprehensive revisions to the CPI have occurred in 1940, 1953, 1964, 1978, 1987, and 1998. Other improvements have been made over the years that reflect not only BLS's own experience and research, but also the criticisms and recommendations of outsiders. The goal of each improvement has been to effectuate in the index a more accurate representation of contemporaneous buying habits and consumption costs. A unifying framework for dealing with practical questions that arise in the construction of the CPI, in assessment of the CPI’s quality, and in guiding improvements

---

3 For an early example, see, Report of The President's Committee on the Cost of Living (Washington, Office of Economic Stabilization, 1945).
made to the index, has been the economic approach to index numbers and, specifically, the concept of the cost-of-living index (COLI).

While the use of index numbers to measure price change dates back to 1707, the economic theory of index numbers is of much more recent vintage. The theory underlying the COLI was developed by A. A. Konus in 1924. Under the assumption of utility maximizing behavior, a COLI is defined as the ratio of the minimum expenditure required to attain a particular level of satisfaction in two price situations, a comparison period and a base period. The CPI-U and the CPI-W are modified Laspeyres indexes that hold the standard of living constant in the span between major revisions by keeping quantities fixed at the level consumed in the base period, but allowing prices to vary. The restriction imposed on these CPI's - holding the quantities fixed and not allowing substitution among goods in response to relative price change - results in a divergence between the CPI (or any other index with fixed quantity weights) and the COLI. In the case of a Laspeyres index, the effect is such that it is greater than or equal to the true cost of living. Indeed, it is well known that a Laspeyres index is an upper bound to the true COLI. In analyses of the CPI by outside reviewers, it often has been argued that the BLS should establish a cost-of-living index as the objective in measuring consumer prices. The BLS has long said that it operates within a COLI framework in producing the CPI. There are, of course, a number of differences between the CPI and a complete COLI other than the ability to reflect how consumers adjust their consumption to changes in relative prices. For example, the CPI ignores the fact that household preferences extend to choices between labor and leisure and among different types of leisure. It also ignores time by assuming that all consumption takes place in a single period. It generally ignores the impact of both the environment and government goods on household welfare. Reflection of substitution in response to changes in relative prices is, therefore, only part of what is required for a complete COLI. It is, however, that aspect of a COLI that has been the focus of the technical criticism of the CPI. The recent study by the Committee on National Statistics stated that: "Within the general conceptual framework of cost-of-living indexes, the appropriate theoretical concept for the CPI is a conditional cost-of-living index that is restricted to private goods and services and in which environmental background factors are held constant." 

The theory underlying the C-CPI-U was based largely on the work of W. Erwin Diewert, who demonstrated that a family of indexes, termed superlative, could be calculated that provided a

---

4 The first quantitative study of price levels was apparently made by Bishop Fleetwood in Chronicon-Preciosum in 1707. For a historical survey of price measurement, see W. Erwin Diewert, "The Early History of Price Index Research," W. Erwin Diewert and A.O. Nakamura (eds.), Essays in Index Number Theory, Volume I, 1993.


6 In general, a Laspeyres index is only an upper bound to a COLI when comparing the "comparison" period with the "base" period, not with an intermediate period.


9 Schultzze and Mackie, At What Price?, p. 73.
close approximation to a COLI using only the observable price and quantity data.\textsuperscript{10} That is, it would not be necessary to econometrically estimate the elasticities of substitution of all of the items with each other. The most widely known index number formulas that belong to the superlative class identified by Diewert are the Fisher Ideal index and the Tornqvist index. The Fisher Ideal Index is a geometric average of a base-period-weighted Laspeyres index and a current-period-weighted Paasche index. The Tornqvist index utilizes expenditure data in both the current and base time periods in order to reflect the effect of any substitution that consumers may make across item categories in response to changes in relative prices. Hence, both indexes arrive at an estimate of average price change over two time periods by utilizing in some fashion the expenditure experience in both periods as weights.

Although funds were initially appropriated in 1998, the development of the C-CPI-U has its roots back at least as far as 1961, with the recommendation from the Price Statistics Review Committee, known as the Stigler Committee, that a constant utility index is the appropriate index for the main purposes of the CPI.\textsuperscript{11} At the same time, the Stigler Committee recognized the need to fund research organizations within the price collection agencies to deal with price and index number issues outside of a production framework. Following the formation of BLS’s Division of Price and Index Number Research in 1966, research on developing a COLI was instituted. Much of the theoretical work on the COLI was done by Robert Pollak, first working at BLS on sabbatical from the University of Pennsylvania and later as a consultant.\textsuperscript{12} On an empirical side, several studies have examined the extent of the substitution effect between a Laspeyres measure and the COLI. On an aggregate level, these include studies by Steven D. Braithwait,\textsuperscript{13} Marilyn E. Manser and Richard J. McDonald,\textsuperscript{14} and Ana A. Aizcorbe and Patrick C. Jackman.\textsuperscript{15} Braithwait utilized annual price and quantity data on Personal Consumption Expenditures from the National Income and Product Accounts for 53 commodities and found that for the fifteen year period 1958-73, the Laspeyres index overstated the COLI by 1.5 percent, or about 0.1 percent a year. The work by Manser and McDonald covered the period from 1959-85 and utilized data on Personal Consumption Expenditures for 101 commodities. On an annual basis this resulted in an estimate of the substitution effect of .19 percent per year. In addition to the numerical estimate of substitution effect, the Manser-McDonald study established the upper and lower bounds for the amount of substitution effect in the Laspeyres index - between .14 and .22 percent per year in the 1959-1985 period. The Aizcorbe-Jackman study examined the substitution effect issue using detailed expenditure data from the Consumer Expenditure (CE) Surveys, the official source used for the CPI. Consumer Expenditure data include a finer level of disaggregation in the commodity classes as well as geographic detail than do data from the Bureau of Economic Analysis. At that time, there were 207 item categories and 44 areas, which allowed the different formulas to be applied at a lower level of aggregation than the earlier studies. Aizcorbe-Jackman’s estimate of the substitution effect for the period from

\textsuperscript{10} See, W. Erwin Diewert, “Exact and Superlative Index Numbers,” Journal of Econometrics 4, pp. 114-45, 1976. By a close approximation, Diewert states that the functional form for the price index is “superlative” if it is exact for a function that can provide a second order approximation to an arbitrary twice differentiable linearly homogenous function.

\textsuperscript{11} The Price Statistics of the Federal Government, p. 52.


1982 through 1991 was .20 or .27 percent per annum, depending on whether a fixed base or chained index was used.

3. Estimation Methodology

Specification and development of an official superlative index presented many challenges for the BLS. Foremost among these were the issues surrounding the sampling errors in CPI price and expenditure data. In most theoretical discussions of superlative indexes, price indexes and expenditure shares are taken as known with certainty, as if the data reflected a single representative consumer who allocates his or her spending in response to the observed price indexes. The CPI pricing surveys and the Consumer Expenditure Survey, however, are based on finite and distinct samples of outlets and consumers, respectively. Sample sizes are especially small at the level of basic item-area indexes and weights, such as were used in the Aizcorbe-Jackman work discussed above. The existence of sampling variation in the underlying CPI data, and the independence of the sampling errors in price indexes and expenditure shares, can affect the expectation of a superlative index computed from those data. 16

The impact of sampling variance could have been mitigated by constructing the C-CPI-U using only national-level data, but this would have tended to underestimate the true substitution effect. 17 Alternatively, the BLS could have chosen to produce only an annual C-CPI-U, using price and expenditure data averaged over the year. 18 This option was also rejected, on the basis that the usefulness of the index would have been sharply reduced if it were not made available on a monthly basis.

It was also determined that, rather than publishing the C-CPI-U only with a lag, the BLS would publish the index in preliminary form and revise it when more timely expenditure data were received and processed. As with the decision to produce a monthly C-CPI-U, this decision was motivated in part by an expectation that some users would desire a monthly, timely superlative index and would estimate such indexes themselves if the BLS did not publish the C-CPI-U in that form.

The BLS was thus faced with the problem of choosing how to employ volatile price and expenditure data from small samples to compute a monthly superlative series. After much discussion, it was further decided to chain the monthly index values, despite the volatile and potentially seasonal nature of the underlying data, rather than attempting to impose chaining on an annual or other frequency. Many of the specific design features described below reflect difficult choices made among several imperfect alternatives, without the benefit of guidance from index number theory. Fortunately, analyses of historical data indicated that simulated C-CPI-U series were surprisingly robust to their specific design features, particularly at the aggregate level.

17 See Greenlees, op. cit.
18 Erwin Diewert, "Notes on Producing an Annual Superlative Index Using Monthly Price Data," University of British Columbia Economics Department Discussion Paper No. 00-08 (July 2000) discusses issues in constructing an annual superlative index.
A second significant issue was determining how to compute the preliminary C-CPI-U indexes, which could not use a true superlative formula. The objective was to develop a specification that would yield accurate forecasts of the Final C-CPI-U. Again, a variety of alternative formulas were evaluated, but simulation studies showed that the differences in the resulting preliminary index values were relatively small.

3.1 Construction of the Consumer Price Index

The CPI is built in two stages. In the first stage, price changes for roughly 80,000 specific items per month are averaged to yield 8,018 estimates of aggregate price change. This stage is often referred to as “lower-level aggregation” or “elementary-level aggregation” as it involves averaging the most fundamental component of the index - observed price change for specifically defined consumer goods, services, and products. For example, the prices of approximately ten different brands and styles of watches at various locations in Chicago are observed each month, compared to the prices observed in the previous month, and averaged together to produce an index of price change for watches in Chicago. Watches (ITEM=AG01) is one of 211 elementary items, and Chicago (AREA=A207) is one of 38 elementary areas in the current CPI market basket structure. The Chicago-watch index is one of the 8,018 (211 items x 38 areas) elementary indexes produced in the first stage of CPI construction.

In the second stage, the elementary indexes are averaged together to yield various aggregate indexes and ultimately the All-Items, U.S. City Average index of price change. See Figure 3.1.

Within the two-tiered scheme of calculating the CPI, consumer substitution can and does occur at both levels. Ideally, a superlative formula would be employed at both stages of CPI index construction to account for consumer substitution that might occur intra-item, that is among specific products within an elementary item (e.g., a leather band watch versus a stainless steel band watch or whole wheat bread versus white bread) and inter-item, that is across elementary items (e.g., theater admission versus video rental or beer versus wine). However, the BLS is currently precluded from using a superlative formula at the first stage because reliable monthly expenditure data for each of the 80,000 lower-level quotes are not readily ascertainable. As an alternative to a superlative index as a means of addressing intra-item substitution, the BLS began using a hybrid combination of Geometric Mean indexes and Laspeyres indexes for lower-level aggregation in 1999. Zero elasticity of substitution within item categories is assumed for the small number of item categories in which the Laspeyres formula is used. Unitary elasticity of substitution intra-item is assumed for items using the Geometric Mean formula.

The use of a superlative formula for upper-level aggregation is designed to address inter-item substitution. In order to use a superlative index formula at the upper-level, monthly expenditure estimates for each of the 8,018 elementary item-area combinations are required.

19 In BLS terminology, the specific goods and services at the lower-level are called price quotes.

20 Collecting monthly information from CPI outlets on the sales of individual items would be extremely costly and would impose an unacceptable respondent burden.


22 Laspeyres items are Local telephone service, Rent, Housing at school, Owners’ equivalent rent, Electricity, Natural gas, Water and sewerage service, Physicians’ services, Dental services, Eyeglasses and eye care, Other medical professional service, Hospital services, Nursing homes and adult daycare, Cable television, and State and local vehicle registration, license, and motor vehicle property tax.
Expenditure weights for CPI upper-level aggregation are derived from the Consumer Expenditure Surveys.\textsuperscript{23} Effective with data collected in 1999, the CE sample size was increased by 50 percent, in part to accommodate production of the C-CPI-U.

\section*{3.2 Availability of Requisite Expenditure Data}

Consumer Expenditure Survey data are processed annually and made available for CPI use with a substantial lag. For example, data for calendar year 2001 were not available until the fourth quarter of 2002. Data for calendar year 2002 will not be available until the fourth quarter of 2003. This lag in data availability prevents BLS from calculating and publishing a superlative index in real time, i.e., on the same contemporaneous publication schedule as the CPI-U. BLS could have opted to simply calculate and publish the C-CPI-U on a two-year lag schedule. For example, superlative indexes for all months of calendar year 2002 could have been released in early 2004. However, as noted above, demand for the superlative indexes concomitant with the release of the CPI-U was anticipated to be high. Therefore, BLS developed the following general estimation and publication schedule for the C-CPI-U index for a given month (t) in year (y):

- **Following Month (t+1):** Calculate and publish an \textit{Initial} estimate of the month (t) superlative index using lagged expenditure data
- **February of Following Year (y+1):** Revise and publish each Initial monthly superlative index from year (y) with an \textit{Interim} estimate based upon updated, but still lagged expenditure data
- **February of Following Year (y+2):** Calculate and publish a \textit{Final} superlative index for each month of year (y) using the now-available monthly expenditure data from year (y)

Hence, the C-CPI-U index for any given month will be published simultaneously with the CPI-U as an initial estimate. This initial estimate will be revised and republished the following February, in the news release containing CPI data for January. Then, a final estimate will be published in the February of the following year. Accordingly, there will be three versions of each monthly C-CPI-U index: an Initial, Interim, and Final version. At any point in time, the historical C-CPI-U index series will be comprised of (a) Initial values for all months in the current year, (b) Interim values for all months in the previous year, and (c) Final values for all months prior to the previous year. See Figure 3.2 for an illustration of the publication and revision schedule.

\section*{3.3 Estimation of Upper-Level Price Change}

Aggregation of elementary CPI data into published indexes requires three ingredients: input elementary indexes, input elementary expenditures to use as aggregation weights, and a price index number formula that employs the expenditures to aggregate the sample of elementary indexes into a published index.

**Input Elementary Price Indexes.** All three versions of the C-CPI-U index utilize the exact same hybrid Laspeyres and Geometric Mean lower-level indexes as are currently used in the

---

\textsuperscript{23} The Consumer Expenditure Surveys have been conducted continuously since 1980 by the U.S. Bureau of the Census under contract with the BLS.
Equation 3.1.a. Laspeyres items:

\[ L_{i,a}^{L} = \sum_{k \in a} k s_{0} \left( \frac{k P_{t}}{k P_{0}} \right) \]

Equation 3.1.b. Geometric Mean items:

\[ G_{i,a}^{L} = \prod_{k \in a} \left( \frac{k P_{t}}{k P_{0}} \right)^{k s_{0}} \]

Estimation of Lower-Level Price Change in the CPI

Publishing the Final C-CPI-U with a substantial lag presents the opportunity to adopt an alternative approach for imputing off-cycle indexes and to eliminate any potential bias associated with using previous-month imputation. Specifically, off-cycle indexes used in

---

24 Equations 2.1.a and 2.1.b provide the general estimator formulas for most commodity and services items. Rent and owners' equivalent rent are estimated using a slightly different approach. See Frank Ptacek and Robert M. Baskin, “Revision of the CPI housing sample and estimators,” Monthly Labor Review, December 1996.

25 BLS field economists do not price the samples of quotes in all 8,018 elementary item-area cells on a monthly basis. Approximately 40 percent of the elementary item-area combinations are priced on a bi-monthly schedule, with half of these priced every even month and the other half priced every odd month. Months for which prices are not collected are called “off-cycle” months.

26 For example, the index for “Sports vehicles” (ITEM=RC01) in “Seattle” (AREA=A423) in December 2001 was 116.4, relative to June 1985. This elementary index is off-cycle in odd months. Hence, prices are not collected in January, and the January 2002 index was set equal to the December 2002 index value of 116.4. The February 2002 index is computed using prices observed in February. The index value was 116.8 relative to June 1985.
Final C-CPI-U aggregation are set equal to the geometric mean of the immediately previous and subsequent on-cycle month indexes. See Equation 3.2.  

\[
i_{a,i}I_X[t] = \left( i_{a,i}I_X[0,t-1] \times i_{a,i}I_X[0,t+1] \right)^{1/2}
\]

where,

- \( a \) = CPI elementary area
- \( i \) = CPI elementary item
- \( t \) = year and month
- \( 0 \) = base-period reference month

Imputing off-cycle elementary indexes with geometric averaging of bounding month indexes, rather than using previous-month index values, is not expected to have a dramatic impact on published Final C-CPI-U indexes. To measure the impact, Final C-CPI-U indexes were estimated using both methods of off-cycle bimonthly elementary index imputation for the 1987 to 2000 time period.  

Use of geometric averaging of off-cycle elementary indexes produced, on average, a 0.001 percent increase in the All-Items, U.S. City Average Final C-CPI-U index per annum over this time frame. See Table 3.1. The differences ranged from -0.033 percent in 1992 to 0.025 percent in 1996. The fact that geometric averaging yielded a larger index value for some years (in 1990, 1993-1999) and a smaller index value in other years (in 1988, 1989, 1991, 1992, and 2000) is evidence that previous-month imputation does not result in any systematic effect on inflation measurement.

**Input Elementary Expenditure Weights.** In order to aggregate elementary indexes into published indexes, an aggregation weight for each elementary item-area combination is required. The function of the aggregation weight is to assign each elementary index a relative importance or contribution in the resulting aggregate index. The aggregation weight corresponds to consumer tastes and preferences and resulting expenditure choices among the 211 items in the 38 areas comprising the CPI sample, for a specified time period, by the population the index is designed to represent. This section compares the estimation and use

---

27 For example, the January 2002 index for “Sports vehicles” in Seattle was set equal to 116.6, the geometric average of the December 2001 index (116.4) and the February 2002 index (116.8).

28 For simulation analysis, a laboratory of input elementary price indexes and elementary expenditures was created using official CPI lower-level indexes and Consumer Expenditure Survey data, from December 1986 to December 2000, in order to evaluate features of C-CPI-U index construction. The CPI item structure changed significantly in January 1998, with the number of item categories increasing from 207 to 211 and the number of areas decreasing from 46 to 38. See Walter Lane, “Changing the Item Structure of the Consumer Price Index,” Monthly Labor Review, vol. 119, no. 12, December 1996, pp. 18-25. In order to achieve a continuous time series of data from 1986 to 2000 for each of the 8,018 elementary item-area combinations currently in the CPI sample, item-area index and expenditure levels were roughly approximated for all months in the 1986 to 1997 time span, using available indexes and expenditures based on the 1987 CPI market basket structure. The monthly expenditure data were adjusted according to official methodologies adopted for Final C-CPI-U construction. However, the monthly expenditure data for years 1986 through 1999 in the simulation laboratory are based on underlying CE sample sizes significantly smaller than that achieved in 2000, when the CE sample was increased by 50 percent. This affects the variance of the simulated indexes discussed herein. The sampling error associated with monthly expenditure data from 2000 and beyond is expected to be lower than that observed in the experimental lab.
of aggregation weights in the Laspeyres CPI-U, preliminary C-CPI-U series, and the Final C-
CPI-U.

a. CPI-U. In the CPI-U, aggregation weights are defined as:

\[
AW_{i,a,p} = \frac{i,a,p \hat{P}_a \times i,a,p \hat{Q}_\beta}{100}
\]

where \(i,a,p \hat{P}_a\) is the estimated price of item \((i)\) faced by population \((p)\) in area \((a)\) in time period \((\alpha)\), and \(i,a,p \hat{Q}_\beta\) is the estimated quantity of item \((i)\) purchased by population \((p)\) in area \((a)\) in time period \((\beta)\). Time period \((\alpha)\) is the base period of the corresponding elementary item-area index—i.e., the period at which the index equals 100.\(^{29}\)

Time period \((\beta)\) corresponds to the reference period of the expenditures used to derive the implicit quantity weights needed for Laspeyres aggregation. Currently, the CPI-U has an expenditure reference period of \(\beta=1999-2000\). Historically, the CPI expenditure reference period has been updated approximately every ten years (see Table 3.2). In 1998, BLS announced that it would institute a biennial rotation schedule for updating the expenditure reference period. Effective with the January 2004 index, the expenditure reference period will change from \(\beta=1999-2000\) to \(\beta=2001-2002\); effective with the January 2006 index it will be updated again to 2003-2004; and so forth. Note that a change in the expenditure reference period results in a change in the implicit quantity \((Q)\) assigned to each elementary index, but not the implicit price component \((P)\) of the aggregation weight \((AW)\).

Aggregation weights for the CPI-U are derived from estimates of household expenditures collected in the Consumer Expenditure Survey data. Despite an increase in the CE sample size in 1999, expenditure estimates at the elementary item-area level would be unreliable due to sampling error without the use of statistical smoothing procedures. BLS uses two basic techniques to minimize the variance associated with each elementary item-area base-period expenditure estimate. First, data are pooled over an extended time period in order to build the expenditure estimates upon an adequate sample size. The current reference period \((\beta)\) uses 24 months of data. Second, elementary item-area expenditures are averaged, or composite-estimated, with item-regional expenditures.\(^{30}\) This has the effect of lowering the variance of each elementary item-area expenditure at the expense of biasing it toward the expenditure patterns observed in the larger geographical area.\(^{31}\)

The CPI-U aggregation weight for item \((i)\) in area \((a)\) in reference period \((\beta)\) is computed by first calculating an aggregate annual expenditure estimate, \(i,a (PQ)_\beta\), for each year \((\beta_n)\) in reference period \((\beta)\). This estimate is derived directly from Consumer Expenditure Survey

\(^{29}\) For example, the “Sports equipment” \((ITEM=RC02)\) in Seattle \((AREA=A423)\) index has a base period of \(\alpha=June 1985\). CPI elementary indexes have varying base periods that are not updated on a regular basis. Most published indexes have an index base period of \(\alpha=1982-1984\).

\(^{30}\) Elementary areas area grouped into region x city-size classifications for the purpose of composite-estimation. There are four regions and two city-size classifications for a total of eight region-city-size classifications.

\(^{31}\) Aggregation weights for the CPI-U and CPI-W are each derived separately according to the steps outlined in the text.
data.\textsuperscript{32} Next, the share of total area expenditure is computed for each item in each area, for each year. Similarly, the share of total expenditure in each major-area \((m)\) is computed for each item for each year. A composite-estimated share of total expenditures is computed for each item for each year by taking a weighted average of its area share and corresponding major-area share. The weight \((\delta)\) assigned to the major-area \((m)\) and the weight \((1-\delta)\) assigned to the elementary area \((a)\) is a function of the variance and covariance of each measure.\textsuperscript{33} The resulting average share \(\tilde{\beta}_i\) is then multiplied by the sum of all expenditures in the elementary area in the corresponding year, to obtain a composite-estimated item expenditure in year \((\beta_n)\). This estimate is in turn multiplied by a “raking factor” which is equivalent to the ratio of unadjusted expenditures \((i,a, (PQ)_{\beta})\) summed to the expenditure-class, major-area level, to the composite-estimated expenditures \((i,a, (\tilde{PQ})_{\beta})\) summed to the expenditure-class, major-area level. The raking factor is designed to limit the degree to which composite-estimation can change relative expenditures among item-area cells. Next, the composite-estimated-and-raked expenditures for each year \((\beta_n)\) are averaged to obtain the final estimate of annual aggregate expenditures in reference period \((\beta)\).

The CPI-U aggregation weight for each item-area combination is then derived from the composite-estimated-and-raked expenditure estimate by first multiplying it by the index of price change from reference period \((\beta)\) to pivot-month \((v)\).\textsuperscript{34} The resulting product is a cost weight: an estimate of item-area expenditures in pivot-month \((v)\), based upon quantities purchased in reference period \((\beta)\). Finally, the cost weight is divided by the corresponding pivot-month index to obtain the aggregation weight: an estimate of item-area expenditure based upon quantities purchased in reference period \((\beta)\) and prices of time-period \((\alpha)\).

### Estimation of Monthly Expenditures at the Elementary Level in the C-CPI-U

Equation 3.4. Estimated monthly expenditures

\[
_{i,a} (\tilde{PQ})_t = \sum_{(i,a)\in(i,A)} (PQ)_t \times \left( \frac{\sum_{i,a} (PQ)_t \times \sum_{i,a} (PQ)_{t}}{\sum_{i,a} (PQ)_t} \right)
\]

where,
- \(a\) = CPI elementary area
- \(i\) = CPI elementary item
- \(A\) = all CPI elementary areas; "U.S. City Average"
- \(P\) = price
- \(Q\) = quantity
- \(t\) = month
- \(T\) = time period covering month \((t)\) and 11 months prior to month \((t)\)

\textsuperscript{32} For a detailed explanation of how aggregate expenditure estimates are computed from CE data, see BLS Handbook of Methods, Bulletin 2490, 1997, Chapter 17.


\textsuperscript{34} The pivot-month is the first month in which expenditures from reference period \((\beta)\) are used in the CPI.
**b. Final C-CPI-U.** For the Final C-CPI-U, which uses the Tomqvist index for upper-level aggregation in a monthly-chained construct, monthly expenditure estimates for each elementary item-area combination are required as aggregation weights. Like the biennial data used for CPI-U aggregation, adequacy of the underlying CE sample size from which the expenditure weights are estimated is an issue for C-CPI-U aggregation. In order to minimize the variance of the elementary item-area-monthly expenditures, a ratio-allocation procedure is used to estimate each item-area-monthly expenditure from item-U.S.-monthly expenditures. See Equation 3.4.

Moreover, the CPI-U estimate is refined by averaging area data with major-area data, and the C-CPI-U estimate is refined by allocating U.S. expenditures to each area based upon previous-year expenditure patterns among the areas. Hence, both methods “borrow” information across time and geography in order to estimate item-area-reference period expenditures. Once summed, the estimated item-area-monthly expenditure data are equal to the composite-estimated-and-raked expenditure data at the item-U.S.-year level. See Figure 3.3 for an illustrative example of monthly weight estimation of an elementary item-area cell in the C-CPI-U.

c. Initial and Interim C-CPI-U. Lacking a satisfactory method to forecast the requisite monthly expenditure data, BLS opted to select an aggregation methodology for the Initial and Interim versions of the C-CPI-U that would best predict the Final C-CPI-U Tomqvist version – constrained by the use of the most contemporaneous expenditure data available at the time of index publication, i.e., expenditures from the CPI-U expenditure reference period (β). An adjusted Geometric Mean index formula was ultimately adopted. See the discussion of Aggregation Formula for Initial and Interim Indexes below.

Since the Initial version of the C-CPI-U is published simultaneously with the CPI-U, it uses expenditure data from the same expenditure reference period (β) as the CPI-U as aggregation weights. In contrast to the CPI-U, however, it is not necessary to adjust the expenditures forward to a December “pivot” month and rebase them such that the implicit price corresponds to the corresponding item-area index base period (α). Rather, the estimated expenditure weights with implicit prices of time period (β) and implicit quantities of time period (β) are used as aggregation weights. This is consistent with the underlying assumption behind a geometric mean price index aggregator: consumers respond to changing relative prices by holding their expenditure shares constant over time. Hence, it is implicitly assumed that each item-area expenditure share derived from reference period (β) will be equal to each monthly item-area expenditure share over the time period in which aggregation weights based on reference period (β) are used to construct the Initial and Interim C-CPI-U indexes. In other words, $i,aS_β = i,aS_{t1} = i,aS_{t2} = \ldots = i,aS_{t24}$, where $t1 \text{ to } t24$ are the 24 months for which aggregation weights derived from (β) are used to construct the Initial and Interim C-CPI-U indexes.

The Interim version of each monthly C-CPI-U index will be published in February of the ensuing year. If the ensuing year is a weight update year, then the Interim version of each monthly C-CPI-U will be based upon more contemporaneous expenditures than its Initial version. For example, 2002 Initial indexes produced in 2002 will use $\beta=1999,2000$. Interim indexes for 2002 will be produced in 2003 and will likewise use $\beta=1999,2000$. Initial indexes for 2003 will also use $\beta=1999,2000$. However, 2003 Interim indexes will be produced in 2004, a weight update year. Hence they will be constructed using $\beta=2001,2002$.
Aggregation Formula.

a. CPI-U. The Laspeyres price index is used to aggregate elementary indexes into published CPI-U indexes. The Laspeyres index uses quantities from the predetermined expenditure reference period \((\beta)\) in order to weight each elementary item-area index. These quantity weights remain fixed for a two-year period and are then replaced each January in each even year when the aggregation weights are updated. Zero elasticity of substitution within item categories is assumed. An aggregate index for any given month is computed as a quantity-weighted average of the current month index divided by the index value in the index base period. See Equation 3.6.a. Month-to-month price change is then calculated as a ratio of the long-term monthly indexes. See Equation 3.6.b.

### CPI-U Upper-level Aggregation Formula

**Equation 3.6.a. Long-term Price Change**

\[
I,A IX^L_{[z,\tau]} = I,A IX^L_{[z,\nu]} \times \frac{\sum_{i,a} AW\beta_i_A IX_{\text{LoG}}^{L} \left[\alpha,\tau\right]}{\sum_{i,a} AW\beta_i_A IX_{\text{LoG}}^{L} \left[\alpha,\nu\right]}
\]

**Equation 3.6.b. Month-to-Month Price Change**

\[
I,A IX^L_{\{t-1,\nu\}} = \frac{I,A IX^L_{\left[\nu,\tau\right]}}{I,A IX^L_{\left[z;\nu\right]}}
\]

where,

- \(a\) = CPI elementary area
- \(A\) = all elementary areas; "U.S. City Average"
- \(i\) = CPI elementary item
- \(I\) = all elementary items; "All-items"
- \(t\) = month
- \(z\) = base period of the aggregate index (NOTE: the All-Items, U.S. City Average CPI-U index has a base-period of \(z=1982-84\))
- \(\alpha\) = base period of the elementary index \((i)\) in area \((a)\)
- \(\nu\) = year and month, usually December, prior to the month expenditure weights from reference period \((\beta)\) are first used in the CPI
- \(i,a IX_{[a,\tau]}\) = lower-level index of price change from period \((\alpha)\) to month \((t)\) for item \((i)\) in area \((a)\)
- \(i,a IX_{[a,\nu]}\) = lower-level index of price change from period \((\alpha)\) to pivot-month \((\nu)\) for item \((i)\) in area \((a)\)
- \(i,a AW\beta\) = aggregation weight from reference period \((\beta)\) for item \((i)\) in area \((a)\)
- \(i,A IX_{\{a,\nu\}}\) = aggregate-level CPI-U index of price change from period \((z)\) to pivot-month \((\nu)\) for aggregate item \((I)\) in aggregate area \((A)\)

b. Final C-CPI-U. In contrast, the C-CPI-U is built by chaining together indexes of one-month price change. For the Final C-CPI-U index, each monthly index is computed using the Tornqvist formula and monthly weights from the current month and previous month. Consumer substitution behavior is not assumed by the Tornqvist formula, but rather implicitly accounted for by use of current and base-month expenditures. An index of one-month price change is calculated and then multiplied by the previous month index value to obtain the current month index value. See Equations 3.7.a and 3.7.b.
Final C-CPI-U Upper-level Aggregation Formula

Equation 3.7.a. Long-term Price Change

\[ I_A IX_{[z:t]}^T = I_A IX_{[z:t-1]}^T \times I_A IX_{[t-1:t]}^T \]

Equation 3.7.b. Month-to-Month Price Change

\[ I_A IX_{[t-1:t]}^T = \prod_{(i,a) \in (I,A)} \left( \frac{I_A IX_{[a:t]}^{Log}}{I_A IX_{[a:t-1]}^{Log}} \right)^{z_{a-1:t-1}} \]

where,

- \( a \) = CPI elementary area
- \( A \) = all elementary areas; "U.S. City Average"
- \( i \) = CPI elementary item
- \( I \) = all elementary items; "All-items"
- \( z \) = base period of the aggregate index (NOTE: the All-Items, U.S. City Average C-CPI-U index has a base-period of \( z = \) December 1999)
- \( a \) = base period of the elementary index (\( i \)) in area (\( a \))
- \( I_A IX_{[a:t]} \) = lower-level index of price change from period (\( a \)) to month (\( t \)) for item (\( i \)) in area (\( a \))
- \( I_A IX_{[a:t-1]} \) = lower-level index of price change from period (\( a \)) to month (\( t-1 \)) for item (\( i \)) in area (\( a \))
- \( I_A IX_{[a:t]} \) = expenditure in month (\( t \)) for item (\( i \)) in area (\( a \)) as percent of total expenditures in month (\( t \)) for aggregate item (\( I \)) in aggregate area (\( A \))
- \( I_A IX_{[a:t-1]} \) = expenditure in month (\( t-1 \)) for item (\( i \)) in area (\( a \)) as percent of total expenditures in month (\( t-1 \)) for aggregate item (\( I \)) in aggregate area (\( A \))
- \( I_A IX_{[z:t]} \) = aggregate-level C-CPI-U index of price change from period (\( z \)) to month (\( t \)) for aggregate item (\( I \)) in aggregate area (\( A \))

**c. Initial and Interim C-CPI-U.** Selecting an aggregation formula for the Initial and Interim versions of the C-CPI-U was more complex. As discussed above, construction of the Initial and Interim versions of the C-CPI-U is constrained by the use of aggregation weights from a lagged expenditure reference period (\( B \)). The Geometric Mean index and the Constant-Elasticity-of-Substitution (CES) index were identified as possible aggregation formula alternatives, since both of these indexes can be produced in real-time using expenditures from the lagged base-period (\( B \)).

The CES long-term index is defined by Equation 3.8. The CES function was developed by Arrow, Chenery, Minhas, and Solow in 1961.\(^{35}\) Use of the CES functional form as a price index number formula was derived independently by Lloyd (1975) and Moulton (1996).\(^{36}\) In 1996, Shapiro and Wilcox advocated the use of the CES as a feasible, “real-time” preliminary


index for the Tornqvist, arguing it allows for substitution effects while not requiring current expenditure data.37

**Constant Elasticity of Substitution Index**


\[
1_a \times I_{(a;1)}^C = 1_a \times I_{(a;1)}^C \times \left[ \sum_{i,a} \alpha_{(i,a)} \beta_{(i,a)} \left( \frac{I_{(i,a)}^C}{I_{(i,a)}^C} \right)^{1-\sigma} \right]^{1-\sigma}
\]

where,

- \(a\) = CPI elementary area
- \(A\) = all CPI areas (U.S. City Average)
- \(i\) = CPI elementary item
- \(I\) = all CPI items (All-items)
- \(z\) = base period of the aggregate index
- \(t\) = month
- \(\alpha\) = base period of the elementary index \((i)\) in area \((a)\)
- \(\beta\) = expenditure reference period
- \(\gamma_{(i,a)}\) = expenditure in period \((\beta)\) for item \((i)\) in area \((a)\) as percent of total expenditures in period \((\beta)\) for aggregate item \((I)\) in aggregate area \((A)\)
- \(\gamma_{(i,a)}\) = lower-level index of price change from index base-period \((\alpha)\) to period \((\beta)\) for item \((i)\) in area \((a)\); NOTE: because period \(\beta\) encompasses 24 months, this index is computed as an unweighted arithmetic average of the 24 monthly indexes contained in \((\beta)\)
- \(\gamma_{(i,a)}\) = lower-level index of price change from index base-period \((\alpha)\) to period \((t)\) for item \((i)\) in area \((a)\)
- \(\sigma\) = elasticity of substitution parameter
- \(\gamma_{(i,a)}\) = aggregate-level CES index of price change from period \((z)\) to month \((t)\) for aggregate item \((I)\) in aggregate area \((A)\)

The distinguishing feature of the CES is the elasticity of substitution parameter \((\sigma)\). Roughly speaking, substitution elasticity measures the proportionate change in the relative quantity demanded of a commodity over the proportionate change in its relative price. The CES is a pliable functional form, collapsing into various other price index formula in the limiting values of \((\sigma)\). When \(\sigma=1\) (unitary elasticity of substitution), Equation 3.8 reduces to the Geometric Mean price index (Cobb-Douglas preferences). When \(\sigma=0\) (zero elasticity of substitution), Equation 3.8 reduces to the Laspeyres price index (Leontief preferences).

The CES was analyzed in detail by BLS as a potential Initial and Interim C-CPI-U aggregator. For a variety of reasons, however, the CES was judged ill-suited as an approximation of the Final C-CPI-U. First, estimation of \((\sigma)\) is problematic from a theoretical perspective. In principle, to aggregate the 211 elementary CPI indexes into aggregate CPI indexes the elasticity of substitution among the 211 items must be estimated in each of the 38 elementary CPI areas. This requires estimation of a complete demand system, i.e., estimation of \(\sigma\) for all 22,155 possible pairs of elementary items, in each of the 38 areas (i.e., 841,890 possible combinations). A representative price index aggregator could have been selected from a class

of variable-elasticity-of-substitution functions to accomplish this task. However, a variable-elasticity aggregator suffers from several undesirable index qualities, most notably inconsistency in aggregation, and is infeasible to produce with the data currently available to BLS.

Alternatively, following Shapiro and Wilcox, the CES functional form with a single value of (σ) could be assumed to hold across all possible pairs of item strata. The CES is much more feasible to produce than a variable-elasticity aggregator. One difficulty with using the CES, however, resides in the selection of the optimal value of (σ). Using the laboratory of elementary price indexes and expenditure weights from December 1986 to December 2000, monthly CES indexes were simulated in order to find solutions to the fitting parameter, i.e., the value of (σ) yielding a CES index best-approximating a Tomqvist index. The simulations revealed that the optimal elasticity of substitution parameter was unstable across time (ranging from a low of σ=0.06 to a high of σ=2.78). See Figure 3.4. Accordingly, the assumption of a constant value of (σ) across month could result in specification error. This was judged a major weakness of the CES.

Moreover, the conceptual underpinnings of the CES was judged ill-suited for use in building a monthly-chained time series. A CES constructed with expenditure shares derived from the same base-period month as used in the Tomqvist, i.e., $s_{t-1}$, could be used to approximate the Final C-CPI-U month-to-month index for month (t). These CES month-to-month indexes could then be chained together to produce the long-term Initial and Interim C-CPI-U index series. There are two problems with this approach: lags in data availability preclude producing a (t-1) weighted CES in real-time, and a monthly-chained CES could be impacted by chain drift. Instead, expenditure shares derived from available data, i.e., $i_{t} s_{t}$, must be used in a biennially-chained construct.

The CES functional form further requires the expenditure shares (s) to be measured over the same time period as the denominator of the price relative. If the only available expenditure shares are those derived from expenditure reference period (β), which encompasses 24 months, it follows that the denominator price index in the CES relative must be some average index that is representative of the 24 discrete indexes available in time (β). The choices (weighted versus unweighted average, arithmetic versus geometric average, etc.) introduce additional estimation complexity and potential specification error.

Due to these impediments surrounding use of the CES, the Geometric Mean index was selected as a plausible, and simpler, approximation of the Tomqvist in real-time. The general functional form of the Geometric Mean price index is given by Equation 3.9.

If consumers exhibit Cobb-Douglas utility preferences by holding expenditure shares constant over time, i.e., $s_{1}=s_{2}=\ldots=s_{m}$, then the Geometric Mean is an exact approximation of the Tomqvist. Empirical evidence from CE expenditure data suggests that expenditure shares do not change radically over time, at high levels of aggregation at least. See Table 3.5.

---

39 See Moulton, op. cit.

\[ I_{A}^{G} = \prod_{(i,a) \in (I,A)} \left( \frac{IX_{(a,i)}[0,t]}{IX_{(a,i)}[0,0]} \right)^{i,a} \]

where,

- \( a \) = CPI elementary area
- \( A \) = all CPI areas (U.S. City Average)
- \( i \) = CPI elementary item
- \( I \) = all CPI items (All-items)
- \( 0 \) = base period of the aggregate index
- \( t \) = current month
- \( a \) = base period of the elementary index \((a)\) in area (a)
- \( s_{0} \) = expenditure in period \((0)\) for item \((i)\) in area \((a)\)
- \( s_{t} \) = lower-level index of price change from index base-period \((a)\) to period \((0)\) for item \((i)\) in area \((a)\)
- \( s_{\beta} \) = lower-level index of price change from index base-period \((a)\) to period \((t)\) for item \((i)\) in area \((a)\)

The optimal value of the elasticity parameter \((\sigma)\) that yields the minimum difference between the CES and the Tornqvist averages approximately 0.9 in CPI data. Hence, the assumption of unitary elasticity of substitution over short time intervals may well be a satisfactory approximation. Second, under the assumption of Cobb-Douglas preferences, \( s_{\beta}=s_{t-1}=s_{t}=s_{\beta} \). Because it is implicitly assumed \( s_{\beta}=s_{t-1} \), the denominator price index in the Geometric Mean relative need not be an average of \( \beta \) period indexes. Month-to-month indexes can be computed directly using base-period price indexes \((t-1)\) in the denominator, thus evading any potential specification error caused by use of an average price relative in the denominator.

For any given month \((t)\), the Geometric Mean month-to-month index will differ from the corresponding Tornqvist index to the extent that the \( i,a s_{t-1} \) differ from \( i,a s_{t} \) and to the extent that \( \beta \) poorly predict \( i,a s_{t-1} \). Empirical evidence of upper-level aggregation in the CPI suggests that a Geometric Mean index is biased slightly below a Tornqvist index, when computed using expenditure shares derived from the same base-period as the Tornqvist. Therefore, it is anticipated that use of \( i,a s_{t-1} \) in a Geometric Mean index would tend to produce a lower measure of price change than would result from use of \( (i,a s_{t-1} + i,a s_{t})/2 \) in a Tornqvist index on a consistent basis (e.g., the Geometric Mean month-to-month index averaged 0.006 percent below the corresponding Tornqvist month-to-month index in CPI data over the 1990 to 2000 time period.) A notable exception to this general rule appears to occur in December, where the Geometric Mean index tends to produces a higher measure of price change than the Tornqvist. See Figure 3.5. Moreover, evidence from CPI data further suggests that use of lagged expenditure shares \( i,a s_{\beta} \) in a Geometric Mean index consistently over predicts use of \( i,a s_{t-1} \), albeit by a small amount. See Figure 3.5.

40.93 from 1987 to 2000.
In order to mitigate any specification error associated with use of a Geometric Mean index built with lagged expenditure shares, the BLS decided to adopt an “Adjusted Geometric Mean” approach for the Initial and Interim C-CPI-U. That is, elementary indexes are first aggregated using the Geometric Mean index. Then, the resulting measure of price change is multiplied by an adjustment factor ($\lambda$) that represents the historically observed difference between Tomqvist and Geometric Mean upper-level aggregation of CPI elementary indexes.\(^{42}\) See Equation 3.10.c and 3.10.d. The function of the adjustment factor is to more closely align the Geometric Mean month-to-month index, computed with lagged base-period expenditure weights ($\beta$), to a Tomqvist month-to-month index, computed with contemporaneous monthly expenditures ($t-1$ and $t$).

Finally, the adjusted Geometric Mean month-to-month index is multiplied by the previous-month C-CPI-U index value to obtain the current month C-CPI-U index value. See Equations 3.10.a and 3.10.b. Note that each Interim month-to-month index is chained onto an Interim long-term index value, with the exception of the January index which is chained onto the previous year December index, which is in Final C-CPI-U form. Each Initial month-to-month index is chained onto an Initial long-term index value, with the exception of the January index which is chained onto the previous year December index, which is in Interim C-CPI-U form.

For all months of 2002 and 2003, the adjustment factor has been set equal to unity. BLS plans to use Initial and Interim indexes calculated for 2002 and 2003, in conjunction with Initial and Interim versions of 2000 and 2001, to evaluate further how the Geometric Mean behaves relative to the Tomqvist. A permanent methodology for calculating the adjustment factor will be implemented at a future date.

4. Comparing And Contrasting C-Cpi-U Versions

Difference in Estimation. Initial versus Interim Indexes. The long-term Interim C-CPI-U index for a given year and month can differ from the corresponding and previously released Initial version for three reasons. First, the historical time series to which the Interim index is chained will contain an additional year of Final C-CPI-U indexes. Consequently, the terminal December index value to which the January Interim index is chained may be different from the terminal December index to which the Initial index is chained. Second, the relative expenditure weight patterns used for aggregation may be different. This is highly probable in odd-numbered years – when the expenditure reference period ($\beta$) used for the calculation of Initial and Interim indexes will be different. For example, 2003 Initial indexes will use $\beta = 1999,2000$, but 2003 Interim indexes (released in 2004) will use $\beta = 2001,2002$. The aggregation weights for even-year Initial and Interim indexes will be identical. Third, the adjustment factor ($\lambda$) applied to the Initial Geometric Mean aggregation and the Interim Geometric Mean aggregation may be different.

\(^{42}\) The set of data available to compute the adjustment factor is limited to all time periods for which the Final C-CPI-U has been computed. In 2002, for example, Final C-CPI-U indexes are available only for the 12 months of 2000. In 2003, an additional 12 months of data will become available. Lacking a sufficient time-series of historical data, the adjustment factor for 2002 and 2003 Initial and Interim C-CPI-U indexes will be set equal to $\lambda = 1$. Official methodology for calculating the adjustment factor will be implemented with the calculation of January 2004 Initial indexes, when 36 months of historical data will be available.
Initial and Interim C-CPI-U Upper-level Aggregation Formula

Equation 3.10.a. Initial C-CPI-U
Long-term Price Change

\[ I_A IX_{(x,y,t)}^{G} = I_A IX_{(x,y-1,12)}^{G} \prod_{n=1}^{t} I_A IX_{(y,n(y-1),n)}^{G} \]

Equation 3.10.b. Interim C-CPI-U
Long-term Price Change

\[ I_A IX_{(x,y,t)}^{G} = I_A IX_{(x,y-1,12)}^{T} \prod_{n=1}^{t} I_A IX_{(y,n(y-1),n)}^{G} \]

Equation 3.10.c. Initial C-CPI-U
Month-to-Month Price Change

\[ I_A IX_{(t-1,t)}^{G} = \lambda \prod_{i,a \in I_A} \left( \frac{i,a IX_{(a,t)}^{Log}}{i,a IX_{(a,t-1)}^{Log}} \right) \]

Equation 3.10.d. Interim C-CPI-U
Month-to-Month Price Change

\[ I_A IX_{(t-1,t)}^{G} = \lambda \prod_{i,a \in I_A} \left( \frac{i,a IX_{(a,t)}^{Log}}{i,a IX_{(a,t-1)}^{Log}} \right) \]

where,

- \( a \) = CPI elementary area
- \( A \) = all elementary areas; "U.S. City Average"
- \( i \) = CPI elementary item
- \( I \) = all elementary items; "All-items"
- \( z \) = base period of the aggregate index (NOTE: the All-Items, U.S. City Average C-CPI-U index has a base-period of \( z=December \ 1999 \))
- \( \alpha \) = base period of the elementary index (\( \alpha \)) in area (a)
- \( I_A IX_{(a,t)}^{(i)} \) = lower-level index of price change from period (\( \alpha \)) to month (t) for item (\( i \)) in area (a)
- \( I_A IX_{(a,t-1)}^{(i)} \) = lower-level index of price change from period (\( \alpha \)) to month (t-1) for item (\( i \)) in area (a)
- \( I_A IX_{(i)}^{(a)} \) = expenditure in reference period (\( \beta \)) for item (\( i \)) in area (a) as percent of total expenditures in reference period (\( \beta \)) for aggregate item (\( I \)) in aggregate area (A)
- \( I_A IX_{(i)}^{(a)} \) = aggregate-level C-CPI-U index of price change from period (z) to month (t) for aggregate item (I) in aggregate area (A)

Final versus Interim Indexes. Similarly, the long-term Final C-CPI-U index for a given year and month can differ from the corresponding and previously released Initial and Interim versions for three reasons. First, the historical time series to which the Final index is chained will contain additional years of Final C-CPI-U indexes. Consequently, the terminal December index value to which the January Final index is chained may be different from the terminal December indexes to which the Initial and Interim indexes were chained. Second, the relative expenditure weight patterns used for aggregation most likely will be different. The difference in aggregation weights is the primary distinction in functional form between the Final and two preliminary versions of the C-CPI-U. Variation in the relative monthly expenditures used for the Final from the lagged constant-within-year relative expenditures used for the Initial and Interim may result in differing estimates of aggregate price change. Third, off-cycle
elementary index values may be slightly different between the Final and preliminary versions, as the Final version will use geometric averaging of bounding month indexes.

5. Comparing And Contrasting The C-Cpi-U And The Cpi-U

**Difference in Estimation.** Because the CPI-U and all three versions of the C-CPI-U differ in the set of input elementary price indexes and expenditures used for aggregation, as well as in aggregation formula, each index may yield a different measure of aggregate price change for a given year and month. Table 5.1 summarizes the differences between CPI-U and C-CPI-U index construction.

The Initial C-CPI-U index is published in the same news release as the CPI-U. The long-term index values are not directly comparable, as the CPI-U will be on a 1982-84=100 base and the C-CPI-U will be on a December 1999=100 base. The Initial C-CPI-U month-to-month index for a given year and month will differ from the corresponding CPI-U index by upper-level aggregation method only. The input prices and expenditures will be the same. Similarly, the Interim C-CPI-U month-to-month indexes will differ from the CPI-U month-to-month indexes in aggregation method. In addition, odd-year Interim indexes will differ in input expenditures used for aggregation. Final C-CPI-U month-to-month indexes will differ from the CPI-U month-to-month indexes in all aspects of index construction: (a) the input elementary price indexes will be the same, with the exception of off-cycle bimonthly indexes; (b) input elementary expenditures will be different, and (c) aggregation method will differ.

**Index Simulations.** Simulating official estimation methodology, a biennially weight-updated CPI-U index series was calculated and compared to a Final C-CPI-U index series over the 1990 to 2000 period in order to measure the anticipated difference between the two series. See Figure 5.1. The average difference between the weight-updated CPI-U and the C-CPI-U was 0.32 percent per year over this time period. See Table 5.2. This estimate is at the upper end of the 0.1 percent to 0.4 percent range of upper-level substitution bias estimated in prior BLS research, in which the average annual percent difference was closer to 0.2 percent. Simulations for the 1990 to 2000 period define the range at 0.1 percent to 0.5 percent.

There are several factors contributing to this result. First, the prior average of 0.2 percent was based on data from 1991 to 1995 using the 1987 CPI market basket structure of 207 elementary items and 46 elementary areas (i.e., 9,522 elementary cells). The estimated average in Table 5.2 of 0.3 percent is based on the 1998 CPI market basket structure of 8,018 elementary cells from 1990 to 2000. In order to obtain data on the 1998 structure for years prior to 1998, price indexes and expenditures were approximated using a rough concordance between the old and new structures. Moreover, the current estimate is based on composite-estimated-and-raked biennial expenditures whereas the prior estimates were not. These differences function to produce an average annual substitution effect estimate for the overlap years of 1991 to 1995 that differs by 0.05 percent, i.e. 0.24 percent versus 0.19 percent.

Second, the gap between the weight-updated CPI-U and C-CPI-U appears to have widened in the later part of the decade. The average annual percent difference between the two indexes rose to 0.40 percent in 1996 to 2000, almost double that observed from 1991-1995. Analogously, the percent difference in simulated weight-updated CPI-U and C-CPI-U 12-month indexes steadily increased over the decade. See Figure 5.2.
A likely contributor to the growing gap is increased dispersion in relative elementary index changes. In general, the CPI-U and the C-CPI-U will diverge to the extent that (a) component elementary indexes have rates of inflation that differ from each other, and (b) expenditure shares reflect a shift in consumer purchases toward those item categories that have fallen in relative price. Consequently, when there is more variation in price movement among elementary indexes, there is more room for the Laspeyres-based CPI-U and the superlative-based C-CPI-U to diverge.

Price change in CPI elementary indexes varied more widely during the later part of the 1990s. See Figure 5.3. Two examples of indexes with unusual index movements in 1999 and 2000 are computers and natural gas. The series for ITEM=EE01 “Personal computers and peripheral equipment” decreased by 22.7 percent from December 1999 to December 2000. In contrast, the series for ITEM=HF02 “Utility natural gas service” increased by 36.7 percent over the same interval. The median elementary index change over this time period was 2.2 percent.

The significance of outlier elementary index series can by quantified by excluding them from CPI-U and C-CPI-U calculations and measuring the gap between the resulting “trimmed” indexes. This exercise was performed using the December 1999 to December 2000 12-month simulated index for the C-CPI-U and weight-updated CPI-U, at varying outlier thresholds. See Table 5.3. When computing the indexes using the set of elementary indexes in the middle quartile, i.e. trimming the lower and upper quartiles from the calculations, the percent difference between the CPI-U and C-CPI-U is diminutive, 0.07 percent. The gap increases to 0.2 percent when the set of elementary items used in the calculations is limited to those that increased or decreased in price by 10 percent or less (roughly 75 percent of all elementary items over the December 1999 to December 2000 period). When the set is expanded to include all items exhibiting 20 percent price change (90 percent of all elementary items) the gap increases to 0.3 percent. The difference between the CPI-U and C-CPI-U is greatest when restricting aggregation over the lowest quartile of elementary index price change (1.5 percent) and highest quartile (0.7 percent). These trimmed indexes demonstrate that extreme changes in elementary price indexes cause the CPI-U and C-CPI-U to diverge, and suggest that deflationary outliers contribute heavily to the gap.

6. Inaugural Published Indexes

C-CPI-U indexes were published for the first time in August 2002. Indexes are published for the urban population only. There are no plans at this time to calculate and publish a C-CPI-W. Published C-CPI-U indexes are available for the U.S. City Average only. No regional or local area indexes are published. Moreover, a limited set of indexes are available. See Table 6.1.

Table 6.2, and Figure 6.1, display the inaugural published values of the C-CPI-U in relation to the CPI-U. 43 The most surprising result was the 0.8 percentage point gap between the two estimates of 12-month change during calendar year 2000, which was the only year for which the C-CPI-U index values were published in Final form. The reasons for this were discussed in 5 above; additionally, rounding played a role in exaggerating the differences. The Interim

---

43 It should be noted that the published values for 2000 differ from the simulated values reported in section 5 for a variety of minor reasons. For example, the CPI-U values in Table 5.2 were simulated using a biennial weight update process, whereas the published CPI-U in that year used 1993-95 base period weights. The values in Table 5.2 also are annual-average figures, rather than December-to-December changes.
and Initial values for 2001 and 2002 indicated some significant narrowing of the gap in those years.

7. 2003 Revisions And Most Current Published Indexes

In accordance with its previously-announced schedule, the BLS issued the first set of revised values of the C-CPI-U effective with the release of January 2003 CPI data. While the magnitude of these revisions is expected to be small in the aggregate, in theory each monthly index is subject to two revisions until monthly expenditure data have been generated for that period and introduced two years later. With the availability of consumer expenditure data for 2001, the monthly C-CPI-U index values for 2001 became Final, and the values for 2002 moved from Initial to Interim status. The revised index values and 12-month changes are shown in Table 7.1.

In most cases, the revisions were upward and had the effect of narrowing the differences between the C-CPI-U and the CPI-U index series. At the All Items level, the Final C-CPI-U index for December 2001 increased from 103.6 to 103.9 and the 12-month change rose from 1.0 to 1.3 percent, compared to a 1.6 percent increase in the CPI-U. It should be noted that rounding exaggerates the impact of revisions in this case: the unrounded change in the December 2001 index level is only 0.22 (103.65 to 103.87).

The bulk of the 2001 revisions can be attributed to a small number of item categories. The three largest contributors—owner’s equivalent rent, natural gas, and gasoline—account for about 92 percent of the revision to the All Items C-CPI-U for the period December 2000 to December 2001. In each case the item’s expenditure share moved in the same direction as its relative price change, indicating less than unitary elasticity of substitution. Owners’ equivalent rent rose somewhat more than the average CPI during 2001. Its overall impact results from its very large weight. Natural gas and gasoline prices both rose sharply early in 2001, then fell sharply. Their impacts result from their extreme price movements.

C-CPI-U index levels for 2002 change only because they are spliced onto the revised 2001 Final values. Rates of change in months after December 2001 are unaffected by the revisions, except for rounding. Percent index changes during 2002 and 2003 are still based on the geometric mean formula and weights from 1999-2000, as they were before the revisions. At the All Items level, the CPI-U increased 2.4 percent from December 2001 to December 2002 and the C-CPI-U increased 2.1 percent.

Comparisons with Personal Consumption Expenditure Index. In 1996, the Bureau of Economic Analysis (BEA) at the Department of Commerce introduced chain-type, annual-weighted indexes as its featured measures of real output and prices of the National Income and Product Accounts (NIPA’s). The chain-type price index for personal consumption expenditures (PCE) is similar in concept to the C-CPI-U in several aspects. Both employ a superlative aggregator—BEA’s PCE price index is constructed using the Fisher Ideal index formula. Both are chained, albeit the C-CPI-U at monthly intervals with monthly weights, while the PCE index is chained quarterly, with annual adjustments. On the other hand, there are significant distinctions between the two series. Notably, the PCE index is broader in scope and includes spending on behalf of consumers by employers and government health agencies. The PCE index also uses NIPA weights rather than Consumer Expenditure Survey values, and relies in part on price series other than those used in the CPI.
Table 7.2 compares December-to-December changes in the CPI-U, the C-CPI-U, and the PCE index for the three years for which the C-CPI-U has been published. As the table shows, the C-CPI-U and PCE series are remarkably close given the differences in their scope and construction. This lends some indirect support to the C-CPI-U. Note that the 2002 C-CPI-U estimate is not a Final number.

8. Summary And Directions Of Further Research

The release of the C-CPI-U marks a significant step forward in the U.S. CPI program. A superlative index formula, by employing expenditure data from two separate points in time rather than from a single base period, is designed to reflect substitution among items in response to changes in relative prices and to provide a closer approximation to a cost-of-living index than is possible using the Laspeyres formula. The unavoidable delays in receiving expenditure data result in the inability to produce the final C-CPI-U on the same schedule as the fixed-base CPI-U and CPI-W indexes. The BLS therefore designed the C-CPI-U as its first CPI series to be produced in preliminary form and subject to revision. In August 2002 the inaugural C-CPI-U series were released, followed in February 2003 by the first set of annual revisions.

One notable difference between the C-CPI-U and the existing fixed-base indexes is the interpretation of the sub-indexes—e.g. food, energy, all items less food and energy, etc. In the fixed-base CPIs, the subindexes are separable and additive. In the C-CPI-U the subindexes are conditional upon the behavior of the expenditures and prices of items not included within the subindex, and are therefore not additive. For example, in the C-CPI-U, the behavior of the food index is a function not only of the expenditure and price movements of the food items included within the food aggregate, but also of the expenditure and price movements of all other components within the All Items C-CPI-U. In addition, the All Items C-CPI-U in general is not the weighted average of the subindexes that comprise it, either for all 211 item strata or for sub-aggregates such as food, energy, and all items less food and energy. That is, the C-CPI-U is not consistent in aggregation.

As discussed above in section 3, many challenges faced the BLS in the course of designing the C-CPI-U. Corresponding to each of these challenges are issues that can be addressed in the future with the aid of growing time series of published data. Further analysis may, in some cases, lead to modifications in the way the C-CPI-U is constructed. Among the issues that the BLS is planning to study are:

- Analyze the gap between the CPI-U and the C-CPI-U. Sections 5 and 6 described how the difference between the two estimates of annual inflation widened in the late 1990s. In the first year of published C-CPI-U data, the gap of 0.8 percentage point greatly exceeded the earlier BLS prediction range. The gap was smaller in 2001 than in 2002, and it is reasonable to expect that during the next few years CPI-U inflation estimates will exceed those of the C-CPI-U by more than 0.2 percentage point but less than 0.8 percentage point. BLS staff are now working to develop alternative means of decomposing the inflation differences, in order to more easily identify the contributions of individual index series.44

• Determine the best estimator for the Initial and Interim indexes. The BLS has no experience yet in the C-CPI-U upon which to base estimates of the future size of revisions between Initial and Final index values. The revisions to 2001 Interim indexes were not unduly large, but the fact that they were upward revisions was relatively unusual compared to most years in historical simulations. Thus far, the simulation studies have been insufficient to estimate an adjustment factor reliably, and thus far the BLS has employed a factor of one—i.e., using an unadjusted geometric mean formula—to produce the preliminary index values. Future work will address how best to estimate the adjustment factor when longer time series are available for that purpose. Other potential topics include the use of different factors for different months of the year and, more generally, the analysis of alternative formulas and approaches to predicting the final superlative values.

• Evaluate the monthly weights. The issues of seasonality, volatility, sampling error, and other special features of the Consumer Expenditure Survey data comprise a rich and important area of research. The BLS plans to study how best to deal with seasonal and durable goods, and whether alternative methods of composite weight estimation could improve index performance.

• Evaluate the elementary price indexes. Parallel to the analysis of weight data, research on elementary price indexes would include outlier mitigation, collapsing of areas, composite estimation, and other means of reducing volatility and making the price series more appropriate to their use in a monthly superlative formula.

In the future, with the production and publication of a longer time series, the BLS also will revisit the issue of generating seasonally adjusted C-CPI-U indexes. It should be noted that the method employed in the CPI-U, dependent seasonal adjustment (i.e., aggregating a selected set of seasonally adjusted components to obtain seasonally adjusted aggregates), is inappropriate for the C-CPI-U, since it is not consistent in aggregation. The Bureau also intends to produce and publish statistical measures of reliability for the C-CPI-U such as are presently published for the fixed-base CPIs.
Figure 3.1: 2-Stage Construction of the CPI

Upper-Level Aggregation

80,000

Price Quotes

Elementary Indexes

Aggregate Indexes

8,018

Lower-Level Aggregation

Figure 3.2: Publication Schedule of C-CPI-U Index Versions

<table>
<thead>
<tr>
<th>YEAR OF INDEX</th>
<th>YEAR OF PUBLICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2003</td>
</tr>
<tr>
<td>2003</td>
<td>INITIAL</td>
</tr>
<tr>
<td>2004</td>
<td>INITIAL</td>
</tr>
<tr>
<td>2005</td>
<td>INITIAL</td>
</tr>
</tbody>
</table>
Table 3.1: Percent Difference in Final C-CPI-U annual average indexes:
Geometric averaging versus previous-month imputation of off-cycle elementary indexes.

<table>
<thead>
<tr>
<th>Year</th>
<th>Geometric Mean of Bounding Months</th>
<th>Previous Month Value</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>1.03719</td>
<td>1.03730</td>
<td>-0.011</td>
</tr>
<tr>
<td>1989</td>
<td>1.04480</td>
<td>1.04497</td>
<td>-0.017</td>
</tr>
<tr>
<td>1990</td>
<td>1.04890</td>
<td>1.04879</td>
<td>0.010</td>
</tr>
<tr>
<td>1991</td>
<td>1.03927</td>
<td>1.03952</td>
<td>-0.024</td>
</tr>
<tr>
<td>1992</td>
<td>1.02682</td>
<td>1.02716</td>
<td>-0.033</td>
</tr>
<tr>
<td>1993</td>
<td>1.02779</td>
<td>1.02769</td>
<td>0.010</td>
</tr>
<tr>
<td>1994</td>
<td>1.02442</td>
<td>1.02438</td>
<td>0.005</td>
</tr>
<tr>
<td>1995</td>
<td>1.02409</td>
<td>1.02395</td>
<td>0.014</td>
</tr>
<tr>
<td>1996</td>
<td>1.02727</td>
<td>1.02701</td>
<td>0.025</td>
</tr>
<tr>
<td>1997</td>
<td>1.01819</td>
<td>1.01800</td>
<td>0.019</td>
</tr>
<tr>
<td>1998</td>
<td>1.00912</td>
<td>1.00903</td>
<td>0.009</td>
</tr>
<tr>
<td>1999</td>
<td>1.01714</td>
<td>1.01705</td>
<td>0.009</td>
</tr>
<tr>
<td>2000</td>
<td>1.02782</td>
<td>1.02783</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

Average Annual Difference: 0.001

NOTES:
1. Indexes are for U. S. City Average, All-items and are not seasonally adjusted.
2. Each index is computed as the average annual index divided by the previous-year average annual index.

Table 3.2: Expenditure Reference Periods in the CPI.

<table>
<thead>
<tr>
<th>Expenditure Reference Period</th>
<th>Month Introduced</th>
<th>Terminal Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1917-1919</td>
<td>1919</td>
<td>1939</td>
</tr>
<tr>
<td>1934-1936</td>
<td>Jan 1940</td>
<td>Dec 1952</td>
</tr>
<tr>
<td>1950</td>
<td>Jan 1953</td>
<td>Dec 1963</td>
</tr>
<tr>
<td>2001 - 2002</td>
<td>Jan 2004</td>
<td>Dec 2005</td>
</tr>
</tbody>
</table>
Figure 3.3: Monthly expenditures for ITEM=HF02 (Fuel Oil) as percent of All-item expenditures: Washington, DC and U.S. City Average, 1999 and 2000

Figure 3.4: Value of \( \sigma \) minimizing (Tornqvist – CES)\(^2\)
Table 3.5: Percent distribution of CPI market basket expenditures by major group

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing</td>
<td>38.5</td>
<td>38.5</td>
<td>39.1</td>
<td>38.8</td>
<td>39.3</td>
<td>39.2</td>
<td>37.5</td>
<td>38.6</td>
<td>39.0</td>
<td>39.8</td>
<td>40.1</td>
</tr>
<tr>
<td>Transportation</td>
<td>17.4</td>
<td>16.8</td>
<td>16.9</td>
<td>17.4</td>
<td>18.3</td>
<td>18.1</td>
<td>18.6</td>
<td>18.2</td>
<td>18.0</td>
<td>17.9</td>
<td>18.2</td>
</tr>
<tr>
<td>Food and beverages</td>
<td>16.9</td>
<td>16.6</td>
<td>16.4</td>
<td>16.6</td>
<td>15.7</td>
<td>16.0</td>
<td>16.3</td>
<td>15.9</td>
<td>15.7</td>
<td>15.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Recreation</td>
<td>6.4</td>
<td>6.8</td>
<td>6.5</td>
<td>6.3</td>
<td>6.2</td>
<td>6.4</td>
<td>6.6</td>
<td>6.7</td>
<td>6.8</td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td>Apparel</td>
<td>6.0</td>
<td>6.3</td>
<td>5.9</td>
<td>5.5</td>
<td>5.3</td>
<td>5.4</td>
<td>5.4</td>
<td>5.0</td>
<td>4.8</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>Medical care</td>
<td>5.6</td>
<td>5.7</td>
<td>5.9</td>
<td>5.8</td>
<td>5.6</td>
<td>5.1</td>
<td>5.4</td>
<td>5.4</td>
<td>5.5</td>
<td>5.5</td>
<td>5.6</td>
</tr>
<tr>
<td>Education and communication</td>
<td>4.9</td>
<td>5.1</td>
<td>4.9</td>
<td>5.4</td>
<td>5.5</td>
<td>5.7</td>
<td>5.9</td>
<td>6.0</td>
<td>6.1</td>
<td>6.1</td>
<td>6.0</td>
</tr>
<tr>
<td>Other goods and services</td>
<td>4.3</td>
<td>4.4</td>
<td>4.3</td>
<td>4.3</td>
<td>4.2</td>
<td>4.2</td>
<td>4.2</td>
<td>4.0</td>
<td>4.3</td>
<td>4.3</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Figure 3.5: Average percent difference in Tornqvist and Geometric mean month-to-month indexes, 1990 to 2000, by month
Table 5.1: Components of CPI-U and C-CPI-U Index Construction

<table>
<thead>
<tr>
<th>ELEMENTARY PRICE INDEXES</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower-level Aggregation Formula</td>
<td>Hybrid</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Imputation Method for Off-Cycle Indexes</td>
<td>Previous Month</td>
<td>Previous Month</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ELEMENTARY EXPENDITURES</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure base-period</td>
<td>Biennial, lagged 2-3 years</td>
<td>Biennial, lagged 2-3 years</td>
</tr>
<tr>
<td>Even-year indexes</td>
<td>Biennial, lagged 2-3 years</td>
<td>Biennial, lagged 2-3 years</td>
</tr>
<tr>
<td>Odd-year indexes</td>
<td>Biennial, lagged 3-4 years</td>
<td>Biennial, lagged 3-4 years</td>
</tr>
<tr>
<td>Expenditure current-period</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Weight Update Frequency</td>
<td>Biennial</td>
<td>Biennial</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AGGREGATION METHOD</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper-level Aggregation Formula</td>
<td>Laspeyres</td>
<td>Adjusted Geometric Mean</td>
</tr>
<tr>
<td>Construction of Long-term Index</td>
<td>Fixed-base</td>
<td>Chained</td>
</tr>
<tr>
<td>Construction of Month-to-Month Index</td>
<td>Ratio of Long-term Indexes</td>
<td>Direct</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PUBLICATION SCHEDULE</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>One month lag</td>
<td>One month lag</td>
<td>2 to 13 month lag</td>
</tr>
</tbody>
</table>
Figure 5.1: Simulated Weight-updated CPI-U and C-CPI-U indexes

Table 5.2: Simulated Weight-updated CPI-U and C-CPI-U average annual indexes, relative to previous year average-annual index, 1991 to 2000

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Simulated Biennially Updated CPI-U</th>
<th>Simulated Final C-CPI-U</th>
<th>Percent Difference</th>
<th>Percent Difference from Prior BLS Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>1.04033</td>
<td>1.03927</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>1992</td>
<td>1.03155</td>
<td>1.02682</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>1993</td>
<td>1.02987</td>
<td>1.02779</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1994</td>
<td>1.02663</td>
<td>1.02442</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1995</td>
<td>1.02624</td>
<td>1.02409</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>1996</td>
<td>1.03047</td>
<td>1.02727</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>1.02188</td>
<td>1.01819</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>1.01374</td>
<td>1.00912</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>1.02116</td>
<td>1.01714</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.03259</td>
<td>1.02782</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

Average annual percent difference 0.32 0.19
Average, 1991-1995 0.24 0.19
Average, 1996-2000 0.40 na
Figure 5.2: Percent difference between CPI-U (biennially weight updated) and Final C-CPI-U Simulated 12-month Indexes

Figure 5.3: Elementary Price Index Extremes: Number of Elementary Month-to-Month Price Indexes less than 0.95 or greater than 1.05, by month
Table 5.3: Percent difference between simulated CPI-U and C-CPI-U trimmed December 1999 to December 2000 12-month indexes

<table>
<thead>
<tr>
<th>Trimmed Index Definition</th>
<th>n</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
<th>Percent difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest Quartile (IX &lt; 0.98)</td>
<td>2004</td>
<td>0.9159</td>
<td>0.9023</td>
<td>1.48</td>
</tr>
<tr>
<td>Middle Quartiles (0.98 &lt; IX &lt; 1.064)</td>
<td>4005</td>
<td>1.0259</td>
<td>1.0251</td>
<td>0.07</td>
</tr>
<tr>
<td>Highest Quartile (IX &gt; 1.064)</td>
<td>2009</td>
<td>1.1353</td>
<td>1.1275</td>
<td>0.69</td>
</tr>
<tr>
<td>All-items where 0.80 &lt; IX &lt; 1.20</td>
<td>7356</td>
<td>1.0272</td>
<td>1.0241</td>
<td>0.30</td>
</tr>
<tr>
<td>All-items where 0.85 &lt; IX &lt; 1.15</td>
<td>6929</td>
<td>1.0261</td>
<td>1.0233</td>
<td>0.28</td>
</tr>
<tr>
<td>All-items where 0.90 &lt; IX &lt; 1.10</td>
<td>6054</td>
<td>1.0242</td>
<td>1.0222</td>
<td>0.20</td>
</tr>
<tr>
<td>All-items where 0.95 &lt; IX &lt; 1.05</td>
<td>4180</td>
<td>1.0184</td>
<td>1.0175</td>
<td>0.09</td>
</tr>
<tr>
<td>All-items where 0.99 &lt; IX &lt; 1.01</td>
<td>1122</td>
<td>1.0012</td>
<td>1.0002</td>
<td>0.11</td>
</tr>
<tr>
<td>All-items where 0.995 &lt; IX &lt; 1.005</td>
<td>692</td>
<td>1.0006</td>
<td>0.9987</td>
<td>0.19</td>
</tr>
<tr>
<td>All-items where IX &gt; 1.20</td>
<td>458</td>
<td>1.3445</td>
<td>1.3223</td>
<td>1.65</td>
</tr>
<tr>
<td>All-items where IX &lt; 0.80</td>
<td>204</td>
<td>0.7192</td>
<td>0.6867</td>
<td>4.52</td>
</tr>
</tbody>
</table>

Table 6.1: List of Published C-CPI-U Indexes

<table>
<thead>
<tr>
<th>All-items</th>
<th>Transportation</th>
<th>Special Indexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and beverages</td>
<td>Private transportation</td>
<td>Services</td>
</tr>
<tr>
<td>Food</td>
<td>Public transportation</td>
<td>Commodities</td>
</tr>
<tr>
<td>Food at home</td>
<td>Medical care</td>
<td>Durables</td>
</tr>
<tr>
<td>Food away from home</td>
<td>Medical care commodities</td>
<td>Nondurables</td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>Medical care services</td>
<td>All items less food and energy</td>
</tr>
<tr>
<td>Housing</td>
<td>Recreation</td>
<td>Energy</td>
</tr>
<tr>
<td>Shelter</td>
<td>Education and communication</td>
<td></td>
</tr>
<tr>
<td>Fuels and utilities</td>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Household furnishings and operations</td>
<td>Communication</td>
<td></td>
</tr>
<tr>
<td>Apparel</td>
<td>Other goods and services</td>
<td></td>
</tr>
</tbody>
</table>
### Table 6.2: Inaugural C-CPI-U published indexes for All-items, U.S. City Average compared to CPI-U

<table>
<thead>
<tr>
<th>YEAR-MO</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
<th>CPI-U minus C-CPI-U</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
<th>12-MONTH PERCENT PRICE CHANGE</th>
<th>CPI-U</th>
<th>C-CPI-U</th>
<th>C-CPI-U minus CPI-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003.01</td>
<td>103.3</td>
<td>100.3</td>
<td>3.0</td>
<td>0.3</td>
<td>0.3</td>
<td>3.4</td>
<td>2.6</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>2003.02</td>
<td>103.9</td>
<td>100.9</td>
<td>3.0</td>
<td>0.6</td>
<td>0.5</td>
<td>3.7</td>
<td>2.8</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2003.03</td>
<td>101.7</td>
<td>101.6</td>
<td>0.1</td>
<td>0.6</td>
<td>0.7</td>
<td>3.5</td>
<td>2.6</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>2003.04</td>
<td>101.8</td>
<td>101.6</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>3.3</td>
<td>2.5</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>2003.05</td>
<td>101.9</td>
<td>101.7</td>
<td>0.2</td>
<td>0.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.8</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>2003.06</td>
<td>102.4</td>
<td>102.1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>3.2</td>
<td>2.5</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>2003.07</td>
<td>102.7</td>
<td>102.3</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>2.7</td>
<td>2.0</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>2003.08</td>
<td>102.7</td>
<td>102.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>2.7</td>
<td>2.1</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>2003.09</td>
<td>103.2</td>
<td>102.8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
<td>2.5</td>
<td>1.9</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>2003.10</td>
<td>103.4</td>
<td>102.9</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>2.1</td>
<td>1.5</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>2003.11</td>
<td>103.4</td>
<td>102.8</td>
<td>0.6</td>
<td>-0.1</td>
<td>-0.2</td>
<td>1.9</td>
<td>1.3</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>2003.12</td>
<td>103.4</td>
<td>102.6</td>
<td>0.8</td>
<td>-0.1</td>
<td>-0.2</td>
<td>1.6</td>
<td>1.0</td>
<td>0.6</td>
<td></td>
</tr>
</tbody>
</table>

### Notes:
1. Indexes are not seasonally adjusted.
2. January 2000 to December 2000 C-CPI-U indexes are in FINAL form.
3. January 2001 to December 2001 C-CPI-U indexes are in INTERIM form.
5. CPI-U indexes are CPI-U published indexes divided by the index in December 1999 (169.3), times 100.

### Figure 6.1: Percent difference in 12-month CPI-U and C-CPI-U indexes
Table 7.1: C-CPI-U ALL-ITEMS, U.S. CITY AVERAGE INDEXES: Differences in Original and Revised Values

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Original Published Value in 2001</th>
<th>Revised Published Value in 2003</th>
<th>Revised minus Original</th>
<th>1-MONTH PERCENT PRICE CHANGE</th>
<th>12-MONTH PERCENT PRICE CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Index Relative to December 1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Original</td>
<td>Revised</td>
<td></td>
<td>Original</td>
<td>Revised</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>Price</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Index Relative to December 1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Original</td>
<td>Revised</td>
<td></td>
<td>Original</td>
<td>Revised</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>Price</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NOTES:**
1. Indexes are not seasonally adjusted.
2. A = Final C-CPI-U version
3. A.1 = January C-CPI-U version
4. A.2 = January C-CPI-U version

Table 7.2: December-to-December percent changes in CPI-U, C-CPI-U and PCE Chain Price Indexes

<table>
<thead>
<tr>
<th>Year</th>
<th>Index Series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPI-U</td>
</tr>
<tr>
<td>2000</td>
<td>3.4</td>
</tr>
<tr>
<td>2001</td>
<td>1.6</td>
</tr>
<tr>
<td>2002</td>
<td>2.4</td>
</tr>
</tbody>
</table>

International Working Group on Price Indices - Seventh Meeting
Session 4 - Financial services, including insurances

Chair: Thierry Lacroix, INSEE

Summary of session

Two papers were presented in this session. Viglino is interested in an important but overly often neglected aspect of the treatment of insurance services: taking into account of excess. Ribe also deals with this problem and is more broadly interested in different conceptual and measurement issues related to insurance and financial service indices. The two studies rely on the methods used for the compilation of French and Swedish CPI or of their variant harmonized at European level, the HICP.

- Insurance

Treatment of excess

The modifications of insurance tariffs go more often through a modification of excess - generally a rise - than through a variation of the amount of the premiums, in particular in car insurance. It is therefore important to correctly deal with the variations of excess in CPIs. Viglino and Ribe propose a similar treatment in the form of quality adjustment, the principle of which is applied in France and in Sweden. The adjustment performed consists in correcting the amount of gross premiums of gain expectation, for the policyholder, linked to the variation of excess. The calculation involves an estimate of the claim frequency and an assessment of the distribution of claims according to their amount. Viglino points out that this information may be obtained from trade associations of insurers or estimated from household surveys. The statistical laws of the distributions of claims according to their amount are thus estimated in the case of the French index whereas the only frequency of claims exceeding the amount of the excess is used in the Swedish index, in view of the greater difficulty to obtain detailed information in Sweden than in France. Silver suggests an overlap method as an alternative, based on the assessment by the insurers of the cost of the new policy in the conditions of the previous invoicing. This solution has however two drawbacks: the first one is to make heavier the work requested from the insurers, the second is that the insurer may not be in a position to perform such an assessment.

Ribe thinks that an analogous procedure should in principle be applied to maintain constant over time the value of excess (a nominal value unchanged means a decrease of actual value) but this adjustment can practically be neglected in a low inflation period.

The robustness of the estimations at a given date seems good (Viglino) but frequent updates of the distributions are required to take into account the possible adaptations of risk behaviour (frequency of claims) to the modifications of the level of excess when it is important (Ribe).

Concept of price and weight

A consistent and satisfactory measure of prices and weights is difficult in the case of insurance indices. Ribe presents the method adopted in the European HICP, which combines
an estimate of weights consistent with National Accounts (net approach) and an estimate of prices resting, for practical reasons, on the follow-up of the premiums (gross approach). In relation with a follow-up of net service price, the bias made is widely unknown but is potentially important because the amount of net premiums is of a much lower level than the one of gross premiums. Lacroix points out that a French recent study, with a few simplifying hypotheses, shows that both approaches can lead to very different annual variations but that their evolutions at middle term appear more similar. For Ribe, a gross/gross approach, should be more legitimate in the case of a compensation index because better reflecting the consumer’s point of view whereas the approach retained (net/gross) is more justified in the case of an inflation index. Several contributors acknowledge that the gross approach is not ideal for the prices but that there exists no satisfactory alternative: on the one hand it cannot be contemplated to observe the net service for a monthly index, on the other hand the choice of a gross/gross approach would raise as Turvey has shown other difficulties at the level of sub-indices concerning the goods which enter the field of the products covered by insurance.

- **Banking services**

*Coverage*

The HICP as well as many national indices exclude the expenditures indirectly measured known as FISIM (financial intermediation services indirectly measured). Are this way part of FISIM the fundings performed thanks to the difference between exchange rates for purchase and for sale (currency exchange services) or between interest rates of mortgages and of loans (financial services). For Ribe, the exclusion of FISIM may be justified through conceptual and practical considerations but presents drawbacks. The service price only measures a part of the expenditure and may therefore be difficult to interpret if a substitution occurs between the components explicitly and implicitly measured of the price. It can be however pointed out that this type of problem arises for other expenditure followed-up in CPI such as health, social protection or education expenditure when the part of the cost funded by the state or compulsory social security varies.

*Treatment of stock-brokerage and fund services*

Stock-brokerage of securities (shares, bonds) or fund services are generally invoiced proportionally to the value of the transaction or of the amount of the assets held. The choice of the service unit the price of which will be followed-up in the index is therefore not obvious. Now Ribe shows that this choice has important consequences, in view of the great amplitude of fluctuations of financial markets these last years.

The basic question is the choice between a unit defined in volume (a portfolio of shares defined by its composition - alternative A) or a unit defined in value (the value of this very portfolio - alternative B). To keep constant over time the unit transaction leads to an indexing, either on stock-exchange indices (alternative A), or on the general price index (alternative B). Ribe prefers this latter choice for two reasons. The first one is conceptual, the service appearing more linked in the eyes of the purchaser to the value of the securities held rather than to their number. The second concerns the relevance of the indices obtained with alternative A: the strong variations of the price index of financial services are annoying in that case and their impact on the overall index is excessive (as far as 0.5% over the annual change).
Consistency of the choices between price and volume index

To a price index may always be associated a volume index. With the A alternative, Ribe thinks that the volume index may be rather stable, whereas it will be probably very variable with alternative B. The point is then which of these alternatives best reflects the activity of banking services: unfortunately the answer to this question is not obvious.

Consistency of choices between consumer price indices and producer price index

Ribe considers that the objectives of the two indices are not the same, which may justify different approaches. Nevertheless he proposes, after an analysis conducted from the producer’s point of view and calling for considerations in terms of production and productivity to also retain the B alternative for the producer price index.

The discussion shows that many participants judge that different approaches may justify different conceptual choices between price indices and volume indices on the one hand, consumer price index and producer price index on the other hand.

Recommendations for statistical agencies

- Insurance

1. In the compilation of the index of insurance services, the excess variations should be treated as quality adjustments. An appropriate method is to perform explicit assessments of quality differences through the use of statistical laws of risk distribution according to their amount.

2. In combination with weights reflecting the service value, the follow-up of gross premiums is the method commonly used for practical reasons (net/gross approach). The constraints of a monthly index like the CPI make difficult any other choice. It would nevertheless be advisable to study more in-depth the extent of the bias linked to the use of this proxy.

- Banking services

3. The inclusion of FISIM in the scope of banking services indices remains controversial, for theoretical and practical reasons. The objectives of the index, the quality of information sources and the cost of treatments are to be taken into account in the choices made.

4. The selection of the service unit is difficult in the case of transactions on financial securities and the management of portfolio. It is however crucial because its impact on overall index may be strong in case of important fluctuations of financial markets. The two possible options are to define the constancy of the unit in volume or in value: the second appears to be preferable from the point of view of the consumer’s behaviour and because the results obtained are more credible.

5. The decisions made on the choice of the service unit for the CPI and the PPI should be considered together. The same way the consequences of these choices on volume indices should be examined.
Insurance and quality adjustment: excess and option-cost method

Lionel Viglino
Eurostat, Luxembourg

Abstract: This paper shows an option cost method for excess quality adjustment in the field of insurance index. Concept, estimator and estimates are given. Estimates are rather old but come from true statistical production.

Keywords: Index, quality adjustment, insurance, option-cost, excess, lognormal distribution, gamma distribution.

Excess is the first part of a damage, which stays to the policyholder burden.

Excess might a fixed amount. An excess of 1000 € means that the policyholder receives nothing for an 800-€ damage and receives 5000 € for a 6000-€ damage.

Excess might be proportional or proportional with limit. 10% with a maximum of 1000 € means that the policyholder would receive 720 € for an 800-€ damage.

A variation of excess is clearly a variation of quality. Often insurance companies increase the excess instead of raising the premium.

It is possible to assess the expected cost for a policyholder due to excess. With this estimate it is possible to quality adjust the premium with an option cost method.

1 The views expressed in this paper are those of the author and do not necessarily reflect the policies of Eurostat.
1. Expected cost due to excess

1.1 Fixed excess $F$

$F$ is the excess and $C$ is the value of damage random variable and $dP^C$ is the joint probability density function.

\[ \text{Excess average cost} = \int_0^F c \times dP^C + F \times \text{Prob}(c>F) \]
1.2 Proportional (Tx) excess with minimum (Fmin) and (Fmax)

\[ \int_{F_{\min}}^{F_{\max}} c \times dP^c \]

\[ T_x \times \int_{F_{\min}}^{F_{\max}} c \times dP^c \]

\[ F_{\max} \times \text{Prob}(c>F_{\max}) \]

Fmin Fmax value of damage

average cost due to excess = \[ \int_{F_{\min}}^{F_{\max}} c \times dP^c \]

\[ T_x \times \int_{F_{\min}}^{F_{\max}} c \times dP^c + F_{\max} \times \text{Prob}(c>F_{\max}) \]

2. Distribution of damage

Actuary and other Staff of insurance companies evaluate or use these frequencies and distributions of damages. This is the core of their work. Also with household survey, National Institute of Statistics can collect the necessary information to estimate these distributions.

The experience shows that damages are always following a lognormal distribution except for fire. In case of fire, value of damage can be very high more often because a fire can spread to another place and increase the destruction. The frequency distribution for fire must have a heavy end. The Gamma distribution is possible.

With a major insurance confederation: APSAD, and an INSEE household survey (PCV 96), damage parameters had been estimated in France in 1995 and 1996 for cars and household insurances.
2.1 "Cars"

2.1.1 Distribution of damage (expressed in Euro)

- **Thefts of cars and fires**: lognormal distribution with $m=7.84$ et $\Gamma=0.98$ (Household survey PCV 96)

- **Thefts of accessories**: lognormal distribution with $m=5.91$ et $\Gamma=0.96$ (PCV 96)

- **Plate glass**: lognormal distribution with $m=5.40$ et $\Gamma=0.83$ (PCV 96)

- **Car crashes**: lognormal distribution with $m=7.09$ et $\Gamma=1.08$ (PCV 96 and insurance confederation APSAD). The expected value is $\exp(m+\Gamma^2/2)$, around 2150€.

- **Civil liability in car crashes (bodies and cars)**: lognormal distribution with $m=6.67$ and $\Gamma=1.0$ (APSAD 94)

2.1.2 Frequencies of damage (Insurance confederation APSAD 95 and household survey 96)

- **Thefts of cars and fires**: 2.35 % (burning is about 0.2 %)
- **Thefts of accessories**: 0.70 %
- **Plate glass**: 9.40 %
- **Car crashes**: 10.5 %
- **Civil liability in car crashes**: 6.00 %

2.2 "Houses"

2.2.1 Distributions (in Euro)

- **fires**: Gamma distribution with $\gamma=1/2$ and $\alpha=1/1220$ (APSAD 93). The end of distribution is rather heavy, because some fire can cause important value of damage)

- **water damage**: lognormal distribution with $m=6.59$ et $\Gamma=1.0$ (PCV 96)

- **Plate glass**: lognormal distribution with $m=6.57$ and $\Gamma=0.97$ (PCV 96)

- **Thefts**: lognormal distribution with $m=7.50$ and $\Gamma=1.04$ (PCV 96)

- **Civil liability**: lognormal distribution with $m=6.61$ and $\Gamma=1.18$ (PCV 96)

2.2.2 Frequencies (APSAD and PCV)

- **Fires**: 1.1 %
- **Water damage**: 3.5 %
- **Plate glass**: 2.5 %
- **Thefts**: 1.2 %
- **Civil liability**: 1.5 %

3. Cost-option estimator of excess

3.1 with fixed and absolute excess

- if value of damage $i \rightarrow N(m_i,\Gamma_i)$ with excess $F_i$,
\[
\text{cost}_{\text{sinistr}}^i = \exp(m_i + \frac{\Gamma_i^2}{2}) \times \Phi_n \left( \frac{\log F_i - (m_i + \frac{\Gamma_i^2}{2})}{\Gamma_i} \right) + F_i \times \left[ 1 - \Phi_n \left( \frac{\log F_i - m_i}{\Gamma_i} \right) \right]
\]

with \( \Phi_n \) distribution function of standardized normal distribution (Gauss, 0, 1).

- if value of damage \( i \to \gamma(p_i, \theta_i) \) with excess \( F_i \),

\[
\text{cost}_{\text{sinistr}}^i = \frac{P_i}{\theta_i} \times \Phi_\gamma(\theta_i F_i; p_i + 1) + F_i \times \left[ 1 - \Phi_\gamma(\theta_i F_i; p_i) \right]
\]

with \( \Phi_\gamma \) distribution function of Gamma distribution.

### 3.2 with fixed but relative excess

- if damage \( i \to N(m_i, \Gamma_i) \) with excess \( F_i \),

\[
\text{cost}_{\text{sinistr}}^i = \exp(m_i + \frac{\Gamma_i^2}{2}) \times \Phi_n \left( \frac{\log F_i - (m_i + \frac{\Gamma_i^2}{2})}{\Gamma_i} \right)
\]

with \( \Phi_n \) distribution function of standardized normal distribution.

- if damage \( i \to \gamma(p_i, \theta_i) \) with excess \( F_i \),

\[
\text{cost}_{\text{sinistr}}^i = \frac{P_i}{\theta_i} \times \Phi_\gamma(\theta_i F_i; p_i + 1)
\]

with \( \Phi_\gamma \) distribution function of Gamma distribution.

### 3.3 with proportional excess without maximum and minimum

- if damage \( i \to N(m_i, \Gamma_i) \) with excess \( F_i \) (%),

\[
\text{cost}_{\text{sinistr}}^i = F_i \times \exp(m_i + \frac{\Gamma_i^2}{2})
\]

- if damage \( i \to \gamma(p_i, \theta_i) \) with excess \( F_i \) (%),

\[
\text{cost}_{\text{sinistr}}^i = F_i \times \frac{P_i}{\theta_i}
\]
3.4 with proportional excess limited by minimum and maximum

- if damage \( i \rightarrow N(m_i, \Gamma_i) \) with \( F_{\text{min},i} \) minimum excess, \( F_{\text{max},i} \) maximum excess, 
  \( \alpha_i \) rate between maximun and minimum.

\[
\cos^i_{\text{sh.min}} = \exp(m_i + \frac{\Gamma_i^2}{2}) \times \left[ (1 - \alpha_i) \times \Phi_n \left( \frac{\log F_{\text{min},i} - (m_i + \Gamma_i^2)}{\Gamma_i} \right) + \alpha_i \times \Phi_n \left( \frac{\log F_{\text{max},i} - (m_i + \Gamma_i^2)}{\Gamma_i} \right) \right] \\
+ F_{\text{max},i} \times \left[ 1 - \Phi_n \left( \frac{\log F_{\text{max},i} - m_i}{\Gamma_i} \right) \right]
\]

with \( \Phi_n \) distribution function of standardized normal.

4. It is not negligible

Experience shows that adjustment for excess worths few percents on elementary index, and very often for an entire company, which is not negligible.

This result is consistent with a fast comparison made with frequency of damage, excess and gross premiums.

References


Astin Bulletin.

Bulletin de l’Institut des Actuaires Français.


International Association for the study of insurance Economics (Genève).


L’Argus (Journal international des assurances).
Financial Services in Swedish Price Indices

Martin Ribe

Statistics Sweden

The author is grateful to colleagues at Statistics Sweden for helpful comments and stimulating discussions.

Abstract: The paper deals with issues that have been found important for the treatment of financial services in Swedish price indices. In particular the definition of the unit transaction for fund and stock-brokering services with charge proportional to transaction value is crucial. Such services have more than a negligible extent in Sweden, also for consumers, and different definitions of the unit transaction could lead to notably different results for the overall index of consumer prices. The definition of the unit transaction has great impact also on a volume index for financial services.

1. Coverage of financial services

The Swedish CPI (Consumer Price Index) covers insurance services since 1993 and partially banking services since a few years later. The HICP (Harmonised Index of Consumer Prices) for Sweden also covers those services except building insurance, and since 2001 the HICP furthermore covers stock-brokering and fund services. The treatment in the CPI was basically designed by Lönnqvist (1992) and is described by Statistics Sweden (2001). Presently a Swedish Producer Price Index of banking services is being developed.

For 2003 the weight of insurance is 0.96 per cent in the national CPI and 0.64 per cent in the HICP, and the weight of banking services etc. is 0.67 per cent in the national CPI and 1.80 per cent in the HICP. Building insurance is covered in the insurance weight of the national CPI but not that of the HICP, while conversely stock-brokering and fund services are covered in the weight of the HICP but not that of the national CPI.

It may be noted that the mentioned weights for financial services are not really negligibly small, whence large movements in recorded prices for these services may have a notable impact on the overall indices for Sweden.

The present paper focuses on some specific potentially problematic matters of principle which have had to be dealt with and which have also been given consideration in connection to the international harmonisation work for the HICPs and other statistics.

1 The views expressed in this paper are those of the author solely.
2. Insurance – adjustment for excess, gross vs net principle

Insurances covered in the CPI are home insurance, building insurance, motor-car insurance, and private sickness and accident insurance. The “home insurance” customary in Sweden is a comprehensive insurance for households covering damage or loss of movable property by fire, theft or flooding from pipes, and to some extent also liability etc. The coverage of motor-car insurance includes the compulsory traffic (liability) insurance as well as insurance against material damage and theft. Building insurance is covered in the CPI but excluded from the HICP, as being regarded a part of the capital costs of owner-occupied housing (cf. Council of the EU, 1998).

For the index computation, premium amounts for insurance policies are followed over time. This is done for a sample of policies from the few major insurance company groups. For the premium amounts (the prices) collected, the conditions of each sampled policy, as well as the characteristics of the policyholder, are kept constant over time. In this way the prices followed become comparable over time. (Cf. Statistics Sweden, 2001, p. 58f.)

2.1 Adjustment for excess change

To keep the conditions of the insurance policies constant may not be directly feasible in cases where an insurance company changes the excess. The excess is a fixed amount that the insurance company deducts from the value of damage in its compensation for a case of damage. When the excess for an insurance policy is changed, the tariff may not include any option comparable to the previous excess. Then the change in the excess is taken account of by a form of quality adjustment.

The adjustment is done in the way that the annual premium as of the base period is decreased by subtraction of an amount computed as the difference between the current and the previous excess, multiplied by the annual claims rate. By rough approximation, the subtracted amount corresponds to the change in the actuarial risk premium, i.e., in the mathematical expectation of claims receivable to the policy-holder, due to the excess change. In plain words, it is roughly how much less (or more) the insurance protection becomes worth to the policyholder, by the excess change.

A practical difficulty of this method is that the required data on the claims rate may not be readily available, and previously estimated figures may have to be used.

Those gradual real-term changes in the excess, that result from the inflation, could and should be accounted for in a similar way. This is presently not done and is also less urgent in a time of slow inflation.

To elaborate a little on the conceptual basis, let $F(x)$ denote the cumulate probability distribution function of the values of damages. And let $r(x)$ denote the annual rate of damages of value exceeding $x$, or in other words, the annual rate of claims if the excess of the insurance is set to $x$. Then if the excess is set to $b$, the annual actuarial risk premium, i.e., the mathematical expectation of annual claims, becomes

$$r(0) \int_b^\infty (x-b) \, dF(x).$$
Now assume that the excess is raised from $b$ to $c$, say (with $c > b$). Then the resulting drop in the risk premium becomes

$$r(0) \int_{b}^{c} (x - b) \, dF(x) + r(0) (1 - F(c)) (c - b)$$

$$= r(0) (1 - F(b')) (c - b)$$

$$= r(b') (c - b),$$

for some number $b'$ such that $b \leq b' \leq c$, under the reasonable assumption that the function $F$ is continuous.

The method used for quality adjustment with respect to excess change is approximate, since an observed (or guessed) claims rate has to be used as a proxy for $r(b')$, which would ideally have had to be used. In view of this approximation, which might be rather coarse, it might in a way perhaps be just as well if the rate used is not quite recently observed. For a recent estimate would be likely to approximate $r(b)$ rather than the wanted $r(b')$, which might be notably lower.

Viglino (2003) gives a more exact approach, using a parametric estimation of the distribution of the values of the damages. But then again a practical difficulty might be to obtain the statistics on damage values needed for the estimation.

### 2.2 Gross and net insurance premium amounts

A crucial issue concerns the choice between gross and net premium amounts, for the purposes of the price index. As defined by the Commission of the EU (1999), the gross premium is the insurance premium in the usual sense, i.e., the price paid by the policyholder for the insurance policy. The net premium amount or service charge is the net amount obtained when claims and changes in actuarial provisions are subtracted from the insurance company's premium income, including the yield on premium reserve etc. ("premium supplements").

By the rules for the HICPs (Commission of the EU, 1999), the weights for insurance shall reflect the service charge, while the prices followed shall be the gross insurance premiums. The use of service charge for weights follows the rules for national accounts. The apparent discrepancy between use of net for weights and gross for prices in the HICPs is obviously motivated by the unfeasibility of collecting monthly data on the current service charge. That is, the readily available gross premium is apparently taken as a proxy for the less available service charge, in the price collection. As is noted by Eurostat (2001, Sect. 4.9.2), it may however be questioned whether the changes over time in the gross premium well reflect those of the service charge.

Admittedly, at least for the Swedish CPI no investigation has been made whether it might be feasible to obtain quarterly data on the service charge from the quarterly financial reporting of the insurance companies. For the use of such quarterly data for a quarterly price index, it would have to be shown that that the available data are both detailed enough and conceptually suited to that use, which a priori does not seem certain.

In contrast to the HICPs, the national CPI for Sweden is meant primarily for compensation purposes (cf. Statistics Sweden, 2001, p. 8; SOU, 1999; Ribe, 2001). Actually, a true cost-of-
living index (coli) is since very long time explicitly stated to be an ideal target for the Swedish CPI. This implies that it should be most relevant to consider the gross premium, rather than the net premium, in the index. The gross premium is the cost for the insurance that is seen and felt by the consumer. From a consumer perspective it seems abstract and artificial to consider only the service charge as a price for insurance, and the rest of the gross premium as some “transfer” rather than a consumer cost.

2.3 Practical implications

Consequently, when insurance was first introduced into the Swedish CPI, the weights as well as the prices were based on the gross premiums. This had some practical implications for the computation of the weights, which is done annually, as the Swedish CPI is a chain index with annually updated weights. The annual weight computation is based on data on household consumption expenditure from the national accounts, where only the service charge is recorded as consumer expenditure for insurance. On the other hand, e.g. expenditure of motor-car insurance claims on car repair is included as household consumption expenditure on car repair, and similarly for other kinds of claims.

So to make the CPI weights follow the gross principle, the CPI staff had to re-distribute some amounts among the expenditure amounts obtained from the national accounts. That is, estimated amounts corresponding to insurance claims were re-distributed, from expenditure on car repair, furnishings, etc., to expenditure on insurance premiums.

From year 2000 however the net principle is followed for the CPI weights, as well as for the HICP weighs. This change was made for practical reasons, although it was not ideally desirable in view of the aims of the national CPI. The practical advantages are of course that double procedures, differing between CPI and HICP, are avoided, and that the weight computation is simplified. The governing CPI Board was not entirely happy about the change, but the majority of the Board considered the theoretical drawbacks less important than the practical advantages.

2.4 Negative weights?

An implication of the net principle is that weights for insurance may potentially become negative. Apparently negative weights in a price index would not make very good sense.

In Sweden no problems with negative weights have been encountered. Although claims may exceed premium income, the net becomes positive when the “premium supplements”, i.e., the yields on reserves etc. are also included.

3. Banking services – delineation of the coverage

Banking service prices covered in the CPI include charges for common financial consumer services of banks and the Post, such as transaction charges for payment services, and annual charges for credit card, Internet banking, and safe deposit box. A judgmentally made sample of such services is followed in the price collection, comprising the few major banks and the Post.

Following the rules for the HICPs (Council of the EU, 1998), the coverage in the Swedish CPI and HICP excludes FISIM, i.e., Financial Intermediation Services Indirectly Measured. The FISIM comprise those parts of financial services that are charged for by way of the net
interest earnings of banks etc., obtained in principle by the gap between the interest rates for lending and borrowing.

In banking practice there is however hardly a sharp borderline between FISIM and explicitly charged services. For instance, payment services may be charged for partly by explicit charges, and partly in a hidden way by the interest earned by the service provider on the money handled in the payment flows.

So in a way the price index may show only a part of the price, namely, the part charged for explicitly. It may then be somewhat problematic to interpret the index for bank charges, as changes in the index reflect changes in the explicit part of the price, rather than the full price.

Diagram 1 in the annex displays the HICP sub-index for “Financial services n.e.c.” (Coicop 12.6.2, or essentially banking services), together with the overall HICP index. It is seen that the prices of the financial services in question have shown dramatic rises in the past few years, far beyond what has been the case for consumer prices on average.

At first sight the indicated huge price rises for financial services may seem somewhat strange. But with regard taken to the fact that they may reflect parts of prices rather than full prices, they may not be too unnatural. There may of course also be other explanations, for instance possibly that service providers might raise charges on laborious teller services in order to encourage customers to swap to more cost-efficient services such as internet banking. As teller services and Internet services are treated as different products, the index does not recognise consumers’ possible substitution gains from such swaps.

It may be noted that plans have been reported for an inclusion of FISIM in the Australian CPI (Frost, 2001; Woolford, 2001).

3.1 Currency exchange

Currency exchange services in Sweden are still provided by banks but nowadays mainly by specialised currency exchange companies. Usually no explicit charge is made, but the exchange services are entirely charged implicitly by the spread between the selling and buying exchange rates.

By the rules for the HICPs, such implicit charges for exchange services are to be regarded as FISIM and are thus to be excluded from the HICPs (Eurostat, 2002, Art. 3(2c)). They are also excluded from the national CPI, like brokering and fund services (cf. below).

From a theoretical point of view one might perhaps question whether currency exchange ought not anyway to be covered in a price index, even if indirectly charged for. In practice it might however be problematic to collect prices for the indirectly charged exchange services. The difference between the selling and buying rates might possibly not be an entirely adequate measure. Namely, there might possibly be some asymmetry between selling and buying in the exchange companies’ business aims and turnover vis-à-vis consumers.

4. Stock-brokering and fund services – the issue of unit transaction

The total savings in stocks and in stock and security funds of Swedish households amounted in mid 2001 to 120, and in mid 2002 to 70 per cent of the total annual household
consumption. The fall between the two years is largely due to steeply falling stock-market prices. Anyhow it is seen that Swedish households possess considerable savings in stocks and funds. Accordingly stock-brokering and fund services have more than a negligible role for Swedish households.

Stock-brokering and fund services are provided by banks and specialised brokering companies. Households probably still mostly access these services by way of banks.

Stock-brokerage is charged as a flat minimum rate for small selling and buying transactions, and as a charge proportional to the transaction value for transactions with value exceeding some €4 500.

Fund savings are charged for by a charge proportional to the asset value, and in many cases also by charges proportional to the transaction value at selling and/or buying of fund shares.

From 2002 the coverage of the HICPs is extended to include also services with explicit charges proportional to transaction value ("ad valorem" charges), and rules for their treatment have been established (Commission of the EU, 2001; Eurostat, 2002). This applies to the Swedish stock-brokering and fund services, for which thus price-collection is now implemented.

For stock-brokerage price collection is made from the few major banks, for a small judgmental sample of service levels. For fund services price collection of charge rates is done from a magazine, by a systematic sample (with random starting-point), from a comprehensive list of savings funds.

A proposal has been put forward to extend the coverage also of the national CPI to include such services (Sjögren, 2000). The governing CPI Board however rejected the proposal for the time being, as it was felt that further consideration was needed whether these services are appropriately to be included in consumption as covered by a compensation index.

4.1 The unit transaction — definition

For charges defined as a proportion of the transaction value, the price followed by the index shall be taken as that proportion of a representative unit transaction. This implies that changes in the value of the representative unit transaction, as well as changes in the proportion defining the charge, shall be shown as price changes in the index (as stipulated by Commission of the EU, 2001, Art. 2(2) and 3(1c)).

In practice this means that a sub-index for a service charged in this way is computed as a product of two factors, both indices: One index for the proportion defining the charge, and one index for the value of the transaction unit (cf. Eurostat, 2002, Art. 2).

A crucial issue then consists in the choice of definition for the transaction value. This definition is logically linked to the choice of an adequate index for the value of the transaction value. E.g., for stock-brokerage, a priori a couple of potential main alternatives may here be identified:

(A) Alternative: The unit transaction would be defined as the buying or selling of a constant number of shares. The index for the value of the unit transaction should then be given by a stock-market price index.
(B) *Alternative:* The unit transaction would be defined as the buying or selling of an amount of shares with a constant value in real terms, that is, an amount of shares with a value that can buy a constant amount of consumer products. The index for the value of the unit transaction should then be given by the overall HICP (or CPI).

The Guidelines agreed for the HICPs (Eurostat, 2002, Art. 4(1b)) imply that alternative (B) is to be followed for the HICPs.

4.2 *The unit transaction – motivation*

There have been some discussions on the choice between alternatives (A) and (B), in international harmonisation work and elsewhere (cf. Woolford, 2001). However there appears to be strong reasons to consider alternative (B), the one chosen for the HICPs, to be the appropriate one for a consumer price index.

The stock-brokerage is a charge for a service, not for the shares bought or sold. But the benefit to the consumer of this service is directly related to the purchasing power corresponding to the value of the shares that are bought or sold. By alternative (B) this benefit is kept constant, in the price comparison over time in the index.

And to keep the consumer benefit constant for prices followed in the index, as is met by alternative (B), is what is needed in a consumer price index. This apparently holds equally well for a pure price index as for a cost-of-living index. Namely, in order to show pure price changes, a pure price index has to compare like with like over time. Here “like” has to be interpreted in a consumer perspective, as a consumer price index is an input index to the consumer sector. The need to compare like with like is reflected in the urge for appropriate quality adjustment in the HICPs (as stated by Commission of the EU, 1996, Art. 5). So alternative (B) is apparently the adequate way, at least for a consumer price index.

Alternative (A) on the other hand seems hard to motivate, for a consumer price index. Lack of theoretical motivation may be acceptable in cases where the practical consequences are negligible. But for Sweden, in the present case, they are not. If alternative (A) had been chosen it could have had notable consequences that could hardly be defended, as will now be demonstrated.

4.3 *Consideration of a stock-market index*

Diagram 2 shows the development of a well-known comprehensive stock-market price index for the Stockholm stock exchange, the AFGX (Affärsvärldens generalindex). This index has been produced since long time by the business magazine Affärsvärlden (2003) and is publicly available for download from their web site. The curve in diagram 2 shows the AFGX only for the 15th of each month (or the first trading day alter the 15th), for comparability with the Swedish CPI and HICP which have the 15th as the target day of the monthly price collection.

As is seen in the diagram the Swedish stock prices have been subject to very dramatic fluctuations over the past few years. At an accelerating pace the stock index skyrocketed by a factor of about four in five years between 1995 and 2000. Then it suddenly stopped, turned, and steeply plunged back almost all the way down, to a level of six years ago.
In alternative (A) such dramatic changes in stock-market prices would be shown as price changes for stock-brokering services. This would apparently be an artificial and imaginary effect which says little about reality. For instance, in a situation when stock-market prices rapidly fall down to one third of their previous level, it indeed seems far-fetched to claim that stock-brokering services have become correspondingly cheaper, just because of that. These artificial consequences of alternative (A) apparently make it essential to avoid that alternative.

Volatility should not be feared if it tells the truth, but volatile developments of stock prices hardly tell any truth about stock-brokerage developments.

4.4 Simulation results

Diagrams 3 and 4 show results of a simulation to demonstrate the potential effect of alternative (A) on the overall HICP and the corresponding inflation rate.

The grey thick curves in diagrams 3 and 4 indicate how the HICP and the inflation rate might have evolved, if fund and stock-brokering services had been covered by means of alternative (A) during the past years. It is seen that the effect would have been quite notable.

4.5 Assumed weights etc.

As is not unusual in index production generally, some rather coarse assumptions have to be made about unknown weights on the lowest levels of aggregation.

For stock-brokerage there is as mentioned both a minimum brokerage and a proportional brokerage. A simple way out is to compute one index for the minimum brokerage and one for the proportional part, and then weight them together. The weights for this weighting have to be rather coarsely “guesstimated”, at best with some support from statistics concerning size-distributions of shareholders’ share possessions. Even so, those weights should in principle be price-updated, to avoid a technical bias.

For fund services there are as mentioned charges proportional to asset value, and often also proportional charges on buying and/or selling of fund shares. Here the index factor capturing changes in the charge proportion is based on a weighted mean price ratio for each fund, computed as a weighted mean of the ratio of asset charge proportions and the ratio of buying plus selling charge proportions. Again the weights have to be “guesstimated”, now from a coarse assumption on the rate of buying and selling of fund shares.

The experience so far shows, for both brokerage and fund services, that changes in the charge proportions may occur but do so very seldom. So the index for fund services with ad valorem charges mostly follows the HICP, even though the very infrequent proportion changes may have a noticeable impact when they do happen.

5. Mortgage interest

For completeness it may be mentioned that changes in mortgage interest are taken account of in the component of owner occupied housing in the national CPI in Sweden, like in some other countries (cf. Statistics Sweden, 2001; SOU, 1999; Goodhart, 1999). However the mortgage interest is here included not as a price of a financial service as such, but rather as a component in one out of some alternative indicators for the capital cost of owner occupied
housing. The treatment of owner occupied housing in price indices is of course in itself a very intricate issue, far from the scope of this paper.

6. Financial services in Producer Price Indices

Presently a Producer Price Index (PPI) for banking services is being developed for Sweden, and one for insurance services is planned. The PPIs for services have quite another main use than both the CPI and the HICP. The PPIs for services are needed primarily in the production of national accounts, for computation of volume indices and time series at constant prices ("deflation"; cf. Eurostat, 2001, Sect. 1.3). The PPIs for services shall cover services to the business sector, not consumer services, which are already covered in the CPI and HICP. Like for the CPI and HICP, FISIM is excluded from the coverage of the PPIs. So for the PPIs too, only explicit charges are to be covered.

Banking services to businesses are to a large extent sold as packages. For small companies there may be standard packages with various options, for which prices might be feasibly followed over time in the index. For large companies on the other hand the banks offer more specialised deals, where negotiated price reductions and special conditions usually come into play. So for the following of prices for banking services to large companies the task seems trickier. Possibly some variant of the "model pricing" method may be used, although elements of judgment may be hard to avoid (cf. Eurostat, 2001, Sect. 3.1.1.1). The development of the survey method has just begun, so it is not possible to be more specific right now.

Like for the CPI and the HICP, the exclusion of FISIM may not be quite unproblematic. In their pricing of service packages to companies, banks are likely to primarily care for their total earnings potential, including earnings both by charges and by interest. So again we may have a price index that covers only a part of the price and could by this show seemingly unnatural movements. Whether such effects may have disturbing consequences of course largely depends on the procedures and contexts where the indices are used. Cf. also TEG-PPI (2002, Sect. L).

6.1 Ad valorem charges in PPIs – the issue

Ad valorem charges, or charges proportional to the transaction value, may be of minor importance in banking services to business companies. But to the extent that they do occur, there is of course again the choice between the alternatives (A) and (B) mentioned before. For PPIs the harmonisation rules of Eurostat (2001, Sect. 4.9.1) imply that here alternative (A) is to be used.

With regard to basic principles, the choice here between alternatives (A) and (B) is not perfectly analogous to that for the HICPs. Namely, the PPIs are output indices, while the HICPs are input indices; cf. Triplett (1983, especially Sect. 5.4.2), who recognises how output as well as input can be disaggregated into characteristics, but differently.

For now here merely some preliminary remarks will be made on the somewhat controversial issue, reflecting the author's personal views.
6.2  Ad valorem charges in PPIs – preliminary discussion

Considering the use of the index for the national accounts, one may notice that while alternative (A) makes the stock-market price fluctuations show up in the price index, alternative (B) on the other hand makes them show up in the volume index.

It may be then argued that a volume index should be adjusted for all effects of price fluctuations, also fluctuations in stock-market prices. This would be an argument in favour of alternative (A), for this context. By alternative (A) the volume series would tend to be stabilised, and so would derived series on productivity. A stable series on productivity would evidently seem better interpretable than a volatile one.

But again, if volatility shows the truth, it should in principle not be hidden. Apparently the current productivity in agriculture has to depend partially on volatile weather conditions, and not purely on cost-efficiency in farming work. It might then possibly also be considered natural that productivity in stock brokering partially depends on volatile conditions of the stock market. Generally, elimination of volatility cannot be an end in itself, but it is merely artificial and misleading kinds of volatility that should be eliminated. In the context of HICPs alternative (A) would lead to an artificial kind of volatility, as was argued above.

So the question now appears to be whether changes in stock-brokering volume due to stock-market price changes are to be considered true or artificial. It may be argued that higher stock-market prices, ceteris paribus, imply a greater importance of both the stock market and stock-brokering services, and thus indeed a greater volume of stock-brokering services. If that argument is valid, it is in favour of alternative (B) for PPIs.

As the PPIs are output indices, the consideration of comparison of “like with like” over time should here be seen from a producer perspective, not a consumer perspective. When ad valorem charges are used, this apparently indicates that the producer effort is somehow related to the transaction value, in terms of e.g. the service provider’s responsibility and risk. Namely, if the producer effort would rather consist in the handling at a flat production-cost per transaction, it would seemingly be both unmotivated and uncompetitive for the service provider to charge ad valorem, and thus it would then hardly be done – but it is. So also with a producer perspective in mind there appears to be good reasons for alternative (B).

But as mentioned the issue is controversial, and the views given here are the author’s.
References

Affärsvarlden (2003), [a Swedish business magazine], http://www.affarsvarlden.se


S. Frost (2001), The construction of price indexes for deposit and loan facilities, Sixth Meeting of the Ottawa Group, in Canberra.


I. Sjögren (2000), Banktjänster: utvidgning av KPI:s täckning, Memorandum for the Swedish CPI Board, meeting no 206 (in Swedish; cf. also the minutes of the meeting).


K. Woolford (2001), Financial Services in the Consumer Price Index, Sixth Meeting of the Ottawa Group, in Canberra.
Annex

Diagram 1: HICP/Sweden, Financial services n.e.c. and Overall index

Diagram 2: HICP/Sweden and AFGX stock index

Note: The AFGX stock index levels followed by the curve are those for the 15th of each month.
Diagram 3: HICP/Sweden (Overall), and alternative (A) simulated

Diagram 4: Inflation rate by HICP/Sweden, and alternative (A) simulated
Session 5 - New products, substitution between products and outlets

Chair: Bert M. Balk, Statistics Netherlands and Erasmus University Rotterdam

Summary of session

The paper by Koskimäki and Ylää-Jarkko reports about two interesting numerical exercises, both using a large set of electronic transaction data (so-called “scanner data”) provided by ACNielsen.

The first exercise replicates to the extent possible the official CPI calculation method on these data, which comprise various commodity groups, and compares the results with officially published figures. Despite differences in price level, simulated and official CPI are close to each other, which is a positive result.

In the second exercise the authors consider the effect of the classifications used to aggregate data for calculating unit value indices. Using two dimensions, the regional dimension (country – province – ACNielsen region – outlet) and the commodity dimension (COICOP5 – COICOP7 – ACNielsen brand – ACNielsen EAN class), the effect of 16 classifications could be tested on the data. The level of aggregation appeared to have a significant impact on the overall index. The results showed that tighter specifications led to higher substitution bias but increasing the size of groupings in the commodity dimension provided higher price index numbers. The last observation could probably be seen as a case of so-called unit value bias.

The paper by Gudnason is essentially a translation of a general purpose paper on the method of the Icelandic CPI. As such it covers many topics. Some noteworthy features are:

- The overall index has the structure of a Lowe price index, comparing prices of the current month (April of year j through March of year j+1) to prices of March of year j, using quantities of year j-2. The index is updated every year, the month of chaining being March.

- Household expenditure data come from shopping receipts provided by households, telephone bills provided by telephone companies, to name just some interesting examples.

- For elementary aggregates a mixture of Jevons and Dutot indexes is used.

The following special topics were considered in this paper:

- A very big inflation during the year 2001 led consumers to move their expenditures to outlets characterized by lower price levels. The move appeared to be so massive that, between the two yearly, general updates of the CPI, an adjustment of the outlet weights was deemed necessary and executed. It led to a downward correction of the CPI by 0.55%.

- The treatment of owner-occupied housing consists in basing the expenditure weight on (an estimate of) user cost and the price relative on the prices of properties sold.
Roman presented a paper written by Makaronidis, dealing with the way in which Eurostat monitors the inclusion of newly significant goods and services in the HICP.

Newly significant goods and services have to be quickly introduced in the sample to keep it up-to-date. They relate to additions to the sample (extension of the coverage of the index) and not to replacements (sample updating with more representative products). It appears to be important, but at the same time very difficult, to clearly distinguish between additions and replacements.

Another issue is the level of appreciation for the “significance” of products: products should be introduced as soon as they are consumed to a significant extent (currently 0.1%), but this threshold should be appreciated at a higher level, the “expenditure group” which gathers products of the same family. A set of criteria is currently under construction, one of which is the notion of purpose. The working of these criteria is discussed with respect to electricity suppliers, e-commerce, and mobile phones.

**Recommendations for statistical agencies**

Although the main purpose of the three papers was to inform the audience about ongoing research and to draw the attention to some important issues in CPI construction, three recommendations could be distilled from the discussion:

1. According to their effect on index calculation, more attention should be paid to classifications used in data aggregation. The aggregation level at which unit values are considered as basic data for CPI construction is in particular an important decision variable. Since some results of the Finnish paper were viewed as counterintuitive, more research is necessary.

2. A big change in household shopping behaviour should (ideally) lead to an immediate adjustment of outlet weights, to prevent bias. Statistical offices should not wait with such an adjustment until the next scheduled general update of the CPI.

3. New products should be represented in the sample as soon as their consumption has reached a significant extent. Operational rules for the conditions of their introduction in the sample have to be further studied.
Segmented markets and CPI elementary classifications

Timo Koskimäki¹ and Mari Ylä-Jarkko² ³

Statistics Finland

Special thanks to Eugen Koev and Antti Suoperä for insightful comments. Professor Yrjö Vartia also provided valuable ideas in the initial stages of the research project.

Abstract: A traditional consumer price index (CPI) is typically constructed using - sometimes very limited - samples of transactions. Also, the standard CPI typically keeps the quality of the products to be compared constant by imposing on the products to be compared rather rigid classificatory rules. These rules determine, which kind of products should be treated as comparable and which not. The availability of more exhaustive scanner data sets opens up an opportunity not only to compile indices utilising more exhaustive data but an opportunity to check the validity of the standard CPI sampling- and replacement practices.

In the paper the performance of traditional CPI sampling, matching and replacement rules are evaluated using an extensive scanner data set as a simulation platform. The structure of the study is as follows:

First, the methodology of a traditional⁴ CPI is replicated in every detail using a representative micro-level scanner data set. The results of the scanner- and traditional price indices are compared.

Then, alternative classification schemes that allow for varying degree of substitution between varieties offered by different manufacturers, product brands, outlets and different regions are constructed and applied to the scanner data. The results are then interpreted in the light of traditional CPI sampling and replacement solutions.

The type of data used does not seem to be of immense importance for the price change comparisons. When the CPI practices were replicated using scanner data and compared with the official CPI, the indices were fairly close to each other.

The choice of classification and replacement rules (whether the products belonging to the same product class are deemed directly comparable or not) do, however, affect the results of the index calculus quite significantly also in the case of rather simple consumer goods.

The problem of quality change has in the research literature mainly been discussed in the context of products where rapid technological change is apparent. On the basis of the results it is apparent that "quality change" issues should be considered also in the case of simple commodities. However, the nature of quality change for simple products is somewhat

¹ Timo Koskimäki wrote the first draft for the abstract, chapters 2 and 3.
² Mari Ylä-Jarkko wrote the first draft for chapter 1.
³ In addition to the authors, Juha Mylläri and Antti Katainen have been active members of the research team.
⁴ In this case the Finnish national Consumer Price index.
different. It is mainly a question of differentiated marketing strategies, i.e. manufacturers attempt to present their own products as distinct as possible. The idea is to try to maintain brand-based segmented micro-markets where substitution between the products should happen within the brand, not between brands. If the segmentation strategy is successful, the manufacturer can adopt quite profitable pricing schemes and plan the obsolescence of their products. The question of how price statisticians should react to this kind of situation in principle is somewhat difficult. However, in practical data collection situation this decision is taken when determining the “sameness” of the products to be priced.

As there do not exist any self-evident statistical rules on how to deal with different types of classification and comparability issues at the most detailed level of the CPI's, measures should be taken to agree on some generally accepted best practices for types of commodities and market situations highlighted in this study

1. The traditional CPI compiled using scanner data

1.1 The data

The data used in this study was provided by ACNielsen Ltd. The "elementary transaction" in the data is sales of a specific product during one week in a single outlet. "Specific product" is identified as a distinct EAN code. Data relating to one elementary transaction are as follows:

- number of packages sold
- package size
- unit of the package (Kilogram, Litre, Ml et.c.)
- total sales during the week
- Product identification code ("EAN")
- brief description of the product
- manufacturer
- ACN's own product class
- "Brand"
- outlet code
- region

For the purpose of index calculations a simple package size adjustment was performed. The elementary price used in all the examples is the "weekly mean price" of one unit sold.

The products covered in this study are typical "fast moving consumer goods" where no rapid technical quality changes is expected to occur. The data covers the following product groups: Butter, Margarine and other vegetable fats, Vegetable oils, Soft drinks, Fruit juices and Detergents.

1.2 The CPI replication approach

In this part of the study, we identify, in the scanner data set, the same products that are being priced in the Finnish national CPI. Indices applying same type of replacement rules as in the current CPI but using the scanner data will be calculated and the results compared with the official CPI results.

The purpose of this part is to examine whether there are differences between indices based on scanner data and the official indices based on standard CPI data collection which stem from
the differences in basic data. To assess this issue is important from the point of view of generalisation of the results of the other analyses.

From the data only the varieties that corresponded to the standard CPI quality descriptions were taken into the compilation of the ACN -index. From every shop the most sold variety in year 1998 that was available in January 1998 was chosen for monitoring. If there were no observations in the outlet in January 1998 the outlet was excluded from the compilation. Only the outlets that stayed in the sample over the three years of monitoring were selected. This resulted to good 250 outlets, which is slightly more than there are in the standard CPI price collection.

If a variety was not available during the monitored month, the price change has been imputed using the mean price change for all the varieties belonging to the same class as proxy information. In the same manner, when a new variety was selected for pricing, the same procedure was used to construct a new base price. In essence, new varieties were chained into the index assuming that the price development has been the same that was observed for varieties where price information for the two periods was available.

The product groups that matched with standard CPI product descriptions are presented in table 1. When aggregating the sub-indices the standard CPI weights were used also for the scanner data.

**Table 1: Sub-indices and their weights**

<table>
<thead>
<tr>
<th>Total Index</th>
<th>COICOP</th>
<th>100,00</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.1.5.1.01</td>
<td>Dairy butter</td>
<td>14,47</td>
</tr>
<tr>
<td>01.1.5.2.01</td>
<td>Butter and vegetable oil mixture</td>
<td>10,72</td>
</tr>
<tr>
<td>01.1.5.2.02</td>
<td>Cooking margarine</td>
<td>3,25</td>
</tr>
<tr>
<td>01.1.5.2.03</td>
<td>Soft margarine</td>
<td>14,81</td>
</tr>
<tr>
<td>01.1.5.2.04</td>
<td>Low fat margarine</td>
<td>1,17</td>
</tr>
<tr>
<td>01.1.5.4.01</td>
<td>Vegetable oil</td>
<td>3,22</td>
</tr>
<tr>
<td>01.2.2.2.01</td>
<td>Vegetable extract drinks cola</td>
<td>16,18</td>
</tr>
<tr>
<td>01.2.2.3.01</td>
<td>Mixed fruit cordial</td>
<td>1,08</td>
</tr>
<tr>
<td>05.6.1.1.02</td>
<td>Synthetic detergent</td>
<td>18,34</td>
</tr>
<tr>
<td>05.6.1.1.03</td>
<td>Dish washing liquid</td>
<td>5,24</td>
</tr>
<tr>
<td>05.6.1.1.04</td>
<td>General purpose cleanser</td>
<td>11,53</td>
</tr>
</tbody>
</table>

Despite the attempt to replicate the standard CPI in every detail, some important differences do remain. Probably the most important difference is the fact that ACN prices are weekly average prices rather than daily on the spot -prices collected for the CPI. This means, among other things, that (nation-wide) sales campaigns, when the variety on sale is purchased in large quantities, clearly show up in the ACN price development. The feature is amplified because of the fact that CPI data is typically not collected during weekends which is the typical time for campaigns. The other factor causing differences is the slightly different treatment of loyal customer sales prices and other targeted campaigns. These types of sales pricing schemes are not covered by the traditional CPI but are covered in ACN data.
1.3 Results (I)

In picture 1 the indices for group "General purpose cleanser" are presented. The ACN index is clearly more volatile. This volatility is most probably due to the differences in the price concept in the two data sources.

Picture 1: Indices of general purpose cleanser. The current CPI methodology, CPI-data and ACN-data.

However, for other product groups the differences are relatively small. Picture 2 shows price indices for "Dish Washing liquid". Here the differences are marginal.

Picture 2: Indices of dish washing liquids. The current CPI methodology, CPI-data and ACN-data.
When aggregating all compiled sub-indices into a total indices, the results are reasonably close to each other. Total indices are presented in picture 3. After three years of surveillance the indices differ only by 0.5 percentage points. The largest difference is 1.41 percentage points. Generally, we conclude that both of the data-sets give the same result when same type of sampling and replacement practices are applied.

**Picture 3: Total indices 1998:01 = 100 using the current CPI methodology.**

As is apparent from above, the raw mean prices calculated from ACN data do differ from mean prices calculated from CPI data. Annex 3 gives some examples on differences. The general trend is somewhat unclear, in some of the cases studied, the mean prices are reasonably close to each other, whereas for certain product groups the difference is huge. Some of the reasons – campaigns (impacts of sales volumes) and also loyal customer rebates etc. have already been pointed out. In addition, in this study it was not possible to match CPI data with ACN-data at the outlet level. Hence, part of the observed difference may stem from differences in the outlet sample.

It is however quite clear that (raw) mean prices collected using the standard CPI methodology do, in general, differ from each other although the indices compiled from the same data are not biased.

As a consequence, mean prices from the traditional CPI data collection should not be compared directly with the mean prices from cash register systems. Such a comparison will most probably be seriously biased.

### 2. CPI elementary classifications

In chapter 1 we compiled the CPI using the scanner data but applying the traditional CPI sampling and matching rules. As we have access to total data, there is in principle no need to
use sampling. Instead, it is natural to construct a form of index compilation strategy that utilises a maximum amount of the data available.

This part of the study presents two possible strategies that may be applied when using scanner-type data in CPI compilation. We denote the two strategies as the **matching approach** and as the **classification approach**. Both of the strategies will be illuminated by empirical analysis of a geographically representative scanner data set covering some 350 outlets and 8 product categories described in chapter 1. Advantages and disadvantages of the approaches will be discussed as well as the conclusions from the point of view of the traditional CPI practices.

### 2.1 Matching, classification and unit value indices

The most obvious strategy to deal with the excessively detailed scanner data is to treat the data in a similar manner as the data is treated in current CPI compilation, i.e. one-to-one matching of the observations at the most detailed level, within outlets, across time. Indeed, most of the current studies utilising scanner data seem to take one-to-one matching as a self-evident starting point without reflecting on other possibilities.

The number of products to be included in the "scanner-CPI" would be essentially larger than in the traditional CPI. However, the products would still be matched in a very detailed level and the non-matches would be omitted from the compilation. In practice, this means matching of the products on the basis of the most detailed product codes available (in Europe the so-called EAN-codes) on the level of single outlets.

From a more general perspective the matched models approach can be characterised as an extremely detailed classification system that has been formed by combining product- and outlet typologies. On the most disaggregated level of index compilation the product classification to be used is often very narrow. Products at the elementary aggregate level are typically classified by package size, manufacturer of the product, flavour or colour, specific brand name etc.

Once we have realised that the traditional one-to one matching is a form of classification we may also consider alternative ways of classifying the data. In essence, we may choose - at our will - the level the detail to be used in the formation of elementary aggregates for the index compilation. The choice of the level of classification should, of course, be justified by some form of rational reasoning (see, e.g. Durkheim and Mauss 1963/1903).

Generally, the term used for this type of approach is "unit value index". The use of unit value indices as elementary aggregates may be perceived as somewhat peculiar (Balk 1998). Diewert (1995, ref Balk 1998), however, seem to give some support for the general idea. Unit values within homogeneous sub-groups are also the standard procedure proposed by some less recent textbooks (see e.g. Allen, 1975). As will be demonstrated, there always exists a multitude of alternatives to form "unit value indices". To define the borderline between "unit value index" and "price index" is always to certain extent an arbitrary choice.

"Price indexes, almost universally, have followed one fundamental methodological principle: The quality of the product sets to be compared is held constant by following the matched pairs strategy. The price index compiling agency chooses a sample of retail outlets or sellers and a sample of products. It collects an initial period, or base period, price for each of the products selected. It then collects at some later date the price for exactly the same product, from the
same seller, that was selected in the initial period. The price index is computed by matching, observation by observation, the price at the later period with the initial price (Tripplet, 2000).

This kind of classification strategy, admittedly, creates very homogenous sets of observations. Triplett, again, summarises the advantages of one-to-one matching quite neatly:

"The great advantages of this matching methodology are sometimes not explicitly stated, and other times not fully appreciated. The "matched model" methodology holds constant many price determining factors that are usually not directly observable. Examples are characteristics of the retailer, such as customer service, reputation of the manufacturer etc. Matching the price quotes model by model (and outlet by outlet) is not just a methodology for holding quality change constant in the items selected for pricing. It is also a methodology for holding constant non-observable aspects of the transaction that might bias the measure of price change.

The disadvantages of the matching approach are quite obvious although seldom explicitly stated. The most general formulation of the drawbacks is that the matched models approach always causes quite a considerable loss of information. This loss of information stems from the simple fact that new and disappearing varieties can not be taken into account in the index calculation.

Quite a number of contemporary problems in the field of price indices are derivatives of this general feature of the matched models approach. The need for specific "quality adjustment" procedures in cases where a variety that has been followed in the CPI disappears from the markets and has to be replaced by another variety is one class of problems. The other class of problems stems from appearance of "new products". As there is no natural pair for comparison, the new product is generally omitted from the calculation.

A third class of problems stem from the practicalities of the matched pairs -approach. It is often the case that price collectors are advised to follow the initially selected variety until it entirely disappears from the markets. This, in turn, leads rapidly to a situation where the CPI sample is no longer representative, i.e. gives a distorted picture of the markets (see Koskimäki and Vartia 2001, Silver 2001).

The main drawback of the classification strategy is often denoted as "the unit value bias". In essence this term refers to a situation where the elementary cells - or elementary classes created using pre-specified classification rules - are not homogenous enough. The idea of comparing "like with the like" may thus be violated and differences in quality might appear as differences in price. The important question in this study is, however, how should we interpret the notion "like with like".

2.2 The classification approach

For the purpose of this study, two weekly data sets - last week of September 1998 and last week of September 2000 were extracted from the material described in chapter 1.

In addition, products were classified in the spirit of the standard classification of individual consumption by purpose (COICOP). The version of COICOP classification used is presented in annex 1. Also brand- and EAN- classifications provided by ACNielsen were used. The number of classes in each level of the product- and geographic classifications used in the index calculations are presented in table 2 below:
Table 2: Classifications used in the study
(Number of distinct product codes (EAN) refers to 1998 data)

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>Number of levels</th>
<th>Product dimension</th>
<th>Number of levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>1</td>
<td>COICOP 5-digit</td>
<td>6</td>
</tr>
<tr>
<td>Province</td>
<td>4</td>
<td>COICOP 7-digit</td>
<td>26</td>
</tr>
<tr>
<td>ACN region</td>
<td>15</td>
<td>ACNielsen brand</td>
<td>266</td>
</tr>
<tr>
<td>Outlet</td>
<td>338</td>
<td>ACNielsen EAN</td>
<td>1028</td>
</tr>
</tbody>
</table>

The regional and product dimensions were then combined to form a typology - or a matrix - to be used in the compilation of indices. For each cell in the matrix an index with differing weighting structure were calculated. The typology is shown in table 3.

Table 3: Typology of different classification strategies

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>COICOP 5-digit</th>
<th>COICOP 7-digit</th>
<th>ACNielsen Brand</th>
<th>ACNielsen EAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>1.</td>
<td>2.</td>
<td>3.</td>
<td>4.</td>
</tr>
<tr>
<td>Province</td>
<td>5.</td>
<td>6.</td>
<td>7.</td>
<td>8.</td>
</tr>
<tr>
<td>ACN region</td>
<td>9.</td>
<td>10.</td>
<td>11.</td>
<td>12.</td>
</tr>
<tr>
<td>Outlet</td>
<td>13.</td>
<td>14.</td>
<td>15.</td>
<td>16.</td>
</tr>
</tbody>
</table>

The idea of the approach is as follows: The intersection of row and column attributes describes the elementary aggregate level of the index to be calculated. Hence, in the index to be calculated for cell number one we only fix weights for the six COICOP 5-digit groups. In cell 5 we fix the weights also according to the province (4 classes in regional dimension and six classes product dimension, i.e. 24 classes.).

In the lower right corner (cell 16) all available data has been fixed to form an entirely fixed weighting structure (338 outlets and 1028 distinct products). Theoretically, if all outlets would sell all products - and stayed in the markets over the entire period of the study - this would mean 347 464 fixed cells.

Turning back to the two basic approaches, the rightmost cells in the matrix (4, 8, 12, 16) are variations of the matching approach. The rest of the cells are examples of different classification strategies.

From the point of view of traditional CPI practices, the design of the cell structure can be interpreted as a set of distinct replacement rules. In the cell 1 we would consider - in a practical price collection situation - any product within the Coicop 5-digit class as directly comparable (say, vegetable oil is vegetable oil) whereas in cell 16 we would require also the
type of oil, exact package size, manufacturer and product line (brand) to be the same before we consider the products to be the same product.

The elementary aggregates below the given stratification levels have been calculated as weighted arithmetic means. The aim is to produce a meaningful mean price taking into account the fact that the market shares of the products under study vary from 0.5 per mill to some 10 ten per cent within a given product category.

From the point of view of traditional CPI practices, the way to construct the elementary prices within aggregates can be interpreted as a re-sampling rule that would advise the price collector performing outlet-level sampling always to select the most sold variety of the product group to be priced.

In essence, elementary aggregates used here are analogous to the price concept inherent in the scanner data on the outlet level. The weights used in the construction of elementary aggregates are allowed to vary between the two periods concerned. This is consistent with the general idea that the elementary aggregates under study are considered to be homogenous, at least from a consumer's point of view.

2.3 Results (II)

The results of the exercise are shown in tables 4 to 7. To summarise:

- Increasing the level of detail (of the elementary classification) in the product dimension tends to lower the observed price increase. Using 6-class structure in the product dimension yields a price increase of 7.9 percent whereas fixing the weights on EAN-level indicates only 2.3 per cent price increase.

- Increasing the level of detail in regional dimension, especially if the calculus is fixed at the outlet level, tends to increase the observed price increase. The phenomenon, however, disappears if product dimension is tightly fixed.

- Classification which keeps the producer - but not the EAN - fixed (ACNielsen Brand classification, 266 classes) tends to give higher price increase when compared to a reasonable consumer-oriented classification (COICOP 7-digit, 26 classes).

- Tight classification - both in product- and outlet- dimensions - tends to increase upper level substitution bias (measured as difference between Laspeyres and Fisher indices).

- The loss of information increases rapidly when classifications get more detailed. Keeping the outlet sample fixed between the two periods decreases the coverage by 10 per cent. The most detailed product classification enables only 80 per cent of the transactions to be included in the comparison. The joint effect - keeping both the products and outlets extremely fixed - excludes almost 40 per cent of the data from the index calculation.
### Table 4: Laspeyres price indices September 1998 - September 2000 (sales during one week)

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>COICOP 5-digit</th>
<th>COICOP 7-digit</th>
<th>ACNielsen Brand</th>
<th>ACNielsen EAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>107,9</td>
<td>103,1</td>
<td>104,6</td>
<td>102,3</td>
</tr>
<tr>
<td>Province</td>
<td>107,8</td>
<td>103,1</td>
<td>104,8</td>
<td>102,3</td>
</tr>
<tr>
<td>ACN region</td>
<td>107,8</td>
<td>103,1</td>
<td>105,1</td>
<td>102,5</td>
</tr>
<tr>
<td>Outlet</td>
<td>108,6</td>
<td>104,0</td>
<td>106,0</td>
<td>102,8</td>
</tr>
</tbody>
</table>

### Table 5: Fisher price indices September 1998 - September 2000 (sales during one week)

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>COICOP 5-digit</th>
<th>COICOP 7-digit</th>
<th>ACNielsen Brand</th>
<th>ACNielsen EAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>108,0</td>
<td>103,2</td>
<td>104,8</td>
<td>101,5</td>
</tr>
<tr>
<td>Province</td>
<td>107,9</td>
<td>103,1</td>
<td>104,8</td>
<td>101,4</td>
</tr>
<tr>
<td>ACN region</td>
<td>107,9</td>
<td>103,0</td>
<td>104,7</td>
<td>101,4</td>
</tr>
<tr>
<td>Outlet</td>
<td>108,9</td>
<td>103,4</td>
<td>104,9</td>
<td>101,1</td>
</tr>
</tbody>
</table>

### Table 6: Difference Laspeyres – Fisher

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>COICOP 5-digit</th>
<th>COICOP 7-digit</th>
<th>ACNielsen Brand</th>
<th>ACNielsen EAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>-0,1</td>
<td>-0,1</td>
<td>-0,2</td>
<td>0,8</td>
</tr>
<tr>
<td>Province</td>
<td>-0,1</td>
<td>0,0</td>
<td>0,0</td>
<td>0,9</td>
</tr>
<tr>
<td>ACN region</td>
<td>-0,1</td>
<td>0,1</td>
<td>0,4</td>
<td>1,1</td>
</tr>
<tr>
<td>Outlet</td>
<td>-0,3</td>
<td>0,6</td>
<td>1,1</td>
<td>1,7</td>
</tr>
</tbody>
</table>

### Table 7: Turnover covered by each calculation (Whole country, COICOP 5-digit = 100)

<table>
<thead>
<tr>
<th>Regional dimension</th>
<th>COICOP 5-digit</th>
<th>COICOP 7-digit</th>
<th>ACNielsen Brand</th>
<th>ACNielsen EAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole country</td>
<td>100,0</td>
<td>100,0</td>
<td>98,2</td>
<td>80,1</td>
</tr>
<tr>
<td>Province</td>
<td>100,0</td>
<td>100,0</td>
<td>97,5</td>
<td>77,4</td>
</tr>
<tr>
<td>ACN region</td>
<td>100,0</td>
<td>100,0</td>
<td>96,9</td>
<td>75,5</td>
</tr>
<tr>
<td>Outlet</td>
<td>90,4</td>
<td>90,4</td>
<td>84,6</td>
<td>61,7</td>
</tr>
</tbody>
</table>

(period-specific weights)
3. Conclusions

On the basis of the results of the CPI replication presented in chapter 1, it is evident, that the type of data used for CPI's is not of immense importance. However, the way data is compiled during the index calculus is clearly of importance.

It is relatively easy to figure out behavioural explanations for the results presented in chapter 2. If we keep "the producer" (ACN brand-classification above) fixed, we seem to exclude from the index price effects caused by new producers or brands entering the markets.

More important, when keeping the brand classification constant, we technically do not allow consumers' substitution between different brands in connection with brand-specific price changes or changes in consumption pattern. I.e. when, within a brand, a new variety of a product is introduced at a higher price we in this design assume that consumers only substitute between the older versions of the same brand (if these still exist in the markets) but do not go over to less expensive brands.

According to the data, however, people tend to substitute between brands as well. This is evident from the comparison of substitution effects (biases) in table 6: When using ACN brand-classification (substitution between brands not allowed), the substitution bias, measured as a difference between Laspeyres and Fisher price indices is 1,1 per cent whereas when applying Coicop 7-digit classification (where substitution between brands is allowed), the substitution bias (or effect) is only 0,6 per cent.

If we keep our sample of outlets constant, we loose the price decrease brought up by new outlets competing on the markets. Also, we loose the effect of consumers changing to less expensive outlets already in the markets.

Similarly, if we keep our product set extremely tightly defined, we loose a considerable part of the transactions in the markets and hence may create a biased index.

The above mentioned substitution effects have traditionally been discussed in the context of different index formulas (see, for example, de Haan 2001). The exercise exposed here should show that the issue can - and should - be treated also as a problem of product classification, or, in traditional CPI context, as a problem of matching, replacement and quality adjustment rules.

The problem of quality change has in the CPI research literature mainly been discussed in the context of products where rapid technological change is apparent. On the basis of the results of this study, it is apparent that "quality change" issues should be considered also in the case of simple commodities.

The nature of quality change for simple products is, however, somewhat different. It is mainly a question of differentiated marketing strategies, i.e. manufacturers attempt to present their own products as distinct as possible. The idea is to try to maintain brand-based segmented micro-markets where substitution between the products should happen within the brand, not between brands. If the segmentation strategy is successful, the manufacturer can adopt quite profitable pricing schemes and plan the obsolescence of their products. The question of how price statisticians should react to this kind of situation in principle is somewhat difficult.
However, in practical data collection situation this decision is taken when determining the "sameness" of the products to be priced.

As there do not exist any self-evident statistical rules on how to deal with different types of classification and comparability issues at the most detailed level of the CPI's, measures should be taken to agree on some generally accepted best practices for types of commodities and market situations highlighted in this study.

References


Durkheim, E, and Mauss, M: Primitive Classification. The University of Chicago Press, Chicago 1963 (Original: De Quelques Formes Primitives de Classification, Annee Sociologique, 1901-02 (1903)).


Annex 1: COICOP classification used in the study

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.1.5.1</td>
<td>Butter</td>
</tr>
<tr>
<td>01.1.5.1.01</td>
<td>Dairy butter</td>
</tr>
<tr>
<td>01.1.5.1.02</td>
<td>Other butter</td>
</tr>
<tr>
<td>01.1.5.2</td>
<td>Margarine and other vegetable fats</td>
</tr>
<tr>
<td>01.1.5.2.01</td>
<td>Butter and vegetable oil mixture</td>
</tr>
<tr>
<td>01.1.5.2.02</td>
<td>Cooking margarine</td>
</tr>
<tr>
<td>01.1.5.2.03</td>
<td>Soft margarine</td>
</tr>
<tr>
<td>01.1.5.2.04</td>
<td>low fat margarine</td>
</tr>
<tr>
<td>01.1.5.2.05</td>
<td>Other veg. fats</td>
</tr>
<tr>
<td>01.1.5.4</td>
<td>Vegetable oils</td>
</tr>
<tr>
<td>01.1.5.4.01</td>
<td>Rapeseed oil</td>
</tr>
<tr>
<td>01.1.5.4.02</td>
<td>Sunflower oil</td>
</tr>
<tr>
<td>01.1.5.4.03</td>
<td>Olive oil</td>
</tr>
<tr>
<td>01.1.5.4.04</td>
<td>Other Oil</td>
</tr>
<tr>
<td>01.2.2.2</td>
<td>Soft drinks</td>
</tr>
<tr>
<td>01.2.2.2.01</td>
<td>Veg. extract drinks (Coke)</td>
</tr>
<tr>
<td>01.2.2.2.02</td>
<td>Soda orange</td>
</tr>
<tr>
<td>01.2.2.2.03</td>
<td>Energy drinks</td>
</tr>
<tr>
<td>01.2.2.2.04</td>
<td>Soda soft drinks, other than orange</td>
</tr>
<tr>
<td>01.2.2.2.05</td>
<td>Other soft drinks</td>
</tr>
<tr>
<td>01.2.2.3</td>
<td>Fruit juices</td>
</tr>
<tr>
<td>01.2.2.3.01</td>
<td>Mixed fruit cordial</td>
</tr>
<tr>
<td>01.2.2.3.02</td>
<td>Orange juice, 100 per cent fruit</td>
</tr>
<tr>
<td>01.2.2.3.03</td>
<td>Other cordials</td>
</tr>
<tr>
<td>01.2.2.3.04</td>
<td>Other juice, 100 per cent fruit</td>
</tr>
<tr>
<td>01.2.2.3.05</td>
<td>Juice, less than 100 per cent fruit</td>
</tr>
<tr>
<td>01.2.2.3.06</td>
<td>Other juices and cordials</td>
</tr>
<tr>
<td>05.6.1</td>
<td>Non-durable household goods</td>
</tr>
<tr>
<td>05.6.1.1</td>
<td>Detergents</td>
</tr>
<tr>
<td>05.6.1.1.01</td>
<td>Dishwasher detergent</td>
</tr>
<tr>
<td>05.6.1.1.02</td>
<td>Synthetic detergent (Clothing)</td>
</tr>
<tr>
<td>05.6.1.1.03</td>
<td>Dish washing liquid</td>
</tr>
<tr>
<td>05.6.1.1.04</td>
<td>General purpose cleanser</td>
</tr>
<tr>
<td>05.6.1.1.05</td>
<td>Other detergents</td>
</tr>
</tbody>
</table>
Annex 2: Example of the brand-classification – detergents

AIRWICK
AJAX
ANDY
ASPI
BECKMANN
BEMINA
CHEVY
CLEANI
DEONET
DETEX
DOMESTOS
EKO
EPEX
EUROSHOPPER
FINN MAID
GREEN 2000
HARDOL
HAVI
HIT Klaro
JOHNSON
KIIITO
Klorin
KLORITE
LASER
METSANRAIKAS
MINIRISK
MONSUN
MR MUSCLE
MR. PROPER
MT-EXIMA
PINJA
PIRKKA
REMOP
SANILAV
SIISTO
SINETTI
SMART
SPAR
SUPER
SUPI
TOILET DUCK
TOLU
WC DUCK
WC-FLOWER
VIM
MUU TUOTEMERKKI
Annex 3: Price levels in Scanner-(ACN) and CPI – data sets

Samples and varieties in different data sets have not been matched, hence the differences may reflect, in addition to differences in measurement, also differences in samples.

<table>
<thead>
<tr>
<th>Orange Juice</th>
<th></th>
<th>Coke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month/year</td>
<td>ACN</td>
<td>CPI</td>
</tr>
<tr>
<td>01/95</td>
<td>4.88</td>
<td>5.09</td>
</tr>
<tr>
<td>01/96</td>
<td>4.99</td>
<td>5.20</td>
</tr>
<tr>
<td>01/97</td>
<td>5.63</td>
<td>5.41</td>
</tr>
<tr>
<td>01/98</td>
<td>5.08</td>
<td>5.48</td>
</tr>
<tr>
<td>02/98</td>
<td>5.40</td>
<td>5.46</td>
</tr>
<tr>
<td>03/98</td>
<td>5.24</td>
<td>5.45</td>
</tr>
<tr>
<td>04/98</td>
<td>5.18</td>
<td>5.44</td>
</tr>
<tr>
<td>05/98</td>
<td>5.31</td>
<td>5.46</td>
</tr>
<tr>
<td>06/98</td>
<td>5.28</td>
<td>5.46</td>
</tr>
<tr>
<td>07/98</td>
<td>5.41</td>
<td>5.44</td>
</tr>
<tr>
<td>08/98</td>
<td>5.38</td>
<td>5.39</td>
</tr>
<tr>
<td>09/98</td>
<td>5.65</td>
<td>5.47</td>
</tr>
<tr>
<td>10/98</td>
<td>5.74</td>
<td>5.48</td>
</tr>
<tr>
<td>11/98</td>
<td>5.83</td>
<td>5.57</td>
</tr>
<tr>
<td>12/98</td>
<td>5.87</td>
<td>5.56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fruit cordial</th>
<th></th>
<th>Mineral water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month/year</td>
<td>ACN</td>
<td>CPI</td>
</tr>
<tr>
<td>01/95</td>
<td>10.35</td>
<td>7.02</td>
</tr>
<tr>
<td>01/96</td>
<td>12.94</td>
<td>8.32</td>
</tr>
<tr>
<td>01/97</td>
<td>16.29</td>
<td>8.36</td>
</tr>
<tr>
<td>01/98</td>
<td>15.45</td>
<td>8.38</td>
</tr>
<tr>
<td>02/98</td>
<td>15.65</td>
<td>8.48</td>
</tr>
<tr>
<td>03/98</td>
<td>15.41</td>
<td>8.32</td>
</tr>
<tr>
<td>04/98</td>
<td>15.99</td>
<td>8.34</td>
</tr>
<tr>
<td>05/98</td>
<td>15.04</td>
<td>8.32</td>
</tr>
<tr>
<td>06/98</td>
<td>14.36</td>
<td>8.26</td>
</tr>
<tr>
<td>07/98</td>
<td>14.11</td>
<td>8.15</td>
</tr>
<tr>
<td>08/98</td>
<td>16.24</td>
<td>8.04</td>
</tr>
<tr>
<td>09/98</td>
<td>15.53</td>
<td>8.40</td>
</tr>
<tr>
<td>10/98</td>
<td>14.81</td>
<td>8.36</td>
</tr>
<tr>
<td>11/98</td>
<td>15.67</td>
<td>8.28</td>
</tr>
<tr>
<td>12/98</td>
<td>15.96</td>
<td>8.24</td>
</tr>
<tr>
<td>Country</td>
<td>Period</td>
<td>Index</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>United States</td>
<td>2023</td>
<td>100</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2023</td>
<td>95</td>
</tr>
<tr>
<td>Germany</td>
<td>2023</td>
<td>90</td>
</tr>
<tr>
<td>France</td>
<td>2023</td>
<td>85</td>
</tr>
<tr>
<td>Japan</td>
<td>2023</td>
<td>80</td>
</tr>
</tbody>
</table>

This table represents the price indices for various countries in the year 2023.
How do we measure inflation? Some measurement problems

Rósmundur Guðnason

Statistics Iceland

I would like to thank: Guðrún R. Jónsdóttir, Hallgrímur Snorrason, Jón Scheving Thorsteinsson, Jökull Mar Pétursson, Rut Jónsdóttir, Örn Ingvarson for their help in the preparation of this paper. I would also like to thank the participants in the Ottawa Group meeting in Paris for helpful comments.

Abstract: The paper describes methods used in the calculation of the Icelandic CPI and resembles a conditional COLI. Substitution is counted for by using the geometric mean and by allowing substitution between stores. Methods used for correction of shopping substitution bias (outlet substitution bias) are described. This bias increased when inflation rose suddenly in the year 2001 and was corrected by quality adjustment, mainly based on evaluation of different selection of goods. Store weights were adopted in April 2002 which makes such corrections easier. The main sources are cash receipts collected in the household budget survey. Additional sources include information gathered from the main retail groups. Corrections of the CPI were made in December 2001, March 2002 and May 2003, leading to 0.55 per cent lowering of the index. Owner occupied housing is calculated as user cost using depreciation and real interest rate. Using the real interest rate is specific for the Icelandic method.

Keywords: Consumer price index, cost of living index, household budget surveys, shopping substitution bias, outlet substitution bias, quality adjustment, owner occupied housing, user cost.

JEL classification: C43, C81, D11, E31.

1. Introduction

It is difficult to define inflation and no unique definition available. It is often said that it is "permanent general rise in prices"1 or as "continuously rising prices or equivalently of a continuously falling value of money"2. The first definition refers to general conditions the other refers mostly to monetary aspects of rising prices. The question then rises how to measure this permanent increase or change in prices. The tool for such measurement are price indices that measure price changes in distinct periods. It is most common to use consumer price indices for that task but they cover households consumption expenditures. They are based on well known practices recognised both domestically and internationally, the dataset is

1 Jón Sigurðsson (1974) 30
based on regular extensive price collection, often bound by law. The consumer price index measure the price change of private consumption. If the aim is to measure the total price changes in the society there is a need to measure, in addition to private consumption, both public consumption and investment. The problem with such a total measurement is that the indices needed are often either not available or published timely enough to be of use. It is for that reason that the consumer price index is used as the reference for inflation.

The tasks in the measurement of inflation are to find the changes in prices that reflect the inflation development in best way possible. Most of the theoretical literature on indices is about witch prices should be measured and what methods are most reliable for that task. To measure prices is a complex and difficult exercise. Arnjóður Ólafsson describes the problem in the following way in the year 1880 as follows and that description is still valid. "The word price is not a long word but it has not brought any luck to economists or other scholars. The right meaning is truly difficult to find, it is difficult to find the way to the origins of prices and their causes, their increase or decrease. No thing is as volatile and changeable as prices. The extent and the immense quantity of transactions mean that the oversight and total information is not available without intensive effort. Estimates are therefore necessary and here indices enter the stage. The basic task in the calculation of indices is how to compile this extensive information so it reflects changes in prices in the right way. If a complete information on both prices and quantities would be available for all goods and services the index problem would be easily solved. The main problem with price indices is how to calculate them in the most sensible way taking into consideration best available information at each time.

In the last ten years there have been a considerable development in the methodology of consumer price indices especially in the field of elementary indices. That work has mainly been conducted in an international working group, the Ottawa group. The working group was originally started in Ottawa in Canada in the year 1994 with the aim to be a forum for research and discussion about price measurement. The group has held seven meetings (2003) and discussed the theoretical side of this issue but has mainly dwelled with practical research in the field of consumer price indices. This work has to a great extent been used in the drafting of two international manuals on indices. One of them is about consumer price indices and is written under the auspices of ILO. The other one is about producer price indices worked under IMF direction.

In the last years there has also been a discussion internationally about bias in consumer price indices especially after the publication of the Boskin report in 1996 that discussed the subject. The discussion that followed and the fact that there has been a considerable lowering of inflation internationally have further pointed to the issues of price measurement, especially if inflation is upward biased.

ILO has sponsored the drafting of resolutions standards about consumer price indices since 1925. Their latest resolution on the issue is from 1987. The main objective of the resolutions is to make it easier for states to build good consumer price indices. Preparations are ongoing for the ICLS (International Conference of Labour Statisticians) that will be held at the end of the year 2003 where a new resolution will be adopted. The resolution will take into account changes in methodology and priorities in the production of consumer price indices.

---

3 Arnjóður Ólafsson (1880) 106.
When preparing for the Euro as a common currency in Europe high priority was put on inflation targets and in the Maastrict treaty there were clauses about them. This clause led to the intensive work on harmonisation that has been conducted in Europe in last years under the leadership of Eurostat. Since the beginning of 1997 the harmonised index of consumer prices has been calculated monthly. It is based on common methods that are stipulated by law\(^4\).

This paper deals with the sources of and methods used in the calculation of the Icelandic CPI and how it resembles a conditional cost of living index. In chapters 2 and 3 there is discussion about methods of calculations used in the index and new data sources deriving from the continuous household budget survey conducted by Statistics Iceland. Chapters 4, 5 and 6 discuss the fields where the calculation of the consumer price index resembles conditional cost of living index. Chapter 4 describes the calculation of the index, chapter 5 correction of a shopping bias and chapter 6 describe the calculation of housing in the index.

2. Methods for calculations

2.1 Fixed base indices and Cost of living indices

Theoretically there are two main methods used in the calculation of indices. On one hand the expenditures are kept fixed and fixed base indices calculated. Fixed base indices are called pure price indices if they measure only price change when the quantities are kept fixed. If the weights refer to older times the indices calculated are called Laspeyres and if the reference is to present time they are referred to as Paasche indices.

\[
(2.1) \text{Laspeyres: } P_L = \frac{\sum_{i=1}^{n} p_i^t q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0}
\]

\[
(2.2) \text{Paasche: } P_P = \frac{\sum_{i=1}^{n} p_i^t q_i^t}{\sum_{i=1}^{n} p_i^0 q_i^t}
\]

Where \( i = \text{goods, } 1, \ldots, n \), \( p_i = \text{price of good } i \), \( Q_i = \text{quantity of good } i \).

Fixed base indices show changes in expenditures between two periods when fixed quantity of goods and services are bought. They can be of different nature and the prices or weights used in their making are diverse. When they are calculated it is not supposed that consumers change their selection of goods when the price relatives change, i.e. there is no substitution. Usually fixed base indices used are of Laspeyres type as information about old weights is usually available and it is sufficient then to collect prices for the index calculation. In the case of Paasche indices there is in addition to the price collection need to continuously gather information about new weights which is both complicated and cumbersome. Theoretically though there are no arguments to select one rather than the other.

\(^4\) Eurostat (2001)
The other main method, cost of living (COLI)\textsuperscript{5}, measures changes in welfare in a wide meaning. It has been defined as "the ratio of the minimum expenditures required to attain particular indifference curve under two price regimes"\textsuperscript{6} All parts of welfare can not be measured with price indices, such as influence of climate, nature catastrophes, terrorism and plagues. Conditional cost of living indices cover domain where price indices can be used. It is assumed that consumers keep their quality of life fixed and the cost of living index measure changes in cost connected with that. Economically it is connected to theories about true cost of living\textsuperscript{7} meaning that consumers maximise their utility at lowest cost. According to this individuals maximise their utility by changing their consumption if prices go up to buy cheaper goods or goods with less rise in prices.

2.2 Superlative indices

Methods for measuring utilities for individuals are not available and that makes the measurement of cost of living indices difficult. It was therefore a great discovery that symmetric indices, so called superlative indices\textsuperscript{8}, reflect in an adequate manner a true cost of living with given assumptions about the functional form of the utility function\textsuperscript{9}. It is therefore possible to calculate a cost of living index without directly measuring other assumptions such as indifference curves.

Cost of living index and pure price index can be defined as the ratio of expenditures in two periods. In the fixed base index the quantity is kept fixed but it can change in a cost of living index. The calculation of cost of living indices is connected to consumption theories. In reality the method of calculations is not so different as it might seem at first sight. "In practice, the real problem for all price indexes, whether they are intended to be measure of inflation or changes in the cost of living, is to get the most appropriate or relevant weights. It should be noted that even when the objective is to measure the changes in the cost of living, the indexes actually calculated in practice are always pure price indexes of one kind or another. When the weights are 'right', it matters little whether the index is intended to be an inflation or a cost of living index."\textsuperscript{10} Superlative indices are symmetric and two periods always taken into consideration, old and new. The problem is that information about new weights is really available until after some time and it is therefore difficult to calculate them in timely fashion. They differ from fixed base indices in that they either use old weights (Laspeyres) or new weights (Paasche). Bias in the consumer price index is measured as the difference between the index results compared to the result from the calculation of a superlative index. Laspeyres index is biased upwards and the Paasche index downwards. The best known of the superlative indices is the ideal index most often related to the American economist Irving Fisher\textsuperscript{11} but that index is the geometric mean of Laspeyres and Paasche indices (2.9). Its popularity can be traced to the axiomatic (test) method that originates from Fisher\textsuperscript{12}, but according it the ideal index performs best. The only superlative index that

\textsuperscript{5}The concept of cost of living is an economic concept. In Iceland the concept is very often mixed with the question about the level of living cost for different types of households. The COLI does not answer that question and it is not found in the calculation of the CPI.

\textsuperscript{6}Pollak (1989) 6

\textsuperscript{7}Konus (1924).

\textsuperscript{8}The concept was first used to classify indices Fisher (1922) 247.

\textsuperscript{9}Diewert (1976)

\textsuperscript{10}Hill (1999b) 10

\textsuperscript{11}Fisher 1922.

\textsuperscript{12}It was Walsh that was the first to research the test method systematically. Diewert (1993) 39.
reconciles with pure price indices is named after Walsh\textsuperscript{13} (2.10), where the weights are the geometric mean of the quantities and the two periods.

2.3 The axiomatic method

Different methods for calculations of indices show different results. The problem with the selection of a method that is best suited for the calculation is connected with that fact. Therefore it is important to have a tool to scrutinise the different technical properties of the index and to observe which mathematical criteria they fulfil. A method for this is the axiomatic or test method, most often connected with Fisher. The conditions or criteria that indices have to pass to be considered as good indices are put forward. Different indices are tested by these criteria and the results indicate which conditions hold. The results can be used for the selection of method for calculations i.e. if you only look at the technical properties of indices. Examples of such tests are\textsuperscript{14}:

1. Positivity, indices (quantity- and price vectors) must show positive results.

2. Identity test, if prices of all goods in two periods are like the index should be unchanged.

3. Proportionality in current prices, if all prices in one period are multiplied by the same constant the new index should be the old index multiplied with the new constant.

4. Invariance to changes in the units of measurement, the index does not change even if the units change.

5. Time reversal test, if data from two periods are crossed then the result should be the inverse of the original index.

6. Quantity reversal test, if quantities between two periods are crossed then the index should be unchanged.

7. Mean value test for prices, the index result falls in between the highest and lowest price relatives.

8. Paasche and Laspeyres bounding test, the index result fall within Laspeyres and the Paasche indices.

9. Monotonicity in current prices, if prices change then the index changes.

10. Test of permutation or price bouncing, if shops cross prices between months and the price in the later month are connected to the prices in the former month then the index should not change. (For example if the price of a good rises from 20kr. to 25 kr. or by 25 per cent and prices than go down from 25kr. to 20kr. or lowers by 20 per cent then the index should be unchanged and show no change).

\textsuperscript{13} Walsh (1901) 398 and (1921) 97

\textsuperscript{14} Extensive summation is to be found for example in Diewert (2002b), chapter 16, where he discusses 20 index tests.
11. Test of transitivity, indices calculated directly between two periods should show the same results if they are chained.

The tests are not equally important and no definitive rules are available to tell which of them are the most important. The results are therefore always dependant on judgement. The time reversal test, for example, says that the same result should be reached in the calculation whether the index is calculated forward or backward which can be of significance and therefore it is important that indices pass that test.

2.4 Elementary indices

Difference is made between calculation methods used for the elementary aggregates (the lowest level where weights are available) and methods used in the calculation of the total index. At the basic heading level weights are not itemised and the results calculated only with prices. At the elementary level the price changes can be viewed either as the average of price relatives (indices) or the relative of average prices. The main elementary indices calculated are:

\[ P_1 = \text{prices in period 1, } P_0 = \text{prices in period 0 and } n = \text{number of price observations.} \]

Average of price relatives (indices) named after Carli.

\[ P_{\text{CAR}} = \frac{1}{n} \sum \frac{P_1}{P_0} \]  

Relative of average prices named after Dutot.

\[ P_{\text{DRA}} = \frac{\frac{1}{n} \sum P_1}{\frac{1}{n} \sum P_0} \]

Geomean of price relatives (indices) named after Jevons.

\[ G_{\text{JAR}} = \left( \frac{P_1}{P_0} \right)^{1/n} \]

Relative of geomean prices named after Jevons.

\[ G_{\text{JRA}} = \frac{\prod P_1^{1/n}}{\prod P_0^{1/n}} \]

Harmonic mean of price relatives.

\[ H_{\text{AR}} = \frac{1}{\frac{1}{n} \sum \frac{P_0}{P_1}} \]

Relative of harmonic means.

\[ H_{\text{RA}} = \frac{\sum n/P_0}{\sum n/P_t} \]
The connection between these methods is that the geometric mean is always lower than the simple averages and higher than the harmonic means.

Carli index (2.3) which is the average of price relatives, is used in some countries but its use has declined considerably in later years. The index has many undesirable properties. "But we shall see that the simple arithmetic average produces one of the very worst of index number. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose"\(^{15}\). It does not pass the time reversal test (5), transitivity test (11) or the permutation test (10) and is therefore not appropriate in chain indices and considerable biased upward. Its use in the harmonised index of consumer prices is prohibited\(^{16}\).

The Dutot index (2.4) the relative of average prices passes most tests except the invariance to changes in the units of measurement (4). Different package sizes influence the results and the index also has indirect weights where more expensive goods have greater influence on the average than cheaper goods\(^{17}\). The Dutot index is appropriate when the goods are homogeneous and was alone used in the Icelandic CPI until March 1997.

The Jevons index (2.5 and 2.6) has been in use in the Icelandic CPI since March 1997. The geometric mean can be calculated as relative of prices or as the relative of average prices and these two methods are indifferent\(^{18}\). The Jevons index is superior in that way as it passes all major tests. Different package sizes do not influence the results and that characteristic is used in the calculation of the Icelandic CPI. The harmonic mean (2.6 and 2.7) can both be calculated as price relative and average of relative prices. The harmonic mean is the inverse of Carli (2.3). The harmonic mean does not pass tests (10) and (11) and is used on very small scale in CPI calculations and is always biased downward\(^{19}\).

### 2.5 Indices for calculation of the total index

When the price changes for the elementary goods have been calculated they are used to calculate the total index. The use of a fixed base index is the most common method. Superlative indices are a better choice because they are symmetric and work against the bias inherited in the Laspeyres and Paasche indices by neither favouring the old or the new weight information. The main reason why they are not used commonly is the lack of information about the present weights. The main superlative indices are:

Fisher price indices is defined as the geometric mean of Laspeyres and Paasche indices

\[
P_F = \sqrt{P_L \times P_P}
\]  

\(^{15}\) Fisher (1922) 29-30.

\(^{16}\) Eurostat (2001) 217.

\(^{17}\) It is very common to use it in price surveys in Iceland but it is ill suited for that purpose. Expensive items have more weights in the results and different package sizes change the results.

\(^{18}\) \(G_J = \prod \left( \frac{P_1}{P_0} \right)^{1/n} = \prod P_1^{1/n} \prod P_0^{-1/n} \)

\(^{19}\) If a geomean of the harmonic mean of price ratios (2.7) and Carli (2.3) is calculated the result lead to elementary indices with very similar properties as the Jevons index. First pointed out by Fisher (formula 101). Fisher (1922) 472, Carruters, Sellwood and Ward (1980) 25, Dalén (1992) 140.
It is the only superlative index that passes all the major tests and is in that way similar to the geometric mean in the elementary indices, Walsh index which is a pure price index with the quantity weights as the geometric mean of the quantities of two periods.

\[(2.10) \quad P_w = \frac{\sum_{i=1}^{n} p_i^t \sqrt{q_i^t q_i^0}}{\sum_{i=1}^{n} p_i^0 \sqrt{q_i^t q_i^0}}\]

Törnquist price index is defined as a geometric mean of price relatives weighted by the average expenditures in both periods.

\[(2.11) \quad P_T = \prod_{i=1}^{n} \left( \frac{p_i^t}{p_i^0} \right)^{S_i} \quad \text{where} \quad S_i = \frac{w_i^t + w_i^0}{2} \quad \text{and} \quad w_i, \quad \text{is the expenditure weight for a good } i \text{ as share of total expenditures} \quad w_i^0 = \frac{p_i^0 q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0}\]

### 2.6 Substitution in indices

"The problem of how to construct an index number is as much one of economic theory as of statistical technique"\(^{20}\). Assumptions that there is an economic connection between indices move the task of measuring inflation into economics. Indices can be looked at economically, i.e. how they show changes in substitution. When demand elasticities are observed the elasticity for the geometric mean is 1 and Dutot equal to 0. Geometric corrects for substitution in accordance with these assumptions simple averages not. The consumer keeps his total expenditure unchanged by exchanging goods that are increasing in price for those goods that are sold at lower prices. Elasticities are very changeable but indices can be corrected with formulas which take into consideration the fact that goods have different elasticities. Theoretically this has been solved by using formulas that use CES functions (constant elasticity of substitution). In a simple version of the Loyd-Moulton\(^ {21}\) formula different elasticities are taken into consideration, \(\sigma\) and weights, \(q\):

\[(2.12) \quad P_{LM} = \left\{ \sum_{i=1}^{n} q_i^0 \left( \frac{p_i^t}{p_i^0} \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \quad \text{for} \quad \sigma \neq 1\]

The elasticity coefficient reflects substitution and corrects the index for it. If Laspeyres weight is used substitution can be corrected without changing weights. Elasticity coefficients are

---

\(^{20}\) Frisch (1936) 1

\(^{21}\) Loyd (1975) and Moulton (1996).
probably unsteady and this method demands heavy data collection and effort and it is unlikely that it will be a workable solution to correct substitution bias.\textsuperscript{22}

3. New methods in the household budget survey

3.1 Detailed data from shopping receipts

Retail stores and other shops give their consumers detailed receipts and Statistic Iceland collects them in the household budget survey. The method was first used in 1995\textsuperscript{23} and after that in the continuous HBS that started in year 2000\textsuperscript{24}. In the survey participants return the receipts instead of copying that detailed information into the diaries. In the survey the total amount bought is written into the diary and the receipt put into a pocket in the diariebook. In the beginning the main idea was to make it easier for the households to participate by allowing them to return the receipts. In addition it turned out to be a new source of information, used to improve the weights for the consumer price index\textsuperscript{25}. “This method allows much more accurate estimates of the composition and quantity of household goods than otherwise would be the case”. The utilisation of this method also enables precise information to be gathered about consumer activities at much lower effort and cost than previous methods and shows a link between the goods purchased and the buyer\textsuperscript{26}.

3.2 Overview of data from receipts

More shops use bar code cash registers now than in the year 1995. This is partly due to the increased concentration in the retail market as there are today three dominating groups in the market. View over the amount of data that come from the receipts can be shown by summing up transactions from the receipts and the diaries\textsuperscript{27}. This can be viewed in two ways, either as the number of transactions or amount of expenditures. In the 1995 survey 41 per cent of all transactions came from receipts. This number was in the year 2000 about 69 per cent and 2001 it was at 74 per cent. For food and beverages 53 per cent of the records came in this way 1995, 84 per cent in the year 2000 and 2001 the records were about 89 per cent. Receipts covered more than 12 per cent of the total expenditure of households in the 1995 survey, 26 per cent in the year 2000 and about 31 per cent in the year 2001.

Considerable increase has been in the coverage of receipts from the year 1995 as they now cover nearly one third of the expenditures and approximately 75 per cent of the transactions.

\textsuperscript{22} Balk (1999), Haan (2001) and Opperdoes (2001).

\textsuperscript{23} This method was first described in the year 1995 when it was pointed out that with the receipts there would be a possibility to measure the expenditures on debt- and credit cards as that information was available on the receipts. Guðnason (1995) 173.

\textsuperscript{24} In the continuous HBS the sample for a three years period is similar in size as it was in the 1995 survey. The number of households in the sample each year is about one third of the number in 1995. Participating households were in the year 1995 1375, in the year 2000 657 and 611 2001.

\textsuperscript{25} Israel and Ireland are the only countries that have used this method systematically in a similar way. Israel in their 1986-87, 1992-1993 HES and as of 1997 in their continuous survey and Ireland in their 1999 HBS. However they did not use the detailed information that can be gathered in the way described here. Some other countries allow the use of the receipts without using them systematically as is done here. This is the case for Australia and New Zealand.

\textsuperscript{26} Rósmundur Guðnason (1997) 129.

\textsuperscript{27} Large and rare expenditures are collected in the quarterly questionnaire, but these records are not counted here. They are not taken into consideration in the final results.
3.3 **Detail of the receipts**

Usually the following information can be found on every receipt:

- Detailed breakdown of the total amount and the number of items. This opens up the possibility of balancing the data and in that way increase the data security. It also makes the estimate of the total amount of transactions easier and it can even be used before the survey is finalised.

- Name of the shop, it is therefore clear where the item is bought and the expenditure share for each household in each shop can be measured. That information is the base for the processing of the data and its use in the CPI, especially for creating the store weights.

- Timing and the date of the transaction which opens the possibility to map in an exact way consumers consumption behaviour, when and at which time of day they shop.

- Description of the item, quantity, price, and total amount. The information includes package size and brands and in some cases the quantities. Fruits and vegetables are often weighted at the cash register and in those cases the quantities are reported. This opens the possibility for making exact quantity weights.

- Form of payment, showing if the items are paid by cash, debet- or credit cards (e.g. Visa, master Card) or with check.

This information shows for example how much the costumer buys in each shop, how often and when. The place where the household lives is known so this information shows where the participants shop regionally. This detailed information have been of significant use in the making of the CPI base. In some cases the amount on the receipt is not detailed and in the 1995 survey this was the case of about 1.5% of the transactions which is less than in the former HBS. When the information from the receipts has been registered they have been balanced and the results can be used for shopping research. They are as of now only used for groceries in retail stores to make very detailed weights but could probably be used for other basic headings as well.

3.4 **Data from receipts and shops compared**

The first scanning of goods was done more than a quarter of a century ago. The development since than has been at a very fast speed and now the biggest part of shops sales takes place in this form. When goods are bought in retail shops they are scanned at the point of sale, the buyer gets detailed receipt for the transaction and the scanner data from each sale is captured in the outlets database. The receipt the consumer get is a mirror of the information available in the stores database. If all the receipts were kept together they would show the same result as the retailer sales information. If this information is compiled from the consumer it can be seen who the buyer is and what is bought which gives this data a special value exceeding the data that can be collected from the shop. In the databases there are also transactions from other sectors than the households and they cover therefore wider range of transactions. The receipt is a bill from the shop and shows always the total amount to be paid by the costumer. When the data is finalised that is a very handy characteristic as this data can

---

28 It was in the year 1974, Hawkes and Smith (1999) 284.
always be balanced\textsuperscript{29}. That is done by comparing the results that have been registered in the database of the survey with the total on the receipt. The information about the place of buying is available immediately when it has been registered in the database and the weights of shops are therefore always at hand.

3.5 \textit{The use of data from receipts}

Research in the 1995 receipts data showed that it was feasible for use in getting more detailed weights than was possible before. Comparison of result from the receipts and scanner data from the biggest stores for the year 1995 showed similar results. In addition to that the results showed that the biggest stores had in the year 1995 considerable market share among households outside the capital area although their activity was mostly concentrated in the capital area\textsuperscript{30}. Today the CPI for groceries is calculated with very detailed store weights.

When the stores weights for groceries were revised in December 2001 the data from the receipts collected in the continues HBS for the period January 2000 to November 2001 were compared with very detailed monthly sales information for the same period from the biggest group and these data set showed very similar results. Based on these data sets a correction of the weight shares for stores were done in December 2001. It is possible to use the total result because they are balanced as soon as they have been registered even if the survey is still ongoing and the final results not available, as was the case in the year 2001. Considerable development is possible with the use of this new very detailed information and the next step in its use could be as described in the following. "Further, shopping habits of households as mapped in the HBS could be used as a source for weights. This would be done by utilising information on detailed expenditure of typical costumers in each type of outlet. Calculation of the average price change would then be based on the expenditure of different households in the outlets. So for each outlet there would be different indices calculated for all types of households"\textsuperscript{31}. There are five types of households in the HBS which would mean that the elementary aggregates for the households buying of groceries in the stores would probably be around 20000. The price collection as it is now would be used to calculate more elementary aggregates than is now possible without any increase of the price collection. This shows in a nutshell the real coverage of the information gathered from the receipts and how they can be further used.

3.6 \textit{Other new uses in the processing- communication expenditures}

The part of private consumption which has increased at greatest speed in last years is telecommunications, especially mobile phone services. Telephone companies give consumers often detailed bills that can be utilised in the HBS. The bills are rather complicated and it is time consuming to register them in the survey as the interviewers do that in the quarterly questionnaire. To make the participation and the interviewers work easier participants were asked for permission to let the bills be electronically collected from the telephone company. This method has been successful and now most of the bills are gathered that way. This change led to increased safety in the data collection and more reliable results. The information gathered in this way was compared with the detailed data from the telephone company about amount of time units sold and both sources showed similar results\textsuperscript{32}. These results were first

\textsuperscript{29}One third of the expenditures is balanced in this way but this is the first HBS that is known to take advantage of this possibility.
\textsuperscript{31}Hallgrímur Snoranson and Rösmundur Guðnason (1999) 337.
\textsuperscript{32}Rösmundur Guðnason (2001) 634.
incorporated into CPI in the year 2001. Information was gathered about new services that had increased in coverage such as SMS messages. The uses of prepaid telephone cards have also increased considerably but as the service is prepaid there are no bills issued and information about the use can not be collected through the HBS but has to be gathered directly from the telephone companies. The number and amount spent on prepaid telephone cards is collected in the HBS. The method of collecting detailed bills has given good results and increased the safety in the measurement of the weight share and price changes in telecommunications in the CPI.

4. The calculation of the Icelandic consumer price index

What kind of index is the Icelandic consumer price index and how is it calculated? The consumer price index is a modified Laspeyres fixed base chain index with yearly links. The index resembles strongly conditional cost of living index. It allows for substitution by using the geometric mean and owner occupied housing is calculated as user cost.

4.1 The calculation of the consumer price index, overview

Elementary aggregate is the smallest unit in the index where only prices are available. It is split into approximately 6000 shop- and expenditure weights. Elementary indices (696) are calculated at the level of basic heading. Five different methods are used in the calculation of elementary aggregates in the consumer price index.

1. Relative of geometric mean prices (2.6) used in the calculation of nearly 39% of the base expenditures.

2. Weighted relative of geometric mean prices (see appendix) for groceries\(^{33}\), covers nearly 18% of total expenditures.

3. Laspeyres (2.1) or relative mean prices (Dutot) (2.4), covers nearly 38% of the index.

4. Superlative index (Fisher) (2.9), covers more than 2% of the expenditures.

5. Indices that cover nearly 3% of the index.

The calculation and price collection in the Icelandic consumer price index in January 2003

<table>
<thead>
<tr>
<th>Method of calculation</th>
<th>Elementary aggregate</th>
<th>Items number</th>
<th>Prices number</th>
<th>Expenditure shares per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative of geometric mean prices</td>
<td>221</td>
<td>2,083</td>
<td>5,436</td>
<td>39</td>
</tr>
<tr>
<td>Weighted relative of geo.mean prices</td>
<td>364</td>
<td>400</td>
<td>10,000</td>
<td>18</td>
</tr>
<tr>
<td>Laspeyres or relative of mean prices</td>
<td>99</td>
<td>1,185</td>
<td>2,891</td>
<td>38</td>
</tr>
<tr>
<td>Superlative index</td>
<td>7</td>
<td>75</td>
<td>203</td>
<td>2</td>
</tr>
<tr>
<td>Indexes</td>
<td>5</td>
<td>27</td>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>CPI total</td>
<td>696</td>
<td>4,170</td>
<td>18,557</td>
<td>100</td>
</tr>
</tbody>
</table>

\(^{33}\) Perishable items sold in food stores.
The number of subindices in the index is 696. Geometric mean is used in the calculation of 585 subindices, covering 57 per cent of the base expenditures, thereof groceries 364 and other 221. Number of subindices calculated as simple averages are 99 and other methods used for 12.

4.2 Sources for the base of the consumer price index

It is important how the elementary aggregates are organised and what sources are used when they are compiled. The main source in that work is the HBS. Data from that survey is directly used in the compilation of aggregates covering about 64% of the base expenditures in the CPI. When the data from the HBS is not adequately detailed more detailed information is gathered from other sources covers approximately 29 per cent of the base expenditures in the CPI. These expenditures are, alcohol and tobacco, medicine, medical services, petrol, local bus traffic, domestic flight, communication, entering fees for swimming pool and TV subscriptions.

Net weights are calculated for expenditure covering about 7 per cent of the index. It is used for expenditure on new cars, insurances and lotteries. Expenditure weights for new cars are calculated as the difference between cars bought and sold a method used in the calculation of the HICP and the national accounts. Insurances are calculated using net weights based on the revenue from the insurances when the claims have been subtracted and the income from financial activities added. The lottery weight is calculated in a comparable way, i.e. the winnings are subtracted from the income of the lotteries.

4.3 The calculation of the consumer price index

The calculation of the consumer price index is complicated and many methods used. Circumstances and available data govern the way of solution. Each year approximately 220 thousand prices are collected for the CPI on the average more than 18 thousand per month.

4.3.1 The relative of the geometric mean prices (2.6)

The method is used in the calculation of nearly 39 per cent of the index expenditure. The geometric mean corrects for substitution that arises when consumers change their consumption because of changes in prices. It differs how many prices are collected for each basic heading but the price change is calculated for all items available in both periods. The last price available is used if the item is not available at the time of price collection. It depends upon the nature of the basic headings how many prices are collected. When the items are heterogeneous, the basic heading not detailed and the prices vary the calculation will be more reliable collecting many prices within this method of calculation. Such is the case for car spares, toys, and books. When the items are homogeneous few prices can be sufficient for reliable price measurement. The results in the calculation of the geometric mean is independent of the package size and thus allows for different package sizes in the same basic

---

34 Similar amount of expenditures is calculated with the geometric mean both in the Norwegian and the US CPI. Dalton, Greenless and Stewart (1998) and Johannessen (2001).
35 The weight shares are calculated as a three years average as is done in the HICP.
36 The price collection covers both the price of the ticket as the ratio of winnings.
37 In addition on the average 1000 more prices are collected that are not taken into the calculation. This can for example be new items which are incorporated into the calculation gradually. Those items have been added after reference from the price collectors, information from the continuous HBS or from scanner data from stores.
38 If item is not available for three months the HICP rule is used and a new item is chosen.
heading. The substitution is only measured within the basic heading not between them according to this method of calculation.

4.3.2 Weighted relative of the geometric mean prices for groceries, perishable items (see appendix).

The method covers nearly 18 per cent of the base expenditures. Retailers are divided into for groups, Baug, Kaupás, Samkaup and Other. Each group is divided into stores that now totals eleven. Stores are the base unit in the calculation of the index for groceries and their activity reaches all regions in the country. Regional weights are not any longer in use. Store weights make the calculation of the index simpler and make it easier to treat changes in shopping habits, especially when one store replaces another. One reason is the concentration in the retail sector. Three groups now dominate the biggest part of the retail market for groceries. The fact is that the prices within a store are very similar, independent of localisation in the country.

In each store the geomean of prices for all items within a basic heading is calculated. The items can be of different size and make and items only available in one store can be included. The total is weighted in accordance with the share of each store in the total sale within each of the 364 basic headings for the groceries. The total number of price quotation is between 9-10 thousand each month. The prices are collected for more than 800 items and the store weights are nearly 4000. In the stores about 5500 average prices are calculated and when they have been aggregated within each basic heading they amount to 3500. For further explanation the subindex for rice can be observed (COICOP group 01111). Seven rice items are priced, in packages of different sizes, types and brands. The average prices in the 11 stores for these 7 items are 49 and they are all used when the average prices are calculated for the basic heading. The weight is based upon the amount of rice sold in each store. If an item is not available within a store the substitution effect is supposed to be as follows. The consumer first searches for another item within the same basic heading in the store. If no item is available he goes to another store and buys the item at the average price of all the other stores where the good is available. Outlet substitution between stores is in that way allowed. In the calculation of the groceries subindices the average is calculated for all items available and then compared to the average price of the same items in the base period of the index. The number of prices collected differs therefore from month to month and the average prices used each time also differs. One of the major advantages of this calculation method is the fact that all the prices collected each time are used in the price measurement. The basic assumption for use of this method is that many prices are underlying each average price calculated. To increase the probabilities that this condition is fulfilled the prices are collected from more than one shop within the biggest stores.

4.3.3 Laspeyres (2.1) or relative of mean prices (2.4)

Laspeyres or relative of mean prices is used for subindices that cover nearly 38 per cent of the CPI expenditures. The method is mainly used when there are detailed weights and exact

---

39 They were 13 in April 2002 in the link month. Since then Nykaup has been merged with Hagkaup and Ká merged with Noatún. In the HBS a total of 155 retail stores is registered.
40 From March 1997 to March 2002 regional indices for groceries were calculated in the CPI. The total was weighted by using regional weights.
41 Store weights have also changed in the year 2002 and 2003. For example two clock stores 10-11 (in Borgarnesi and Egilsstóðum) and one Hagkaup store have been changed into Bónus stores.
42 This differs from the calculation of the relative of geomean prices when the item is not available. In that case it is incorporated in the calculation at the last known price.
additional information available. In many cases no substitution effect occur as is the case with the geometric mean. When the weights are at the item level it does not matter if the geometric mean is used or the relative of averages. These basic headings are alcohol and tobacco, housing, cars, petrol, driving lessons, communications, lotteries, package holidays and insurances.

4.3.4 Superlative indices (2.9, 2.10 and 2.11)
This method is used for basic headings covering nearly 2 per cent of base expenditure. To be able to utilise this method detailed information about the expenditure structure has to be at hand. Fixed base indices do not measure such changes but superlative indices do. If a new weight is available (Paasche) then it is used along with the older weight (Laspeyres) and the result calculated as Fisher, Walsh or Törnquist indices. New tariffs for services often include readjustment of the tariff structure and price changes which can lead to considerable changes in consumptions patterns. The price changes are calculated according to the consumption pattern last available and also with estimated new weights and in that way the substitution effect which can be considerable, is taken into account. The expenditure are: TV tariffs, local buses, domestic flight, swimming and kindergarten.

4.3.5 Subindices calculated with indicies
Approximately 3 per cent of the CPI expenditure is calculated in this way. The indices are mostly used by convenience and their calculation in independent of the CPI calculation. The biggest item calculated in this way is housing maintenance where subindices from the building cost index are used. Other subindices change in accordance with the wage index such as child minding, au pair and home care services. Subindices for food, electricity and heat are used for calculating school accommodation. The CPI is used for calculating expenditure calculated as ratios such as stamp charges, title deeds and loan cost.

It is not appropriate to use the CPI for measurement in the CPI calculation it can have circular effect on the price change and the use of it is therefore an exception. It is difficult to know exactly about its coverage but rent is probably the basic heading mostly affected as about half of all rent contract in the rent sample are indexed by the CPI.

4.3.6 The calculation of the total CPI
The total index is calculated as modified Laspeyres index. The modification is such that the base from year 2002, mainly originating from the year 2000 HBS is price updated from year 2000 to March 2002. The price changes in the CPI are then calculated from the base in March 2002 to the month of calculation.

When the total index is calculated the average price changes are first calculated for each subindices. The subindices are aggregated and presented at different level of aggregation. The index is chain linked in March each year and the results for March on the old and new base used as link.

---

43 There is a substitution effect at hand but the weight is very detailed down to a bottle of alcohol and a packet of tobacco. The base is changed once a year and the price comparison within the year is paired.
44 Substitution effect is at hand but the weight is revised yearly according to a very detailed information about import.
45 Price changes are estimated after an extensive model of the Icelandic insurance market.
46 Usually these subindices do not change frequently.
The risk is always at hand that the index over- or underestimates the price changes in chain indices\textsuperscript{47}. Drift in indices can occur if there are big price changes in the link month. It can be seasonal price changes or other kind of changes and it is of special importance that they are not overestimated. The clothing sales are not over in all shops in March so there is need for special care when replacements are made to avoid drift (in this case under estimation of price change).

5. Changes in shopping habits and households shopping (outlet) substitution bias

The prices of the same or similar goods are often very different in shops and the consumers always need to take that fact into consideration. Consumer price indices measure the price change of private consumption and in reality the prices should be measured in the households. That is not done mainly because sufficient information about the shopping habits of households is not available. In the price collection for the index the prices are measured in the shops and the average prices are weighted by division of stores sales. When the households change their shopping habits the average prices of the goods bought change without anything happening in the shops, the prices there could be unchanged. To take such effect into consideration there is a need to change the weights of the shops and take that effect into consideration in the price measurement\textsuperscript{48}.

It is important that the sample of shops reflect households trade correctly. Shopping is steadily changing and consumers alter their shopping habits. Shops closes and new ones open or they close without any replacement. New shops are also opened but they do not necessary replace another directly.

Consumers organise their buying in accordance with this development. If a shop closes down consumers are forced to change but can by goods at the same place if the shop is replaced by another or in a new store. However if consumers buy the same goods in some other place at a lower price, it need to be taken into account in the index calculation, otherwise an households shopping substitution bias (often called outlet substitution bias) will occur\textsuperscript{49}.

Until now it has not been possible to follow such changes because of lack of information. Not accounting for the shopping substitution in the CPI equals to the assumption that all the price difference between stores is because their service level is different. In that case no price change is measured in the CPI when consumers change their shopping habits. "When pure price difference exists, a change in market condition make it possible for some households to switch from purchasing at higher to lower prices, for example by changing outlets from which they purchase. The resulting fall in the average price paid by households counts as price fall for CPI purposes, even though the prices charged by the outlet may not change."\textsuperscript{50} The discussion about this kind of bias has not been on a big scale internationally and corrections of indices are exceptional\textsuperscript{51}.

\textsuperscript{47} Frish (1936) 8-9, Szhulc (1983) 555-556.
\textsuperscript{48} If the price change would be measured with households weights their share of shopping in different stores would be changed when shopping habits changed.
\textsuperscript{49} In reality the price policy of a shop does not matter in connection with that but the consumption behaviour of the households does. Therefore it seems more appropriate to talk about households shopping substitution bias instead of outlet substitution bias. Outlet substitution bias arises only when a good is not available in a shop.
\textsuperscript{50} ILO (2002b), chapter 1, 38.
5.1 Inflation, changes in the organisation of shops and shopping habits

In April 2001 the inflation in Iceland increased considerably and from April to the end of the year the CPI rose by 7.3 per cent and the twelve month change at that time was 9.4 per cent. In the year 2002 the price changes diminished and from the beginning to the end of the year the CPI rose by 1.4 per cent.

When the inflation rose there were considerable changes in the organisation of shops and shopping habits in the country, especially in groceries as consumers moved their trade to shops where prices were lower. The share of self service petrol stations increased also remarkably in a few months time. Considerable increase in the share of discount stores occurred mainly within the Baugur group as the Bónus stores increased their market shares by approximately 40 per cent in the year 2001. Now they have the countries biggest share in the groceries market.

In addition to inflation few other factors were important. In October 2000 the Bónus stores allowed the use of credit cards in their stores, but before that there was only possible to pay by cash. The number of Bónus stores increased. They were 9 in the first quarter of the year 1999 but in the middle of the year the first shop outside the capital area opened, in Ísafjörður and a Hagkaup store in Kjörgarði was changed into Bónus store. The last months of the year 2000 till the end of 2001 6 new Bónus stores were opened, thereof two outside the capital area, in Akureyri and Selfoss. New stores were also opened in or near the shopping centres of Kringlan and Smárinn at the end of the year 2001. Last part of the year 2002 two clockshops (10-11) were turned into Bónus stores, in Borgarnes and Egilsstaðir. In April 2003 Hagkaup store in Keflavík was turned into Bónus store and the number of shops was then 20. Similar development but on a smaller scale happened with the second largest group, the Kaupás group, as they opened the Krónastores. These great changes in the organisation of shops and shopping habits especially in the year 2001 were so fast and massive that they had to be taken into account in the CPI price measurement.

In the Bónus stores the prices are often lower than in other type of stores and that price difference has remained even if the quality difference between them and other stores have diminishes. In that connection some changes may be pointed out that have incurred at the Bónus stores mainly in the year 2001.

In the beginning the shops were placed in simple housing, away from the main shopping centres, few commodities were available and they had do be paid with cash, few cash registers were available in each store and the opening hours limited. The newest stores are placed in more spacious housing, in or near shopping centres like Kringlan, Mosfellsgarður or Smáralind. The costumers who in the beginning paid for their commodities with cash can now use credit cards. The item selection has increased, there are more cash registers in each shop and they are now open every day of the week.

These shops are now widespread around the country, in Ísafjörður, Akureyri, Selfoss, Borgarnes and in Egilsstaðir. The selection of goods is still more limited than in other stores so consumers have to go to other shops if they wish to have more diverse selection of

---

52 The drift cost in low price shops is not necessary higher in shopping centres than elsewhere. Low price stores attract costumers which can effect the amount the rent that they pay.

53 In the year 1999 grocery items were 1100 but had reached 1400 in the year 2000. Kaupbing (1999) 9, Íslandsbanki (2000) 19.
commodities and the items are often different, especially regarding sizes of packages. Of the basic heading for groceries in the index only 15 are not available in these stores. It is easier for consumers to shop at the Bónusstores as they have grown in number and are therefore easier to reach.

The trademark of the stores has become the best known within the retail trade in the country. That decreases consumers search cost and has certainly actuated these large changes.

Consumers can choose between different level of service when buying petrol. The service stations now offer the costumer the choice between full- or self service. Self service stations have rapidly grown in number. In the year 2001 approximately 10 per cent of costumers moved from full service to self service when buying petrol.

Petrol is a homogeneous good and the difference between the forms of service is less than before. Since the self service stations became more accessible the queues have vanished but they were common before. Thus the consumer now spends the same time at the station whether he uses the full service or not and often the self service is quicker. The quality difference between self service and full service at petrol stations is therefore very small or none at all.

5.2 Quality adjustment and the changes in the outlet sample

5.2.1 Quality adjustment for groceries

In price measurement under estimation of quality change of goods or services leads to over estimation of inflation. The danger is biggest when inflation increases suddenly and the buying pattern of households changes on a large scale. The service level of the shop influences consumers shopping and it includes all factors that influence quality and features that decide the outlet type. Such factors are commodity selection and availability of items, number of stores and their localisation, number of cash registers, opening hours and the method of payment. All these elements have to be reflected correctly in the price measurement. Quality is both subjective and individual and therefore a considerable difficulty is faced in measuring the level of service, except for the selections of goods. The difference between discount stores and stores of other type is less than before regarding the components that were listed above.

There is a possibility to measure quality difference in the level of service by comparing the selection of items available or the goods that are joint in the stores as they reflect the part of service level that can be price measured. Some of the commodities that were available in the old store are not available in the new one, there are other types of packages and brands. The consumer shops in the same place but in a new type of store. The method used in that case is to compare the items available in the shop which closes down with the items in the replacement shop. The differences in price level between the stores for goods available in both stores are used to measure the price change. Difference in weights of the shops does not influence to a great extent the result of the calculations of the CPI if the price changes within the shops are similar. Simulations tests in December 2001 by changing stores weights and recalculating the index gave approximately the same results as the published CPI.

54 For example Bónus stores do not sell tobacco and some types of meat do not have any subindices in Bónus.

55 It can be measured by regression coefficients. Such research has been conducted and shows that prices have the biggest attraction. (Personal communication with Jón Scheving Thorsteinsson, managing director, Baugur ID).
When consumers move between stores and buy the same goods but at a lower price a part of that is a pure price change and it is not until that price change is taken into account that the CPI is lowered.

5.2.2 Changes in the outlet sample

From the year 1997 - 2001 six stores that were in the CPI sample closed down and in all these cases new store was taken into the CPI instead. The price changes were measured by comparing prices of joint goods\(^{56}\) and the difference in the level of service quality adjusted in that way.

The changes in the shopping patterns in 2001 were treated in a similar way. The result was that nearly half of the increase in the market share of the low price stores was taken into account as a price reduction and the other half as a quality adjustment due to difference in assortment. The result was incorporated into the CPI in December 2001 and in April 2002, but then store weight were adopted which make it easier to incorporate new stores into the sample instead of shops that drop out. In three cases have there been changes in the year 2002. First Nýkaup in the Kringla shopping centre were changed into Hagkaup store in May 2002\(^{57}\) and the weight for Nýkaup was moved over to the Hagkaup stores. KÁ in Selfossi was turned into Nóatún store in the middle of the year and the weight for that store moved over Nóatún. In December 2002 two new Bónus stores were incorporated instead of two clock stores (10-11) and their weight moved to Bónusstores. The difference in assortment was some, even if the number of basic heading is similar, especially regarding difference in package sizes and brands\(^{58}\). One of these shops was in the CPI sample so that change would have been measured with older methods. The changes have continued and in May 2003 more than 1% of the total groceries weight were moved between stores\(^{59}\).

Store weights have proved their relevancy and they make treatment of sudden changes in shopping habits much easier.

5.3 Correction of stores weights for groceries

When the inflation escalated in the year 2001 consumers directed their shopping to shops with lower price level. On the whole nearly 10 per cent of consumers changed their shopping habits in a very short time period, buying goods in shops with lower prices. Consumers think that there are advantages in such an exchange. It is not known that such sudden changes in shopping habits have happened in such short time interval in other places\(^{60}\).

Because of this development it was necessary to change the weights effecting the CPI measurement. The CPI base is revised each year but when the changes in the retail market

---

\(^{56}\) Examples of such changes that influenced the CPI was when Nettó store opened in Reykjavik in August 1998, replacing store that was in the CPI sample and the change was measured immediately. Indirect influence was also at hand as other stores lowered their prices when this happened.

\(^{57}\) Hagkaup in Kringlan shopping center had been changed into Nýkaup in June 1998.

\(^{58}\) Approximately half of the goods available in the clock store (10-11) was available in the Bónus store.

\(^{59}\) Hagkaup in Njarðvik was closed replaced by a Bónus store in April 2003. In April 2003 a clock shop (11-11) in Mosfellbær was replaced by Krónustore. The director for the Kaupásgrybuf have announced that more such changes will occur in the year 2003.

\(^{60}\) “I think that this has never happened in the world before. Although it is possible that WaalMaart growth in some regions in the USA may have been similar.” (Personal communication with Jón Scheving Thorsteinsson, managing director, Baugur ID).
were so big it was necessary to take it into account in the calculation of the CPI as soon as the information was available. An exact information about the groceries sale was at hand so mapping the changes was possible. This information about market share came from cash receipts in the continuous HBS and in addition exact information was collected from the biggest group about monthly sales of stores. Information from both these sources was compared showing similar results. Again the correction was made by measuring the level of service reflected in common goods. The correction of the store weights for groceries in December 2001 led to a price cut of the food component in the index by 1.3 per cent, bringing down the CPI by 0.27 per cent. These effects were estimated again in April 2002 leading up to a further lowering of the index, by 0.10 per cent. At the same time the effect of changes in the shopping habit for petrol led to the lowering of the CPI by 0.08 per cent. In April and May 2003 additional changes in the shopping structure were estimate leading up to a 0.07 per cent’s lowering of the CPI.

The total effect on the CPI, by correcting for changes in household’s shopping substitution on groceries and petrol in December 2001, April 2002 and 2003, was a decrease by nearly 0.55 per cent.

6. Housing in the Consumer price index

In the years 2000-2002 about 82 per cent of Icelanders lived in own housing according to the HBS. To buy an own house is the aim of most people in Iceland and is usually the biggest investment of individuals in their lifetime. House is a place to live in and at the same time an investment and to price measure the use has been a problem in CPI calculation. The use of own house is calculated as imputed rent in the consumer price index, but the buying of the house is an investment and therefore not taken into account in the calculation.

6.1 Methods in the calculation of owner occupied housing

Four main methods are used in the calculation of owner occupied housing. Rental equivalence, user cost, net approach and the method of payments. The most convenient method is to calculate rental equivalence where the result for rented housing in the rent market is used. The rental equivalence is like a rent for comparable apartment or houses. The primary assumption is that the rent market has an adequate coverage so market rent can be measured for comparable types and sizes of properties and that the results can be used as equivalence to changes in rent for own housing. Secondly that the rent market is not controlled, rent is not subventioned by governments or the market prices controlled in other ways. This method cannot be used in Iceland because of the small rent market that also has different composition than the stock of own housing. The method is used in Denmark, Germany, Netherland, Norway, USA, Switzerland and Japan.

Another method measures user cost, it is used in the Icelandic CPI. The service of living in own house is measured as the cost. Annuity is calculated of the market price of the house and the imputed rent measured according to a certain real interest rate and depreciation rate. The real interest rate is based on the required rate of return (the opportunity cost) on the capital bound in the property, independent on whether it is owned capital or a loan. The wear of the

61 In April 2002 a third source of information, namely VAT-reports from the Internal revenue directorate, were used. They supported and strengthened the result of other source’s.
62 The fifth possibility is to leave owner occupied housing out of the CPI.
63 This information is based on Boldsen Hansen (2000).
property is taken into account, it is depreciated in accordance with the expected lifetime of the house. In addition minor repairs are taken into consideration and public tariffs connected to the house such as sewer, garbage and water\textsuperscript{64}. It is only the use of the house that is accounted for not the capital gain of the investment. The price change is measured by all properties sold in the country. The consumer price index is a short time measurement tool for estimating price change and assumed that no substitution between living in own housing and renting is possible. Some countries calculate their housing as a user cost but none of them uses real interest rates in that calculation except Iceland. They are Finland, Sweden, Iceland, Ireland, United Kingdom and Canada.

The third method is to measure the housing with a net method. The net housing cost is the value of new houses in excess of depreciation of the stock. The housing is taken as any other expenditure in the index when it is bought or built and treated equally as other durable goods in the CPI. This method is for example used for cars, electrical equipment; they are capitalized and taken into consideration in the CPI at the point of purchase. The price change is, using only price changes of new buildings. Included are houses constructed by the consumer and properties bought directly from a construction company or real estate agents. In addition it’s necessary to take into account properties that are bought directly from other sectors in the economy. This index has many similar properties as a producer price index for buildings\textsuperscript{65}. Each year there is a different amount built of new houses. The net change in housing can therefore be negative in some years and the net weight for housing too if calculated by the this method. It will therefore probably be necessary to calculate the weight as an average over few years. The weights will be more volatile when the net method is used than is the case for rental equivalence or user cost and the net weight will usually be lower\textsuperscript{66}. The method is used in Australia and New Zeeland.

The fourth method is the payment method. The payments when buying the house are counted for, payment on loans, interests, reparations and renovations. This method is similar to the method used in the Icelandic CPI in the years 1988 to 1992. Its main failure is the use of the nominal interest rate, as it in reality partly reflects the inflation, and the ignorance of the fact that the use is spread over a longer period of time.

Some countries consider housing mainly as investment and argue that it should therefore, similar to other investment, not be taken into the CPI. The reason might also be that price information about the property market is not available, making the above mentioned methods impossible to apply. The owner occupied housing is left out of the CPI\textsuperscript{67} in Greece, Italy, Spain, Portugal, Belgium, Austria, Luxembourg and France.

\subsection*{6.2 Owner occupied housing in the Icelandic consumer price index}

The method to calculate housing as user cost and price update it with price changes of all properties sold was adopted in November 1992 and has been mostly unchanged since then. In the beginning the prices were only measured in the capital area but since April year 2000 they

\textsuperscript{64} Additional user cost model is based on cash flow. In that method the nominal interest rates and depreciation are measured.

\textsuperscript{65} The method has been under study for inclusion in the HICP.

\textsuperscript{66} Even half of it (Diewert 2002a) 62.

\textsuperscript{67} Share of owner occupied housing in these countries is: Greece (75), Italy (78), Japan (78), Portugal (66), Belgium (65), Austria (50), Luxembourg (72), France (54). Boldsen Hansen (2000) 12.
cover the whole country. The base for the calculation is the real estate value of the house and that information is collected in the HBS. The user cost is calculated with real interest rate that is now 4 per cent and depreciation rate of 1.25 per cent of the house’s real estate value. The price measurement is monthly updated by price index for properties sold. Owner occupied housing covers imputed rent, minor repairs and other cost, such as tariffs for sewer, garbage and water. The weight for housing cost comes from the HBS as an imputed figure and the monthly weight is calculated as an user cost of the real estate value. There are three factors taken into consideration when the user cost is measured. First the base for the annuity, second the real interest rate and third the depreciation.

6.2.1 Weight as the annuity base

The base is the estate value of the property. “the law about the measurement of the real estate value says that it should be based on the market price of the property. According the 1 paragraph. of the law nr. 6/2001 shall the estimated value be the discounted market value as estimated last November.” In the middle of the year 2001 the real estate value was revised by the Land registry of Iceland after extensive research by regression analysis. The base for the analysis was the capital area and the estimate for other parts of the country was calculated with regional coefficients. The value of all properties in the country are measured in a harmonised way based on information about sold properties. This basic information is the same as used in the price measurement of housing in the CPI and the real estate value is therefore well suited as a base for the user cost calculation.

6.2.2 Real interest rate

Usually when consumers buy properties they finance it partly which own equity and the rest with loans. The user cost model is based on the financing as measured in the sales contract that are the base for the price measurement. This division of financing is used to calculate the real interest rate used in the calculation. The payment according to the sales contracts is taken as the buyers own equity. The real interest rates used for the own equity reflect the long time rate of return of housing investments. In the calculation of the imputed rent it is supposed that own equity is approximately the half of the house price. When deciding the real interest rate for own equity and because the rate of return has a long term character it was set as the estimated rate of return for the pension funds in the country. When these methods of calculation were adopted the long time interest rate of the pension funds was 3 per cent and it has been kept unchanged since then. Other forms of payment according to the contracts are usually new loans or loans that are taken over and the real interest rate used in the calculation is according to the credit terms. These are most often loans from the Housing financing fond or loans that are taken over by the buyer from the old State housing board. Other financing is mainly originating from the pensions funds. The average real interest rate measured in this way has been around 4 per cent in the last years.

---

68 Correction was made for over measurement of price changes of houses in April 2000 lowering the CPI by 0.35 per cent. At the same time under estimation of rent was corrected leading to a 0.34 per cent increase in the CPI.
71 This share could of course partly be financed with loans, not with properties.
72 The long time rate of return for the pensions funds is now in the range of 2.0-3.5 per cent. Long time rate of return according to the liability law is 3.5 per cent.
73 Real interest rates for these payments lie in the range of 5.0 to nearly 9.0 per cent.
74 The capital gain of the investment is not taken into account here. It is added to the rate of return of the investment taken into account in this method. Investments are not taken into the calculation of the CPI. The
6.2.3 Depreciation

Depreciation should reflect the tear and wear of property and is always very uncertain. The maintenance differs and therefore also the obsolescence. The depreciation rate used in the index is 1.25 per cent of the real estate value. Land is not depreciated as it does not wear over the time. The depreciation should therefore only be calculated of the value of the building. By convenience the depreciation is calculated of the whole value of the housing stock. The depreciation is therefore in reality 1.5 per cent of the house setting the life time of the property to approximately 67 years. The housing stock divided after the building year published in the real estate registry at the end of the year 2001 shows that 90 per cent of all property is constructed after the year 1940, more than one third in the period 1960-1980 and one third is constructed later. The depreciation rate seems therefore to be in accordance with the property stock after building year. The depreciation prerequisite depends on the maintenance of the stock but in the CPI only minor reparations are taken into account.

6.3 Price measurement of properties

6.3.1 Real estate prices and price indices for housing

The market prices are gathered from the sales contracts which the Land registry of Iceland has collected for many years. These data are the base for their evaluation of house’s real estate value and is also used for the measurement of market prices used in the CPI. The sales information is collected through the District Commissioners and the change in the ownership of the property can not be registered unless the sales contract is at hand. Between 8-10 thousand sales contracts are collected each year covering 8-10% of all properties in the country. In the sales contracts the form of payment is shown and that information is the base for the calculation of the present value of the contract. The rate of return in the contract after type of payment is based on market information. The rate of return is measured every month and if the change exceed a minimum the rate of return is changed.

Changes in the market price and the rate of return influence the price measurement. When the rate of return goes down the, the present value goes up and increase in the rate of return lowers the present value. The present value of the contract is used for the price updating of properties in the CPI. The price measurement concept is the same as is used in other parts of the CPI and the prices taken into account are those the consumer pay in reality for goods and services.

The prices are the average prices in the country. The total price information from all the sales contracts are used in the calculation of the imputed rent. In the calculation the combination of the house’s size is kept fixed, based on the sale’s volume in each category for the last three years. The price change is measured for houses (13 per cent share) and

---

75 Orn Ingvarsson (2002) 261
76 It is a question if major reparation should be added to the value of the housing stock before the weight is calculated.
77 It does not matter whether the number or sales value is used.
78 The rate of return has been changed four times in the year 2002.
79 This has been in this way since March 2000. The index for the whole country was then calculated backwards to March 1997.
apartments (59) for the capital area (72) and houses (15) and apartments (13) outside the capital area (28).

The emphasis is on the of price change comparison within three groups of properties not between types of properties or between regions. The size groups are eight. In total nine subindices for housing are calculated in the capital area and eight for housing outside it. The main indices are four, for houses and apartments in and outside the capital area. In total the calculation is based in all on 21 subindices, used in the calculation for the total index for property prices. The calculation is based on three month’s average with one month’s time lag. The sales contracts in April refer to the period January to March and in May for the period February to April e.t.c.

6.3.2 Property prices in the consumer price index

The user cost is calculated with assumptions about the rate of return and depreciation. These change very seldom, are approximately. The imputed rent reflects therefore the market prices changes in properties. Asset prices are incorporated into the CPI in that way. It does not matter in reality if the net weight would be used. It reflected the market prices of houses in the same way. The methods will therefore show similar market price change given that the price change of new housing and all properties sold (where the new houses are also taken into the measurement) are similar.

This method of measuring price changes in the property market in the consumer price index has been successful. Research done by the Icelandic Centralbank showed housing as an important part in measurement of future inflation. “On the other hand lot of information is lost if the housing post is taken out of the CPI”. The connection between housing prices and inflation was pointed out. “The significant correlation between housing prices and the CPI for more than two years ahead might indicate that there is unused information about future inflation in the development of housing prices”.

A strong connection can be observed between price changes in the rent market and the changes in the market prices of housing, i.e. both indices seem to develop in similar way in the longer run even though they in shorter time periods can move in dissimilar cycles. Imputed and paid rent showed similar changes from March 1997 until the middle of the year 2002 but that has changed in the last month of 2002.

High rent prices give an incentive for buying own houses. Contraction or resistance in the housing market sales has emerged in grown tendency to let properties go as payment into the sale entailing an increased supply of apartments for renting which again entails to lowering of the rent.

80 Contracts from places outside the capital area arrive with two months timelag.
81 Except that the weight would be lower.
82 One of the main uses of the CPI is for indexation of long term loans for housing. It is therefore proper that property prices are reflected in that index.
85 Share of properties and other capital goods in the contracts were nearly 20 per cent of the value in the year 1995 and about 4 per cent in 2001. Fasteignamat Ríkisins (2002) 31
Appendix: Calculation of elementary indices for groceries in the Icelandic consumer price index

Jevons index is used for the calculation of elementary indices for groceries in the CPI. The stages in the calculations are:

1. For each price observation, i, in store j, within basic heading k, an unweighted geometric mean is calculated:

\[ P'_{jk} = \left( \prod_{i=1}^{n} \left( P_{ijk} \right) \right)^{1/n}, \text{for } P_{ijk} > 0 \]

Where:

- price observations, \( i = 1, \ldots, n \)
- stores, \( j = 1, \ldots, m \)
- basic heading, \( k = 1, \ldots, h \).

To make the calculations technically easier the logarithms are taken on both sides:

\[ \ln(P'_{jk}) = \frac{1}{n} \sum_{i=1}^{n} \ln(P_{ijk}) \]

Operations (1) and (1a) are carried out in the same way for all basic headings in March (base month) each year and in the month of calculation.

The following price tables are available after the first step in the calculations:

- \( P'_{jk} \): Average price for basic heading k, in store j, in the month of calculation.
- \( P^0_{jk} \): Average price for basic heading k, in store j, in the base period.

2. Weights are for the first time taken into consideration in the calculation at this step. Stores have a weight share in each of the basic headings, where \( \sum_{j=1}^{m} W_{jk} = q_{k} \) \(^{66}\). The weight share is as based on:

- \( W_{jk} \): Households expenditure shares in a store for basic heading k, and \( q_{k} \) is the base expenditure share in the index for basic heading k and \( \sum_{k=1}^{h} q_{k} \) is the total expenditure for groceries.

Weighted geometric means \( P'_{jk} \) and \( P^0_{jk} \) are calculated

\[ \overline{P}_{k} = \prod_{j=1}^{m} \left( P'_{jk} \right)^{w_{jk}} / \sum w_{jk}, \text{ for } P'_{jk} > 0 . \]

\(^{66}\) Special case: If \( P_{jk} = 0 \) in a store that has \( W_{jk} > 0 \): then a scaling of the weight is made moving it over to the other stores.
\( \bar{P}_k \) is calculated in the same way.

The logarithm of the ratio of the averages is taken and the equation adjusted to a convenient form.

\[
\ln \left( \frac{\bar{P}_k}{\bar{P}_k^0} \right) = \sum_{j=1}^{m} W_{jk} \left( \ln \bar{P}_{jk} - \ln \bar{P}_{jk}^0 \right)
\]

The result \( \frac{\bar{P}_k}{\bar{P}_k^0} \), is the price change for basic heading \( k \), from the base period of the index to the month of calculation. It is used in the calculation for each elementary index, \( v_k \).

The index for groceries is calculated as \( \frac{\sum_{k=1}^{h} q_k^0 \bar{P}_k}{\sum_{k=1}^{h} q_k \bar{P}_k^0} \), i.e. an Laspeyres index.
References


Arnltódur Ölaufsson (1880), Auðfræði, Hið Íslenkska Bókmenntafelag, Kaupmannahöfn


Balk, M. Bert, (1999), On curing the CPI’s substitution and new goods bias, Statistics Netherlands Research paper nr. Xxxx


Diewert W. Erwin (1996), Sources of Bias in the Consumer Price Index, *The University of New South Wales Discussion Paper 96/4*.


Hawkes, William, Smith, Roger (1999), Improving consumer price measurement through the use of scanner data and market segmentation, *Proceeding of the Measurement of Inflation Conference, Cardiff University, Cardiff, UK, August 31-September 1, 1999*, 283-397, Mick Silver og David Fenwick editors, Eurostat, Cardiff University, ONS.


Íslandsbanki FBA (2000), Baugur Útboðs og skráningarlýsing Desember 2000, Íslandsbanki FBA.

Johannessen, Randi (2001), Mikroindexformel i konsumentprisindex, Notater 2001/64, SSB.


Kaupþing FBA (1999), Baugur april 1999, Kaupþing FBA.


On the treatment of newly significant goods and services in the HICP: the cases of electricity markets, e-commerce and mobile telephony

Alexandre Makaronidis
Eurostat, Luxembourg

Acknowledgements are due to Patricia Klees-Friedrich (DESTATIS, Germany), who contributed during her stay with Eurostat a number of working papers on this issue, on which a large part of this paper is based; Don J. Sellwood (former head of the CPI harmonization project at Eurostat) and Keith Hayes (Eurostat) for their helpful comments.

Abstract: The paper discusses the EU harmonised standard concerning the treatment of newly significant goods and services (NSG&S) in the Harmonized Index of Consumer Prices (HICP) and provides some specific information regarding the treatment of markets in the process of deregulation (e.g. electricity), e-commerce, and mobile telephony.

1. Introduction

The meaning of "new goods" is not always precise. The difference between new and improved is in fact arbitrary and there is no sharp dividing line between new models and varieties and genuine innovations which fulfil needs that could not be fulfilled before. Neither the index form nor the "step" or the frequency of renewing the basket of the index can address the real problem: the treatment of new goods. The risk of bias as a result of inappropriate treatment is rather high because the introduction of new models and varieties is often the "piggyback" to implement price increases.

The requirement in the framework HICP Council Regulation\(^2\) to 'maintain the relevance of HICPs' means that steps must be taken to ensure that HICPs keep broadly in step with each other and are up-to-date in terms of market developments. The corresponding minimum standard on NSG&S is aimed at ensuring that new products are incorporated in the HICP as soon they achieve a sales volume of over one part per thousand of total consumers' expenditure in the Member State\(^3\). Eurostat is monitoring compliance with the HICP framework in order to ensure that the EU harmonized standards are met by EU Member States (MSs for short).

---

\(^1\) The views expressed in this paper are those of the author; they do not necessarily reflect an official position of the European Commission (Eurostat).


\(^3\) Commission Regulation (EC) No 1749/96 of 9 September 1996 on initial implementing measures for Council Regulation (EC) No 2494/95 concerning harmonized indices of consumer prices (OJ L 229, 10.9.1996, p. 4), Article 2(b) "Newly significant goods and services' are defined as those goods and services the price changes of which are not explicitly included in a Member State's HICP and the estimated consumers' expenditure on which has become at least one part per thousand of the expenditure covered by the HICP.".
Early compliance monitoring exercises provided evidence that the NSG&S standard is interpretable in several ways. The aim of this paper is to summarise the procedures on NSG&S and split them into the elements needed to make the standard operational. Furthermore, it introduces briefly the notions of “product offers” and “(basic) user purposes”, currently under development at Eurostat as a means to support, among other things, the operation of the standard in practice. The paper also offers some thoughts on the treatment of specific cases such as electricity markets, e-commerce and mobile phones.

2. Newly significant goods and services

In the context of the compliance monitoring exercise EU MSs reported on the procedures they had put in place in order to systematically identify NSG&S.

Typically, it is the price collectors who are supplying information on new goods and services (whether genuinely new or replacement products). Head offices observe market developments centrally and assess their observations together with the information supplied by price collectors.

In addition, MSs use information from National Accounts, Household Budget Surveys, production statistics or other sources like food surveys or import statistics. Furthermore, they examine magazines and periodicals (trade and consumer journals, commodity publications) and review newspapers and television advertisements. They collect information from industry, trade organisations, and trade representative bodies, institutional users, and associations of retailers and producers and consumer associations.

In MSs with fixed base indices (e.g. basket updated on a three to five year cycle) the search for newly significant goods and services is ongoing but particular focus is paid at each re-basing period when target samples are reviewed and new replacements or additional items are taken into the HICP. In Denmark each month one of the twelve main COICOP/HICP headings is checked, so that over one year all headings are checked. This may be seen as the most systematic procedure reported by a Member State and might be commended as such.

The term ‘newly significant’ was generally interpreted as ‘being new to the index’. There are two ways by which new products are introduced into the HICP if they have gained a significant part of consumption:

either

- Replacement: the new product replaces an already existing product that has lost importance; in other words a more up-to-date representative for a need or purpose already covered by the index is brought into the sample (e.g. a special variety of car tyres is replaced by a more representative one),

or

---


5 COICOP/HICP is the Classification Of Individual COconsumption by Purpose adapted to the needs of the HICP.
• **Addition:** the new product is brought into the index in addition to the already covered products as representative for a newly significant or modified need or user purpose. The distinction is important both operationally and conceptually. The new product is either a replacement serving an existing reference purpose or it is additional in serving a purpose not yet represented in the index. ‘Newly significant goods’ relates to additions and not to replacements.

Additions (i.e. NSG&S) can have two main reasons:

either

- a new product that had not been represented in the index and would not normally be considered as a replacement because it was radically different from the existing variety (e.g. mobile phones). It would be added as a new category within an existing category,

or

- a product that was available but not explicitly represented because its consumption was too low. Inclusion was not a replacement within a category but a new category within a category (e.g. spaghetti in Denmark or lamb meat in the Netherlands).

In case of additions, the price of the new product is collected in addition to the already observed product; the minimum standard on NSG&S offers one of the following treatments:

- either adjusting the weights of the relevant category of COICOP/HICP, or
- adjusting the weights within the relevant category of COICOP/HICP, or
- assigning part of the weight specifically to the new product (i.e. underneath the 4-digit level of COICOP/HICP).

Replacements on the contrary are not affected by this standard because they do not have in principle an effect on the weights.

The differences between newly significant products and replacements are summarised in table 1.

According to the above, the minimum standard on NSG&S implies that:

• there is a certain expenditure group which has not been explicitly covered by the Member State’s HICP because its weight was lower than one part per thousand;

• this expenditure group must be covered explicitly because it has reached one part per thousand of the expenditure covered by this HICP;

• this group is represented by a certain NSG&S which itself does not need to have a weight of at least one part per thousand of the expenditure of this HICP.
Member States need thus to check

- whether the proposed products represent an expenditure group that is not yet explicitly covered by the index and

- whether this expenditure group has reached one part per thousand of the expenditure represented in the index,

- how this expenditure group should be explicitly implemented into the HICP and

- whether it could be represented by more appropriate price representatives than those proposed (some appropriate representatives might be contained in the list as reported by other Member States).

Table 1: Newly significant goods and services versus replacements

<table>
<thead>
<tr>
<th>Additions (NSG&amp;SSs)</th>
<th>Replacements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Identifier:</strong> New category of product (not yet existing in the HICP or implicitly covered) but not necessarily (technically) new (e.g. lamb meat).</td>
<td>Same category as existing product. E.g. new model of a TV set.</td>
</tr>
<tr>
<td><strong>Threshold:</strong> For the first time consumed to a significant extent (one part per thousand).</td>
<td>No threshold set, but current sample should represent all models or varieties within a product category.</td>
</tr>
<tr>
<td><strong>Content:</strong> Can consist of different variants (brands, models); one or many products are introduced (serving the same purpose).</td>
<td>Defined by previous referent. Basically, one to one replacement; one existing price representative is exchanged into another more representative product. However, many to one is not excluded.</td>
</tr>
<tr>
<td><strong>Weighting:</strong> The particular weight is either not yet existing or needs to be adjusted in order to account for the newly significant product.</td>
<td>No immediate impact on weights.</td>
</tr>
<tr>
<td><strong>Price impact:</strong> Introduction does not in itself impact on measured inflation.</td>
<td>Introduction should impact on measured inflation to the extent that the replacement involves price change after potential allowance for quality change.</td>
</tr>
</tbody>
</table>

In order to ensure that HICPs keep broadly in step with each other and are up-to-date in terms of market developments, Eurostat is monitoring regularly the inclusion of NSG&S into the HICP. In the context of this exercise, Eurostat may also suggest NSG&S for inclusion. In case of exclusion of NSG&S proposed by Eurostat, MSs should provide evidence that the maximum expenditure that could be represented by each of the proposed products was below one part per thousand of the expenditure covered by their HICP. Where appropriate, products should be grouped to higher-level expenditure groups (e.g. ecological/organic carrots and
potatoes belong to the same COICOP sub-index and might therefore represent the same expenditure group).

3. Genuinely new products versus new models and varieties

The distinction between genuinely new products and new models or varieties of existing products is not always a clear cut one. The question is how to design an appropriate device, with sufficient discriminatory power, in order to make such a distinction operational. For the HICP in particular, one wishes to ensure that no new models and varieties of reference products are excluded from the current sampling frame and new products which have become significant are included in accordance with the NSG&S standard, i.e. in the December reference sample.

The notions of “product offers” and “(basic) user purposes” are currently being development at Eurostat and discussed at the HICP working party as a means to support, among other things, the operation of the NSG&S standard in practice. They should allow a principled and operationally feasible approach for distinguishing between genuinely new products and new models and varieties of existing ones.

A “product-offer” (PO) is an observable entity comprising a single product (a good or service) offered for purchase at a stated price, in a specific region and location (outlet) under specific conditions, at a specific time and in a certain socio-economic context (i.e. “directly satisfying particular consumer needs”). While PO defines a unique entity at any one time, it does not define one through time where products are modified or replaced by retailers and manufacturers.

Two POs (two reference POs, two current POs, or a current and a reference PO) are said to serve the same “basic user purpose” (or basic purpose for short) if they are chosen:

- by, to a large extent, the same categories of consumers;
- who perceive them as similar in spite of, or after accounting for, differences in their characteristics;
- for use in similar situations;
- on the basis of their predominant use and with similar ends in mind. 6

The aim is to employ the criteria of this definition so that ‘basic purpose’ can serve to harmonise the operational identification of NSG&S by product-area across member States.

Emerging products which do not satisfy at least one of criteria (a) to (d) above are deemed to be genuine new products, thus falling outside the scope or the a-priori coverage of the basic reference purpose in question. As basic purposes are thought to be the lowest level of aggregation, they are by nature rather narrow; however they are broader aggregates compared to POs, which latter may be considered as clusters or sets of transactions.


International Working Group on Price Indices - Seventh Meeting 325
4. New (electricity) suppliers

Following COICOP/HICP 04.5.1 “electricity” is a non-durable good, which may be described by means of basic purposes such as “purchase of a kWh by a certain category of households for domestic uses”. The appearance of a new supplier in the marketplace constitutes in fact a new PO which serves the same basic reference purpose.

From the perspective of the NSG&S standard, the new supplier should therefore be treated as a replacement. It should enter the index as soon as the new PO is deemed to be significant (and information becomes available on the number of households accepting the new PO). This might be done with or without quality adjustment (e.g. depending on diverging contract modalities and whether or not the PO is considered to involve a new outlet).

The introduction of the new supplier at a relatively early stage seems particularly important to the extent that its early price development may be systematically different from the price trend of the old supplier. It should be noted that the user-purpose approach actually precludes chaining in order to show no price change (e.g. the introduction of a new basic purpose on the assertion that electricity offered by the new supplier serves a different basic purpose due to diverging contract modalities, such as terms and conditions of supply).

In principle, the same applies generally to ‘utilities’ markets which are in the process of deregulation such as gas or telecom services. The situation may be described as the one above, where suppliers enter “the market” with POs serving existing basic purposes (apart from negligible changes due to design of bills or methods of payment). Significant changes in the quality of the products, if any (e.g. for gas with a different or modified calorific value, telephone calls at quite different modalities), would be treated as replacements and dealt with by means of quality adjustment. Realistically, price indices for deregulated utilities would fall as more households decide to switch to the new supplier.

It is mainly because of the monopolistic (or oligopolistic) character of such markets that for practicability reasons prices are collected centrally. This means that the sampling and replacement procedures in place will hardly detect the emerging new POs of new suppliers in the market. The user-purpose approach suggests a pro-active detection of eligible POs and ad-hoc gathering of relevant information in order that new suppliers enter the current index.

5. E-commerce (internet shopping)

Recent evidence on internet usage suggests that consumers are increasingly purchasing goods and services by means of internet orders. The HICP standard on newly significant goods and services standard hence suggests that MSs should expand their price collection in order to cover internet orders in their HICPs.

The case of internet shopping does not seem actually that different from the catalogue mail or phone ordering as compared for instance to the traditional (supermarket) shopping.

Basically it is not differences in the physical characteristcits of purchased products as such which make these new POs be perceived as different or dissimilar by the relevant categories of consumers. It is rather the functionality of the POs concerned - i.e. a strong “home service” component - which hand in hand with the necessary computing skills and equipment makes the difference.
POs by internet suppliers though arguably serving the same user purposes cannot be said to serve the same basic purpose; although both can be said to be for use in similar situations, and chosen with similar ends in mind certainly consumers perceive them as different on the grounds of obvious differences in their functional characteristics. They serve thus different basic purposes, and do so presumably at diverging prices and diverging contract modalities (such as terms and conditions of supply).

It may well be that in the long-run internet shopping will replace a large portion of the catalogue ordering or traditional shopping. However, in the short-term, it can safely be said that they create (serve) new basic purposes and thus change or modify the existing upper level user-purpose.

It seems therefore appropriate to cover internet shopping by means of new basic purposes, i.e. as a newly significant product. This will require chaining-in additional elementary aggregates in the December following the earliest possible estimation of a reliable expenditure weight. Potential shifts from traditional shopping or catalogue ordering to internet shopping will in the mid-term be reflected in the HICP by means of updating their respective expenditure weights in the index reference period.

6. Mobile telephony

The emergence of the mobile telephony some years ago clearly modified the range of the existing POs and the basic user purposes they used to serve; such as purchase of equipment for fixed telephone communication and the purchase of the related communications services.

In some cases, it took a number of years until it was decided to let mobile telephony (equipment and services) enter the consumer price index. In some EU MSs the decision was first taken on the grounds of the HICP standard on newly significant goods and services.

The reasoning for the inclusion and the treatment of mobile telephony in the HICP under the NSG&S standard goes similarly to the one employed in the context of internet shopping.

Mobile phones and related services must be included in the HICP as soon as they are shown to be significant and reliable expenditure weights can be established. It is on the grounds of the difference in the way the new POs serve the basic telecommunication purposes of consumers that mobile telephony is seen as defining a new user purpose.

Since COICOP/HICP requires in principle the distinction between telephone service and equipment, the HICP must distinguish between at least two new basic purposes; one for the purchase of mobile phones and one (or maybe more) for the purchase of mobile telephony services.

Thus, mobile telephony should be covered in the HICP by means of new basic purposes, i.e. as a newly significant product. This will require chaining-in additional elementary aggregates in the December following the earliest possible estimation of a reliable expenditure weight.

As in the case of internet shopping, in the mid to the long-run mobile telephone may replace a large portion of the expenditure on fixed telephony as a result of gradual substitution. However, under the HICP perspective, the two POs belong to two distinct basic reference
purposes priced in the current index independently from one another. Here again, potential shifts from fixed to mobile telephony will be reflected in the HICP by means of updating their respective expenditure weights in the index reference period.

7. Conclusion

The operation of the standard has shown that MSs are making considerable progress in monitoring consumers' expenditure on newly significant products. Significant parts of consumers' expenditure have been added to the HICP. Compliance monitoring provided some rough indication that there may have been more products which should be considered beyond those currently reported by MSs.

The formulation of the standard seemed interpretable in different ways and this gave scope for further clarification, subject to its strong interaction with replacement and sampling rules. However, the requirement is that the number of elementary aggregates or the number of price representatives to be added to HICPs should be sufficient to represent the diversity of newly significant items and the variation of their price movements.

The NSG&S standard provides an operational and effective device to cope with emerging products and user purposes. The notions of POs and basic purposes seem particularly helpful to the devised procedure. The standard reflects the philosophy of the HICP: operational rules defined by observables rather than vague concepts actually define what is measured as inflation.

Notwithstanding 'subsidiarity' there can be no principled reason for MSs to differ on where they attempt to draw the line between 'replacements' and NSG&S; a distinction having direct effects on the measured inflation. We are seeking a common distinction which must be formulated in terms of operational rules to guide actual practice.

References


Session 6 - E-commerce

Chair: Rósmundur Guðnason, Statistics Iceland

Summary of session

Five papers were presented at this session. Four of them treated different aspects of e-commerce and one looked into the issue of classification harmonisation.

Weight and coverage of e-commerce

First question that rises is how e-commerce should be defined. OECD has since April 2000 endorsed definition of e-commerce stipulating that it is where the order is placed but not the payment or the channel of delivery that determines whether the transaction is e-commerce or not. If it was included in the definition that the payment should be made electronically, it would cancel out twenty five per cent of the transactions in the case of Canada.

Even if the growth of e-commerce has been considerable its share in consumption or expenditures is rather low as is shown in the papers. The expenditure weight is spread over many products and fields of the CPI and lead to many difficult measurement problems that were discussed at the meeting. It is difficult to anticipate how e-commerce will grow in the future. Information from Finkel's and Lowe's papers shows that it is most common for individuals in higher income groups to buy on the net and there does not seem to be any special inducement for the lower income groups to trade more products or higher priced goods over the internet in the future. The issue about the safety of e-commerce on the web was mentioned in Okamoto's paper. That issue includes both the payment and delivery of goods or services. When it is not absolutely safe to trade on the web regarding both payment and delivery it can be seen as a strong trade barrier.

One problem is the fact that it is often difficult to know where the web firm is based and where the good is shipped from. This is connected with the scope of the CPI and the question if spending is domestic or from abroad.

To register the amount of goods and services sold various sources are needed. Information from household budget surveys (HBS) seems to be a very strong candidate as is outlined in Fenwick's and Finkel's papers, even it is often not enough precise at a detailed level. Consumers in HBS are asked specific questions and it is always known who the buyer is. Information from sellers can also be useful but the problem is to know who is buying. The major part of e-commerce is conducted by businesses so there are considerable difficulties in separating between business and private expenditure. The third source is surveys but experimental survey that has been conducted in the UK about the internet use does not show promising results.

1 "An electronic transaction is the sale or purchase of goods or services, whether between businesses, households, individuals, governments, and other public or private organisations, conducted over computer mediated networks. The goods and services are ordered over those networks, but the payments and the ultimate delivery of the goods or service may be conducted on or off-line.”

2 European Commision, Eurostat (July 2002), E-Commerce in Europe, (7).
E-commerce sampling

Some participants stressed the fact that e-commerce price collection would only be problem in case web prices moved differently from prices in general. If the prices changed in similar way it would ease the pressure of collecting prices from the web sites. Okamoto’s paper showed that PCs sold by the internet had comparable prices with models sold at traditional retail outlets, after appropriate quality adjustments. But the extension of this observation to other products or countries is an issue. Fenwick showed that two different strategies could be applied according to price movements of products:

- where price movements of e-commerce are different (compared to traditional comparable products), selection of these products in the sample is recommended. Most of participants thought that a sampling frame of outlets was generally preferable as e-commerce looks more like a new distribution format (like mail order shopping) than a new kind of products;

- where price movements of e-commerce are similar, sampling of these products is less necessary. But the observation of internet prices appears in this situation to be a cost-effective tool to track and monitor prices of products sold in other outlets.

Problems of e-commerce price collection

The e-commerce trade is still very volatile which can influence the stability of price collection. There is also big variance in the goods and services sold through the web as it covers very different goods or services such as books, CDs, flight tickets or home banking.

Another problem is the measurement of transport cost. It is very often difficult to assign transport cost to an individual item or good as it is in some cases connected to the amount of goods bought. If the transport cost can be safely related to the item it would be a preferable method to price the good with the transport cost included. But very often that is not possible and in that case the transport cost should be included in a separate index. As with other price measurement the choice of method should not affect the result as long as the proper weight is assigned for the transport services.

In the case of goods like air fares and home banking services arguments were expressed that a method of using profiles should be preferred. The vast information about different prices available on the web, make work with profiles easier.

The issue of classification

Woolford’s paper was partly about the ongoing work of harmonising bar code system classifications. A unified classification system connected to databases of goods bar codes could be of great use in statistics e.g. for data sampling or product identification. It was argued that this would probably be of most use in PPI calculations. The view was expressed that if the aggregation was at a higher level, the lower level classification did not matter so much as long as the aggregation could be conducted safely.

The meeting was of the opinion that this work should be steered by UN Statistical Division that has played a leading role in the harmonisation of other classification systems.
Recommendations for statistical agencies

Weight and coverage of e-commerce

1. The OECD definition of e-commerce should generally be used and the form of payment for goods and services should not influence the decision of which items are included.

2. A consistent treatment should be sought for e-commerce from abroad sites according to the scope of the CPI and to practical considerations.

3. The main source for weights should be household budget surveys. Information from providers of e-commerce services and surveys about e-commerce can also be useful.

E-commerce sampling

4. In principle all web outlets should be covered in the sample frame for the price collection if the expenditure shares are big enough for each item to be included in the index. In practical selection of e-commerce significant products is recommended when their price movements differ from those of traditional comparable products.

5. Even if the weight shares for respective e-commerce products are low, the price information available on the web could be used by convenience in the regular CPI price surveys if that information is detailed enough and price movements of e-commerce and ordinary outlet products are similar.

Problems of e-commerce price collection

6. All cost connected with e-commerce buying should be included in the prices collected. The transport cost should preferably be included in the price of the good or service if possible. If that is not a possibility, a separate index should be calculated covering the change in transport cost and the proper weight should be used.

7. In the case of e-commerce of some goods and services such as airline tickets and home banking a method using profiles is to be preferred.

The issue of classification

8. The meeting was very interested in the ongoing work in the business world to harmonise item classification for scanned goods.

9. There was a strong support at the meeting for the continuation of this work. It was argued that it should be conducted preferably under the supervision of the UN Statistical Division.
International Working Group on Price Indices - Seventh Meeting
1. Introduction

According to market research information online shopping in the UK even by 1999 accounted for 10 per cent of computer software retail sales, 5 per cent of books and 3 per cent of music. ONS's own Omnibus Survey shows that by 2002 46% of households had access to the Internet either at home or work.

There are three reasons why we would want to separate out purchases over the internet from other purchases:

- For the statistical integrity of the index. Prices and price trends might be different;
- For quality assurance, for example to help check weights;
- Public expectations and perceptions.

Against this background, the ONS is expanding the scope of price collection to include goods sold over the Internet in the calculation of the Retail Prices Index (RPI), and its European equivalent the Harmonised Index of Consumer Prices (HICP). Currently prices for books, CD roms and supermarket goods are included explicitly from Internet sources, and the Internet is used as a convenient source of prices data for outlets where goods can be bought at the same price on the Internet and from traditional stores. As expenditure through this channel becomes more significant, we will wish to collect prices for more goods and introduce appropriate weights in to the index, to ensure that such purchases are included with the correct relative importance.

Preliminary investigations have been undertaken in many of these areas, and results are included in the main part of this paper.

2. The inclusion of internet prices in the RPI: a critique of the current strategy

The approach currently adopted is essentially a pragmatic one that exploits the systematic processes already in place for sample replenishment. These are described in the paragraph that follows. But first there is a matter of definition and coverage.

What is the definition of internet expenditure? Does it require the money transfer for a purchase to be online or just the commitment to purchase or order to be placed on line? Clearly this is not a trivial question:
• There will be those that order on line and pay by post, fax or in person, e.g. on delivery and receipt of an invoice;

• Consistency of approach with other remote forms of shopping, such as by catalogue, also needs to be considered. It can be argued that internet shopping is distinguished from catalogue shopping by the very nature of it being totally "proactive" and "interactive". Thus, unlike catalogue shopping, the purchaser can pull information off the web site and complete a transaction totally within the web site.

• Consistency between weights and prices domains. Online transaction prices might differ from others.

The definition used in ONS' Expenditure and Food Survey is that the customer has to pay online. However, in the case of services there can be a practical problem of distinguishing between purchase and payment. For instance, if the initial purchase of a service-say insurance-is made over the internet, under what circumstances would we classify the renewal of that service as an internet purchase? Would this be if the initial discount given for purchasing the service over the internet continues even if the renewal is done over the telephone or only if the renewal is paid for online? Should we take each transaction in isolation? How in practice would we differentiate online renewal from the renewal of the same service which is initially purchased by other means? Clearly, we could easily end up with a mismatch in the index between weights and prices (the final bullet point above).

How does ONS cover internet shopping in its consumer price indices? It is essentially covered by our standard procedures for keeping the sample of shops and items up-to-date and representative. The updating of the sample of retail outlets whose prices are sampled for the RPI is done systematically on a rotating basis, using a random sample stratified by region and size of shopping population and by outlet type. A quarter of locations are re-sampled each year with complete replacement over a four-year cycle. ONS conduct a full enumeration of retail outlets in sampled areas to ensure that price collection is taken across the full range of retail outlets found in the chosen locations. In addition market research information and data from other ONS enquiries is used to identify new types of outlets and shifts in purchasing patterns including towards use of the internet. The latter is taken on board in the RPI using the following criteria to assess the potential impact on the index:

• **Expenditure threshold.** As a guideline (but not a hard and fast rule) we refer to a draft HICP sampling regulation. This requires that the sampling frame covers any distinguishable subset of outlets that accounts for more than 1 to 2 per cent of total household expenditure.

• **Dissimilarity in price movements.** Clearly, in order to maintain representativeness, it is particularly important to include new outlets types if the price movements of the goods they sell are different from those observed for outlets already included in the RPI sample. This can sometimes be difficult to monitor and the results not necessarily predictable, for instance in a competitive market where lower Internet prices can encourage a traditional retailer to discount. In such circumstances it is possible that the appearance of Internet shopping will have lead to a one off fall in prices in traditional shops that will already have been picked up by the index prior to any direct inclusion of Internet prices. In the UK this applies, for instance, to books where the advent of Amazon has led to more competitive discounting of "best sellers" in traditional shops.
• **Perceptions.** The integrity of the index can be undermined if there is a perception that the index doesn't cover all appropriate retail activity even though exclusion has no significant numerical impact.

• **Conceptual and definitional considerations.** For example it may be considered inappropriate to include Internet shopping where this involves membership of an organisation. This depends on the rules for compilation of the particular index.

• **Feasibility.** For example it is considered impractical to cover auctioned and bartered goods even if we wanted to (this includes Internet bidding).

Our general feeling is that our current review arrangements pick up new forms of shopping such as online via the internet reasonably well and that the guidelines for inclusion are robust in the current retail environment despite the fact that they are guided by pragmatism rather than a rigorous conceptual argument. However, it is arguable whether our current methods will provide a suitably robust approach in the future particularly if internet shopping increases in significance. For example, from a more strategic viewpoint we need to consider:

• Whether internet shopping should be considered as separate “outlets”, in which case, it will need to be fully integrated into the systematic 4-yearly enumeration process described above. Additionally, such an approach will also have an impact on the structure of elementary aggregates and the RPI sample. In particular, consideration will need to be given to whether internet shopping should be treated as a separate stratum for sample selection and elementary aggregation.

• Whether items purchased from the internet should be treated as different items from those purchased from more traditional outlets on the grounds that a different type of service is being provided. In this regard, it is worth asking whether internet purchases are essentially any different, at least in principle, from mail order by catalogue. Certainly in the UK the latter has been a significant part of retailing over a number of years and has recently experienced a resurgence from active participation in this sector by fashionable retailers such as the UK clothing company Next. A number of considerations arise:

  - The increasing strength of this sector may be attributed, at least in part, to the convenience for which shoppers are willing to pay a charge for postage. In circumstances where exactly the same good can also be bought for exactly the same price from a physical outlet then the additional cost of postage and packaging may be considered as the minimum consumer valuation and the price paid for the convenience. It is therefore indicative of the “quality adjustment” that should be applied to the total gross price. It could thus be argued in some circumstances that delivery charges are netted out of the price index as a quality adjustment.

  - If it is accepted that there is very little difference between mail order shopping and purchasing via the internet, conceptually at least, then as a matter of principle a consistent approach should be followed. Alternatively and more pragmatically, at the very least the experience gained from measuring the former should help to provide practical solutions to the measurement issues associated with the latter. It should be noted that in the current RPI design, mail order shopping is treated as an outlet dimension to sampling and is included as a sub-stratum of the stratum for “multiples”.
The latter also includes standard retail chains (where prices are collected either from local outlets or via Head Office) and mail order.

- The distinctions between internet and mail order shopping and also with traditional shopping are fast becoming blurred in that, for example, mail order shopping often provides the facility to order and pay by post, over the telephone or over the internet. In addition, for a significant number of goods and service customers have a choice over how they shop with, for example, many supermarket chains offering purchasing over the internet "for added convenience" as well as purchasing at a physical outlet. Thus a number of retailers offer a variety of modes of purchase for identical goods, again making the distinction less clear. This also has implications for measurement.

Whichever view is taken of the treatment of internet shopping in relation to other forms of shopping, the practical measurement issues share common ground with other types of shopping and index compilation. Thus, to summarise, whether internet shopping is any different from any other form of retailing is arguable and perhaps the main distinction from other forms is the practicality of measurement, i.e. of collecting price observations proportional to sales. But what is, perhaps, lacking is a strategic approach focussing on the longer-term. This is something that the ONS is currently addressing in its research programme and it will be looked at in the context of a wider consideration of other relevant forms of shopping.

This paper, amongst other things, seeks to stimulate debate on the need and formulation of an overall strategy, and specifically whether there is a need, at this stage, to include expenditure via the Internet as a fully disaggregated separate "outlet" or "item", where the expenditure share is the primary issue. In addition, there are a number of more technical areas that would benefit from further analysis and discussion.

3. Issues for further analysis and discussion

The production of price indices for goods bought over the Internet is an issue of growing importance in the calculation of the Retail Prices Index. In 2000 ONS introduced the explicit pricing of books purchased from the Internet. From 2003 it will be pricing some food, alcohol and household consumables from supermarkets ordered over the Internet. However there is growing evidence from ONS’s Expenditure and Food Survey (see later paragraphs for further detail) that this should be extended to other areas, for example music CDs, and with the projected growth in Internet expenditure this is likely to extend more widely.

However, there are many issues to be considered before we can include Internet expenditure in the index for many items, and in a more systematic way. As mentioned earlier, there are two basic strategic choices: either make internet collection conform with more traditional price collection or extend definitions and compilation procedures to encompass the distinct practices associated with Internet retailing. Several issues need to be addressed, both conceptual and practical, in considering the relative benefits of these two alternative approaches. The most important of these are:

- Initial identification of products where internet sales to households have become significant, and sources of sales information for weighting;
• The treatment of delivery charges. Should they be included in prices, or charged separately? How should they be calculated given that there is often a fixed minimum charge?

• Differentiation between purchases for domestic consumption and those for business;

• The treatment of sales from overseas websites, and whether they are within the scope of the RPI/HICP (i.e. does the transaction take place within the usual geographic boundaries associated with the index);

• Whether there is a quality adjustment issue arising from a change in the mode of shopping (e.g. greater convenience);

• The production of a sampling frame of Internet outlets, from which a random sample of outlets can be taken;

• The volatility of the Internet market, and how to deal with situations where internet outlets become unavailable because they close.

4. Main sources of information for the monitoring the levels of Internet purchasing

4.1 High level monitoring of shopping patterns

ONS's Expenditure and Food Survey is a sample survey of around 7,500 households which collects information on household spending, and is used to provide weighting information for the RPI. From April 2001, information has been collected on purchases over the Internet. The first year's results show that four per cent of households ordered goods or services over the internet (compared with 40 per cent of households who had internet access), and this accounted for 0.9 per cent of expenditure. Details are given in table 1.

A more detailed analysis of the data summarised in table 1 shows that spending is frequently concentrated in only a small number of areas within the broad heading. For example, the highest spending was on transport, but with most spent on airfares, rather than on other kinds of transport goods and services. In general the information so far produced gives a very good idea of which sorts of products we should be looking to price as internet purchases. However, with only a small number of households spending on internet purchases, and just one year's data, we need to take care when using the data for weighting purposes. Despite this, the data is indicative of levels and future trends and of those areas that warrant further consideration.

Table 2 shows where ONS currently price Internet purchases, and where the EFS data suggest that there is potential for further collection. Items in bold in the second column have been introduced this year. The majority of these are part of a collection of non-perishable goods priced at four supermarket chains. This collection was prompted by the analysis of the EFS data, which showed that such purchases were not only found in food and non-alcoholic drinks, but also within alcoholic beverages, household goods, recreation and culture and miscellaneous goods. The scope is limited in the first year, comprising 30 items priced in 2 varieties at 4 supermarket chains offering Internet shopping. If the collection is successful we will expand it in future years. Note that the current procedures do not treat these items as a special category, rather they are integrated within the current structure of the RPI.
### Table 1

<table>
<thead>
<tr>
<th>COICOP group</th>
<th>Recording households in sample</th>
<th>Number of households ordering goods/services over internet</th>
<th>Households ordering goods/services over internet as a % of all households</th>
<th>Average weekly internet expenditure for households ordering goods/services over the internet (£)</th>
<th>Internet expenditure as percentage of all expenditure on the commodity or service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and non-alcoholic drinks</td>
<td>7,444</td>
<td>64</td>
<td>0.9</td>
<td>36.80</td>
<td>0.8</td>
</tr>
<tr>
<td>Alcoholic drinks</td>
<td>3,886</td>
<td>29</td>
<td>0.4</td>
<td>10.50</td>
<td>0.7</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>5,282</td>
<td>34</td>
<td>0.5</td>
<td>27.00</td>
<td>0.3</td>
</tr>
<tr>
<td>Household goods and services</td>
<td>7,010</td>
<td>70</td>
<td>0.9</td>
<td>10.60</td>
<td>0.3</td>
</tr>
<tr>
<td>Health</td>
<td>3,817</td>
<td>8</td>
<td>0.1</td>
<td>11.30</td>
<td>0.3</td>
</tr>
<tr>
<td>Transport</td>
<td>6,578</td>
<td>25</td>
<td>0.3</td>
<td>93.30</td>
<td>0.6</td>
</tr>
<tr>
<td>Communication</td>
<td>7,186</td>
<td>8</td>
<td>0.1</td>
<td>40.30</td>
<td>0.4</td>
</tr>
<tr>
<td>Recreation and culture</td>
<td>7,430</td>
<td>186</td>
<td>2.5</td>
<td>30.50</td>
<td>1.4</td>
</tr>
<tr>
<td>Miscellaneous goods and services</td>
<td>7,315</td>
<td>80</td>
<td>1.1</td>
<td>6.00</td>
<td>0.2</td>
</tr>
<tr>
<td>All expenditure groups</td>
<td>7,473</td>
<td>299</td>
<td>4</td>
<td>45.50</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>COICOP group</th>
<th>PRODUCTS BEING PRICED</th>
<th>POTENTIAL FOR INTERNET PRICING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and non-alcoholic drinks</td>
<td>non-perishables at supermarkets offering internet shopping</td>
<td>full range</td>
</tr>
<tr>
<td>Alcoholic drinks</td>
<td>supermarkets offering internet shopping</td>
<td>full range</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>none</td>
<td>full range (though some purchases are of specialised goods such as motorcycle kit and replica football shirts)</td>
</tr>
<tr>
<td>Household goods and services</td>
<td>household consumables at supermarkets offering internet shopping</td>
<td>mainly household consumables plus some furnishings</td>
</tr>
<tr>
<td>Health</td>
<td>none</td>
<td>small spend on OTC medicines &amp; similar</td>
</tr>
<tr>
<td>Transport</td>
<td>air fares</td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Recreation and culture</td>
<td>books, computers &amp; computer parts &amp; accessories, computer software &amp; games</td>
<td>CDs, DVDs, toys, holidays</td>
</tr>
<tr>
<td>Miscellaneous goods and services</td>
<td>toiletries</td>
<td>insurance, financial services</td>
</tr>
</tbody>
</table>

338 International Working Group on Price Indices - Seventh Meeting
4.2 Data for weights

In order to correctly include Internet expenditure in the RPI it is clearly important to know, for each included item, what proportion of sales take place over the Internet. This will allow us to either correctly weight price indices for goods bought over the internet together with other indices (i.e. introduce a specific elementary aggregate), or collect the correct number of price quotes to include them in the index without explicit weighting. The data available from the EFS is helpful in this respect, but as it covers only a small number of households spending on internet purchases, and just one year’s data, we need to take care in using the data for weighting purposes. This is not a major issue at this stage but could be in the future if the range of Internet shopping increases significantly and the level of accuracy of the EFS remains insufficient. In these circumstances we may need to consider introducing a supplementary sample. But this could be costly due to the need to filter out those that don’t buy online and there would also be a respondent burden issue. Alternatively we could seek other more imaginative approaches to the issue including the use of administrative and accounting data from Internet and associated companies. The experience to date on the latter, based on a joint project with a major credit card company, has not been encouraging. Further details are given towards the end of this paper. The data issues to be confronted are in many ways similar to scanner data but without the convenience of an active group of suppliers who collate and quality assure the primary input.

4.3 Purchasing patterns

In addition to providing information on which products are purchased on the Internet (and the associated expenditure), household budget surveys are also capable in principle of providing more detailed information on purchasing patterns. For instance, they can show what sort of goods were bought at the same time, and how much was spent- information which can be useful in determining delivery charges (see later). In addition they have the advantage over most other sources of generating valuable information on household profiles and can thus provide the necessary data to identify the purchasing habits of households covered by a consumer price index. For example, the UK RPI excludes pensioner households that derive at least three-quarters of their total income from state benefits and pensions and high-income households, defined as those households whose total household income lies within the top 4% of all households. Such a differentiation would not normally be available from an administrative or accounting source.

4.4 Supplementary sources: market research data

Where possible we have tried to verify and supplement EFS data with market research data, though the latter usually includes business as well as household spending, tends to be collated from numerous incompatible sources and can be expensive to purchase. The limited information we have shows similar patterns of purchases. However, one area identified by market research data, but not in the EFS, are internet car purchases, where prices are very different. We will be investigating this further in 2003. The latter is a typical example of an infrequent but high value purchase on which a household budget survey will generally not be capable of providing reliable information. Market research information from retailers is likely to provide a more reliable source.

5. Purchasing services over the Internet: pricing issues

In the case of EFS data collected on services purchased over the internet we have found it difficult to distinguish clearly between true purchases and the payment of a service through
internet banking. In terms of pricing, the issue may appear to be more clear-cut in that what we are concerned to identify is products where different prices are offered for internet purchase. However, this is simplistic without giving recognition to the fact that internet shopping provides a different form of service and may therefore imply a quality difference.

5.1 Delivery Charges

The issue of delivery charges is partly practical (not all sites explicitly state delivery price), and partly conceptual. With the help of Eurostat draft guidelines, we are able to make a broad statement:

"Delivery price should be regarded as part of the price where the good cannot be consumed unless delivery takes place."

Putting the quality change issue to one side this implies that for Internet purchases, where clearly you can not make use of the product unless it is delivered to the home, delivery charges should be included as part of the price of the good. But behind such a statement lies many differences of opinion and interpretation. This can be seen from extracts of the debate that was held and from the fact that there is no regulation on the issue:

"The issue of delivery charges and their assignment to COICOP/HICP concerns the issue of weights as well as the measurement of prices in the HICP. There are some COICOP groups where the delivery charges are explicitly assigned to the same COICOP groups as the delivered products themselves, e.g. delivery of major household appliances (COICOP 05.3.1/2).

COICOP/HICP follows the NA principle only in those cases where COICOP refers explicitly to delivery and installation charges (COICOP 05.3.1/2). Concluding from this explanation the price representatives to be followed in this COICOP group should include delivery and installation if consumers usually buy the products including these services (such as washing machines).

The HICP does not to use the full NA wording. The part "including any transport charges paid separately by the purchaser to take delivery" has left out of the definition of price on purpose. The decision was based on the argument that further consideration was needed as to whether delivery charges should be generally covered and to which COICOP group they should be assigned.

For those COICOP groups where COICOP does not explicitly give advice where the delivery charges should be assigned, illustrated examples have been produced to the assignment of delivery charges by applying the principle established for the NA. Following this principle the delivery services should be assigned to the same COICOP group as the delivered products. This approach was favoured by some but not all Member States because, e.g. it was consistent with the weights which were based on NA data. The UK added that where the consumer had no choice whether the product was delivered or not (like ordering books and CDs from amazon.com), the price to be followed should include the delivery charge. Furthermore, the question of who arranges for the delivery, the consumer or the retailer, might be worth considering in this respect.

For Germany it was preferable to assign delivery charges to COICOP group 07.3.6 (08.1.0) if

a) the invoice that accompanied the delivered product showed the price of the product and separately the delivery charges; and

b) the delivery charge formed a large part of the total price to be paid for the full service (product plus delivery).

Furthermore, Germany suggested applying the following general principle: If the delivery charge is a fixed charge regardless of the number of items bought, then it should be assigned to 07.3.6 (08.1.0). If, in contrast, the delivery charge was proportional to the product itself, e.g. a delivery charge of € y per box of water, then the delivery charge should be assigned to the COICOP group the delivered product was assigned to."
Eurostat’s opinion at the time was as follows:

"a) Concerning the question whether the HICP should take account of delivery charges or not:

Firstly, delivery charges form part of consumption expenditure and should be covered by the HICP. Secondly, representative items should take account of delivery services if the products are usually bought including delivery. For those products where delivery is currently gaining importance, like the delivery of books and CDs, it might be worth adding price representatives that account for delivery.

b) Concerning the question how the delivery charges should be covered and which COICOP groups they should be assigned to:

As a first criterion it seems to be reasonable to ask the question whether delivery is optional or obligatory.

As a general rule, delivery charges should be covered together with the delivered product(s) where delivery is obligatory.

Where the consumer has the possibility to arrange for delivery himself (optional delivery) the delivery charges should be covered in COICOP 07.3.6 (08.1.0 respectively) depending on the underlying service of delivery (probably depending on the size and weight of the delivered product).

From a practical point of view, there would not be any problem if a consumer ordered only one product. Where a consumer orders several products belonging to several COICOP groups, one would have to apportion the delivery charge across COICOP headings. The discussion in Athens showed that most MSs were not in favour of this solution.

In order to provide a practical solution also in this case, a second criterion should be considered, i.e. whether the delivery charge is paid as a proportional charge (per unit, percentage of the price of the delivered product) or a fixed charge regardless of the number of items bought.

If the delivery charge is to be paid as a proportional charge, it should be assigned to the same COICOP group as the delivered product. The same applies to the delivery of several products if the delivery charge is proportional, in which case the delivery charge should be split accordingly.

If the delivery charge is a fixed charge regardless of the number of items bought, it should be assigned to either COICOP 07.3.6 or COICOP 08.1.0. This rule should always be applied, also when the delivery charge forms a large part of the total price to be paid for the full service (product plus delivery)."

The annex gives a draft Eurostat guideline that has so far not been taken forward.

However, this is not the full extent of the issue. Even with a general principle in place there are practical issues that must be considered. One question is how to deal with sites that do not set delivery charges per item, but rather either set charges as a fixed fee for delivery up to a set number of items, and then an added fee per item, or have a scale for delivery, where the unit cost changes depending on the number ordered. Both of these cause problems for price collection - most specifically in deciding whether to cost delivery on a single item or as an average across several. For example if a bookseller charges £7.50 for the delivery of up to 3 books, then do we treat the delivery charge as £7.50 (the cost if you ordered 1 book), or £2.50 (assuming that consumers would only pay the minimum delivery charge per book). Clearly a major challenge is to obtain detailed information from retailers (or customers) on the make up of individual orders and average charges for delivery. This type of information is not readily forthcoming in the UK. The samples used for the Expenditure and Food Survey is generally too small for the purpose and Internet companies are reluctant to supply such information.

One reason for including delivery charges with the item to which they refer is to ensure that the price is on a comparable basis to that for the same good purchased at a conventional outlet. This is clearly valid for a large durable good which is normally delivered free of charge from a conventional outlet. But consider the case of a household’s weekly shop at a
supermarket. The "delivery" charges associated with this kind of purchase – the cost of using a motor vehicle, or possibly taxi fares, are included elsewhere in the index, not as part of the price of the items purchased. So comparability would not be aided by allocating the delivery charge across the items purchased.

Furthermore, in cases where the consumer purchases a heterogeneous set of items simultaneously from one supplier (common with purchases of book/CDs/computer software) a decision to add the delivery charge to the item price, can cause problems. We would first need to arrive at an average basket of items, and then decide how to apportion the delivery charge (i.e. per item, or in proportion to value, or by weight). In these cases a more appropriate approach might be to produce a separate index for delivery charges, based on an average basket. Information from household budget surveys about purchasing patterns may be useful for determining how delivery charges are applied in practice, though there are survey design and response issues to be overcome.

In summary, there is no definitive answer to the question of delivery charges. At present ONS thinking, based on pragmatism and an earlier discussion held in the context of the HICP, is as follows:

- Include internet delivery charge in item price where this is also the case for items purchased through other retail outlets (e.g. furniture) and for any other large items where delivery from conventional outlets is normally "free";

- Also include these charges for single purchases these cannot be consumed unless delivered by the retailer or their agent;

- For bulk internet purchases of a range of goods (e.g. supermarket purchases) construct separate indices for delivery charges based on typical baskets of goods. This will allow changes in delivery charge regimes to be more easily accommodated.

5.2 Differentiation between prices for consumers and for businesses

Clearly the Retail Prices Index is just that, an index based on retail prices. Therefore any prices that are only available to businesses do not fall within the scope of the index, and should be excluded from the calculation. So the question is how to identify these prices so that they can be excluded.

There is no definitive, or easy, answer to this, however we can produce several indicators that a web-site is only offering prices to businesses, though these will need to be applied with appropriate knowledge of the market. Factors considered so far are:

- Clear statements that the web-site is trade only;

- Prices that are quoted for large numbers of a good (e.g. Processor chip prices of £100 per unit, if 1000 or more purchased);

- Prices quoted without VAT;

- Web-sites for known wholesalers, who are known to exclude the public.
Of these we would need to be most careful in applying the restriction on prices quoted without VAT. For some goods, such as Personal Computers, this is standard practice, and indicates nothing more than the prices are available for both retail and small businesses.

5.3 Treatment of overseas websites

The treatment of overseas websites is a complicated issue, not least because of the difficulty in identifying whether a website is UK or based overseas, and whether the transaction is deemed to take place there, at the purchaser’s home or elsewhere. Originally you would be able to identify the source of a website in one of two ways, either the website address (.uk suffix means in the UK), or from the currency in which transactions take place. However this has changed, and neither of these can definitively be used to determine the location of a web company.

In the first instance website names (or addresses) can be sold, and are not regulated by geographic location. The largest example of this is the .tv suffix - which is popular with any company who has a relationship with the visual media. This is nominally the suffix related to the Island of Tuvalu, however the addresses have been sold world wide - with several companies in the UK using this suffix. Therefore identification purely on this suffix will mean that we exclude UK based websites (and potentially include overseas ones).

In the second case companies do not always restrict sales to take place in the local currency. Websites frequently ask what currency you wish to pay in, and use an on-line currency converter to produce worldwide prices from the local currency price. As these are normally rounded to produce "normal looking" prices the original currency cannot be discerned. Therefore the use of this information to establish location of a website is unreliable. Even a combination of these two sources of information does little to improve the situation, as the companies most likely to buy other Countries' web addresses, are also those most likely to be using multiple currency formats.

This leaves us in a position where it is difficult to identify whether all websites are UK or overseas. In this case we must ask the question whether this makes much difference to the price indices, or if it does, whether the definition of prices should be restricted to UK websites only or allowed to include others. The question is perhaps more pertinent to the Harmonised Index of Consumer Prices (HICP) where consistency in geographical coverage is important for the calculation of the Monetary Union HICP from the figures for individual members states. But again this is unlikely to be an issue of great consequence at current levels of online retailing.

The easy (and perhaps pragmatic) solution would be to say that it is not necessary to exclude overseas websites. This can be seen as a reasonable solution for several reasons:

• The prices are available to consumers in the UK (unlike other prices in other Countries), therefore are applicable to the UK market;

• It could be argued that the transaction took place at the purchaser’s home;

• These prices are available world wide, and so would be likely to influence sales specifically in the UK, therefore price movements are unlikely to be significantly different from UK websites;
• There is no automatic link between the site of a website, and where goods are shipped. As a company can be regarded as from the UK if the shipping is done from within the UK, the overseas distinction becomes even more difficult to define.

All of these factors suggest including in the sampling frame all websites that are either clearly UK, or for which there is a possibility that they could be UK based.

5.4 Choice of sampling frame and the effect of volatility in the market

The final two areas are closely related, and so are being considered together. A major concern in the production of an Internet price index is to obtain price quotes from a representative sample of outlets. For traditional stores this is relatively straightforward, as it is possible to visit shops to determine what they actually sell (as only those goods that people purchase will be on display), and there is only a finite area to search within a given location. For the Internet this is different in several ways:

• The potential number of outlets for each good is extremely large, and will range from a small home based company selling a small number of a very limited set of items, to large supermarket chains selling many items;

• It is not easy to determine sales figures, or identify which goods are most important to the outlet;

• It is not possible to know whether a good is stock, or being advertised in the anticipation of stock becoming available - this is important as we can only price a good in the RPI if it is actually available in the store or ordered from known stock;

• There is, currently, a significant possibility of outlets disappearing without warning, causing major problems of continuity in the RPI.

These factors make the identification of a relevant sampling frame more problematical than for traditional outlets, and will lead to the need for innovative solutions.

Work on this area is not as advanced as for those listed above. A start has been made on linking the collection with the traditional selection, in that the large stores represented in the main selection, and known to have a web-outlet (e.g. Tesco, Sainsbury), would always be included in the sample. However, it is less easy to determine how to choose from web-only retailers (e.g. Amazon, Jungle). For the goods already included in the RPI (books, non-perishables at supermarkets), this is done judgementally, using known web-retailers, and perhaps this will continue to be the best solution in the short term. However, as internet sales become more important, it will also be more important to ensure some form of random sampling in order to maintain the representativity of the index.

Concerns also extend to the current volatility in web-retailers. At the moment many sellers via the internet do not make profits, and this leads to a situation where they are always in danger of collapse, and ceasing trading. We have already experienced one example in the UK where a major online retailer has been introduced into the RPI calculation only to disappear some months later as a result of bankruptcy. We need to consider whether the effect of this
can be minimised by controlling initial sampling, and produce procedures to deal with the situation when it does arise. We will also need to consider the mechanisms by which the sample is updated.

6. Credit Card Surveys: a workable solution?

As part of its methodological development programme, ONS assisted in a pilot project on Internet shopping led by a major credit card company in partnership with a market research company. From June 2001 online diaries were kept by a sample of about 800 UK internet users recording detailed online purchasing behaviour including cost, nature of item, site purchased from, delivery cost and method of payment. This was designed to provide, at the very least, a useful indication of Internet sales volumes but its ultimate purpose was to compile an “E-tail Price Index” for internet purchases gathering data on online purchasers and the value and volume of the goods they obtain online.

6.1 Methodology

The survey consisted of a number of stages:

- Recruitment of respondent panel

A number of potential methods were investigated:

- Recontacting respondents from a current Internet User Profile Survey (IUPS). This consists of a telephone survey of 1600 internet users every six months on internet behaviour and associated attitudes. Those that indicate that would be willing to be re-contacted for further projects would provide a contact database and allow ready access respondents for whom data on shopping frequency and behaviour is already available. However, this source was discounted for two reasons: the small sample size and the six-month timelag. The latter can be problematic and an area than can be fast changing. 2001 results from this survey, based on recollection rather than on the keeping of a diary, suggested that very approximately 2% of the estimated £650 billion total retail spend in the UK is online.

- Utilising an existing online panel where email addresses are gathered on an ongoing basis from respondents completing a survey known as Telebus (telephone omnibus survey). All these respondents will have agreed to be contacted via e-mail for the purpose of completing online surveys. The sample has the advantage both of size (11,000 respondents) and of being demographically representative of internet users but again suffers from out-of-dateness with some respondents having been recruited over a year ago.

- Using pre-screened contacts from a list broker such as Experian, who compile targeted lists from the results of “lifestyle” surveys. Such sources of more often than not less up-to-date than the market research company’s own survey lists. It was also felt not to be so cost-effective, perhaps requiring lists from more than one source to produce a representative and large enough survey.

- Random-digit-dialling (RDD) where people are contacted from a random list of telephone numbers and asked if they would be willing to take part in research on internet access and purchasing behaviour. Experience suggests that only about 40% of numbers contacted are “live” residential households.
In the event RDD was used on the basis that it was most likely to generate a truly random of online shoppers.

- **The telephone interview**
  During the initial “recruitment” telephone interview respondents were asked:
  - Whether they use the internet
  - Whether they *ever* purchase online
  - Details of past online shopping experiences.
  - Expectations of future shopping experiences.
  - Demographic details.

  The recruitment targets for the “panel” of online shoppers were:
  - 400 who had purchased online in the past three months.
  - 50 who had purchased online longer ago.
  - 50 who had not purchased online but intended to do so in the next three months.

  It was recognised that this would over-represent heavier online purchasers but it was also important to maximise the volume of sales data gathered. The results would be re-weighted at the analysis stage to the true incidence in the target population using information from the previously mentioned IUPS. Quotas were also set for age, gender and working status.

- **The Diary**
  Respondents were asked to complete online diaries- given that they would all have online access this seemed to be the most sensible and cost-effective approach for everybody including the respondents themselves. It also allowed real-time data monitoring by market research staff and also respondents to monitor their own responses (allowing them to quality assure their inputs more effectively). In addition as the diary is web-based, respondents were able to record purchases from whichever terminal the purchase was made, e.g. at home or at the office. The cost of data cleaning was much reduced by the online approach e.g. querying online in real time odd responses. Incentive payments were also made.

  A copy of the diary page is given below.

6.2 **Evaluation of pilot**

Despite the hard efforts of all involved the response rate was very disappointing. The assumption at the planning stage was that 35% of recruited respondents would actually proceed to the diary completion stage in the first month. In the event, whist the data collected was of good quality, the response rate was much lower than expected and the attrition rate high in the subsequent months. In consequence, the survey did not proceed past the pilot stage.

The question arises, whether the survey design could have been improved to produce satisfactory data. The alternative option of using the Telebus Survey was considered but in the event this was not pursued.
7. Other options

Obtaining price and sales data direct from online retailers is the main alternative option that the ONS has not pursued. This may well provide a partial solution in the longer-term or at least help to plug some of the current data gaps, but there are a number of measurement problems to be resolved even if companies agreed to co-operate in principle. These include those confronted already in our attempts to get sales data, such as the difficult of finding a reliable sampling frame and distinguishing sales to private households from sales to other commercial companies. Indeed, there may be a lot to be learnt from experiences gained in the use of scanner data. Both are administrative sources but there doesn’t appear at present to be a company that is actively engaged in bringing together into one database information on online retail purchasing.

8. Conclusions

The ONS has been expanding the scope of price collection to incorporate into the Retail Prices Index (and the Harmonised Index of Consumer Prices) goods and services sold on the Internet. A strategy is in place for identifying which items should be included. Sales information is obtained from the Expenditure and Food Survey, supplemented by other market research information, and prices are obtained from the corresponding websites. But there are a number of conceptual and data collection issues that need to be resolved and that would benefit from the sharing of national experiences. It is also an area that is lacking internationally agreed guidelines but that will be increasing in importance and greater use is made of online sales in the future.
### Annex: Eurostat proposals on the coverage of delivery charges

<table>
<thead>
<tr>
<th>Purchase of</th>
<th>Delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual product</td>
<td>Same COICOP as delivered product(s)</td>
</tr>
<tr>
<td></td>
<td>COICOP 07.3.6 or 08.1.0 Weighting information is needed.</td>
</tr>
<tr>
<td>Several products, single COICOP group</td>
<td>Apportion delivery charge to the COICOP groups the delivered products belong to</td>
</tr>
<tr>
<td>Several products, several COICOP groups</td>
<td></td>
</tr>
</tbody>
</table>

**For information: Extract from Commission Regulation (EC) No 1749/1999**

- **COICOP 07.3.6** Other purchased transport services (S)
  - funicular, cable-car and chair-lift transport,
  - removal and storage services,
  - services of porters and left-luggage and luggage-forwarding offices,
  - travel agents' commissions, if separately priced.

  *Excludes*: cable-car and chair-lift transport at ski resorts and holiday centres (09.4.1).

- **COICOP 08.1.0** Postal services (S)
  - payments for the delivery of letters, postcards and parcels,
  - private mail and parcel delivery.

  *Includes*: all purchases of new postage stamps, pre-franked postcards and aero grams.

  *Excludes*: purchase of used or cancelled postage stamps (09.3.1); financial services of post offices (12.6.2).
E-Commerce in the Israeli CPI

Yoel Finkel and Merav Yiftach

Central Bureau of Statistics, Israel

The authors would like to thank Ms. Irit Mishali, Ms. Yafit Alfandari and Ms. Inbar Sabag from ICBS for their part in the data collection and editing processes. All errors are sole responsibility of the authors.

1. Introduction

In January 2003, the Central Bureau of Statistics in Israel (ICBS) updated the base period of the Consumer Price Index (CPI) to the annual average of 2002 = 100.0 points. The former CPI, computed during January 2001-December 2002, had the base period of the annual average of 2000 = 100.0 points.

The consumption weights for the basket of goods and services were derived from the annual Household Expenditure Surveys (HES) conducted by ICBS in 2000-2001. These surveys included questions relating to household consumption by electronic commerce and of Internet services.

ICBS has utilized these data and findings from other sources (media, journals, academic research, private firm surveys) to launch a project of pilot indices for specific goods and services consumed through electronic commerce and actual measurement of improved indices for Internet services in the current CPI.

In this paper we discuss the issues relating to the inclusion of E-commerce in the Israeli CPI, presenting preliminary results of the project. In the next section we define the basic objectives and questions that we set out at the beginning for constructing a CPI for E-commerce; this will be followed by sections on expenditure data from private sources and HES and actual measurement of pilot indices for the CPI. Afterwards we raise some additional questions concerning the measurement of Internet services and sum up by recalling on the main results, their limitations and some thoughts for future work.

2. E-commerce project in the Israeli CPI – the issues

The objective of the project was to construct a pilot series of indices for goods and services purchased via E-commerce. This includes analysis of consumer expenditure in the base period of the index (section 3), collection of prices and compilation of price indices for these goods and services and comparison of these indices to the "traditional" price indices from the actual CPI (section 4).
The goods and services to be included in the E-commerce index may be of the following nature:

- Purchased from internet sites that are "extensions" of the traditional shops (food from supermarket chains, items from department store sites, etc.).

- Unique to E-commerce (the goods may be similar to those of traditional outlets but may be sold exclusively on internet sites).

- Consumption patterns of services that did not exist in the pre-internet period (consuming services in different fashion than the regular index – like financial services, banking, stock markets or consumption of public services via internet and consuming goods and services by on-line auctions).

- Consumption of Internet provided services (connection to Internet and ISP added value services).

The two major issues that we were exploring included: analysis of price trends for goods and services provided by E-commerce (whether they differ from the price trends for the traditional outlets); and whether pricing methods for these deviated from the "normal" price schemes (like auctions). A positive answer for either or both of these questions would necessitate inclusion in the current CPI.¹

The general methodology for the project was (1) gathering of data on E-commerce consumption through media, research articles, surveys (public and private), IT related journals and other sources to determine weights (2) building an internet site "register" for current price collection based on data from major credit card companies or IT journals (3) collection of prices over time for regular and unique goods and services (4) compilation of price indices and comparison with the actual CPI.

Many issues had to be resolved in the process: which goods and services to choose, how to collect the same items on a monthly basis and from which sites, how to handle special pricing schemes, etc.

In the following sections we present findings of the project and answers to some of the questions raised. While we have had only partial success in dealing with the more difficult issues, the results indicate that E-commerce is an issue that can not be disregarded, at least in the Israeli CPI.

3. Consumer Expenditure through E-commerce

Stage one of the CPI E-commerce project was to obtain data on market shares of E-commerce. This was to give insight on areas where E-commerce would be considered to be a

¹ Since e-commerce is a fairly new phenomenon, many statistical agencies assume that consumption weights are negligible, and even if they were considerable then present methods for measuring the price changes of goods and services serve as a good approximate for price changes in the e-commerce domain.
significant part\textsuperscript{2} of the consumption patterns. These goods and services would then be labeled as areas of concern for the index makers.

ICBS turned to two major sources to collect this data, (1) utilize data provided by private research firms, internet sites, credit card companies, IT journals, etc. (2) Annual Household Expenditure Surveys conducted by ICBS.

The advantages of private sector sources are:

- They may be conducted by companies with more professional experience than the statistical agencies in the E-commerce domain.
- There are many of them and we may cross-check and verify results.
- They may be based on many more observations than current surveys conducted by statistical agencies in other areas.
- They are usually for free and national statistical agencies are low on funds.

The major disadvantage is that the motivation for these market research data is self-induced promotion by the private companies. Therefore the statistical agency has no control over the methodology or the results. In addition, the classifications of goods and services are not consistent with an international standard (like COICOP). Whereas a survey conducted by the national statistical agency is objective, of sound methodology and according to scientific and international standards.

3.1 Market Shares from private sources

The following is a description of the Israeli E-commerce market in 2002 based on multiple sources from the private sector:

- 41% of the population has internet access either at work or at home.
- 18% (with access) have purchased goods or services, at least once, through the internet.
- 39% of those connected to internet are "light surfers" (up to 4 hours per week); 31% are "medium surfers" (4 to 10 hours) and 30% are "heavy surfers" (at least ten hours per week).
- Site operators report large increases in sales in 2002 (between 40%-80%) with this trend escalating in the first quarter of 2003.
- 70% of revenues are from auction sales and many operators are moving from conventional selling methods to auction sales.
- 66% of auction purchases are by men (50% in "regular" e-commerce purchases).

\textsuperscript{2} This includes traditional goods and services which now may be significantly purchased via E-commerce or the consumption of internet as a new service.
- Consumption patterns of e-commerce lean towards specific days of the week and specific hours during the day (working hours or 10-11 o'clock at night).

- Prices of goods and services purchased through auctions are usually cheaper than "regular" prices. Closing prices of auctions are similar over time, indicating a "mature" market.

- Popularity of auctions leads manufacturers and wholesalers to supply these sites with unique brands. This causes difficulties in comparing prices with traditional outlets.

Figure 1 below presents the distribution of purchases at E-commerce sites:

![Distribution of E-commerce expenditure](image)

It should be stated that while larger quantities of books and smaller household items may be purchased over the internet, steeper prices of electronic equipment (refrigerators, washing machines, freezers, televisions, etc.) and computers lead to a much larger percent of the expenditure. In addition, the more expensive goods are mostly purchased through auction sales.

3.2 Market Shares from Household Expenditure Surveys (HES)

An alternative and more "scientifically respectable" source for consumption of E-commerce are the Household Expenditure Surveys, conducted annually by ICBS. HES were first performed in the early 1950's; until 1997 they took place approximately once every five years. Since 1997, ICBS has conducted the survey on an annual basis, among the total population of households. The survey aims to obtain data on the components of household budgets, as well as additional data that characterize various aspects of the living standard of households, such as consumption patterns, leisure activities and entertainment, level and composition of nutrition, level and composition of income and housing conditions. In addition, the survey is

---

3 For a more extensive description of survey methodology and findings see Household Expenditure Survey 2001, General Summary, Central Bureau of Statistics Israel, Publication #1201, April 2003.
used for market research, for construction of models to predict consumer behavior, for research on the effect of taxes among the various population groups, etc. One of the most important uses of the survey is to determine weights for the consumption basket of the CPI. As of 1997, the survey population includes 94% of the urban and non-urban household population. The investigation unit is the household, i.e., a group of people living in the same dwelling most days of the week with a shared budget for food expenditures.

3.2.1 Collection of survey data in the HES

Data were collected from each household in an integrated fashion, in the following ways:

- A questionnaire on household structure – filled out by the interviewer, providing basic demographic and economic data on each member of the household.

- A bi-weekly diary – in which the household independently records each member's daily expenditures over a period of two weeks.

- A questionnaire on larger expenditures and on income – filled out by the interviewer on the basis of household reporting, related to the three month period preceding the interview date. Of the 7,509 dwellings sampled in 2001 (results are similar for 2002, not yet published), 685 (9.1%) should not have been investigated (not belonging to survey population). The remaining 6,824 dwellings were inhabited by 6,938 households. 5,787 households (83.4%) participated in the final survey estimates.

Estimates from the bi-weekly surveys and quarterly questionnaires are "inflated" into yearly expenditures and divided into monthly expenditure estimates.

3.2.2 Usage of Internet services, HES 1997-2001

Percentage of households that used internet services grew from 4.6 percent in 1997 to 22.5% in 2001.

Table 1: Consumption of Internet services 1997-2001

<table>
<thead>
<tr>
<th>Year</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet Services</td>
<td>4.6%</td>
<td>8.2%</td>
<td>11.9%</td>
<td>19.8%</td>
<td>22.5%</td>
</tr>
</tbody>
</table>

In 2001, the percentage of households that used the internet (had regular access) in the highest income decile$^4$ was 49.2% compared to only 2.5% for the lowest decile.

---

$^4$ Household population divided into ten equal parts according to some variable. In this case, the income deciles are divided according to net income per standard person – a household’s total current gross income (including income in-kind from housing and motor vehicle consumption) after deduction of compulsory payments (income tax, social security and health taxes), divided by the number of standard persons in the household.
Table 2: Consumption of internet services by income deciles, 2001

<table>
<thead>
<tr>
<th>Deciles</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Services</td>
<td>2.5</td>
<td>9.4</td>
<td>18.8</td>
<td>25.3</td>
<td>40.9</td>
<td>49.2</td>
</tr>
</tbody>
</table>

According to these data, potential of E-commerce consumption has grown at high rates over the previous years and is a common phenomenon, especially for the higher income groups in the population.

3.2.3 Expenditure on E-Commerce, HES 2002

Data from private sector sources (section 3.1 above) indicated that 2002 was a "breakthrough" year for E-commerce sites with revenue growth ranging from 40-80% for several operators. Therefore, much effort was put into extracting preliminary E-commerce expenditure data from HES 2002 (yet to be published).

Household expenditure on E-commerce was 2.04% of total household expenditure in 2002. E-commerce expenditure on questionnaire items (the "expensive" items like electronic and durable goods) was 2.48% and expenditure on diary items (the daily, more frequent expenditures, like food) through E-commerce was less than a 1/2% of total diary expenditure.

Table 3: Distribution of expenditure in diary and questionnaire, 2002

<table>
<thead>
<tr>
<th>Percentages</th>
<th>Total for sub-groups</th>
<th>E-commerce</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diary expenditures</td>
<td>21.6%</td>
<td>0.43%</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>78.4%</td>
<td>2.48%</td>
</tr>
<tr>
<td>Total expenditure</td>
<td>100%</td>
<td>2.04%</td>
</tr>
</tbody>
</table>

The expenditure on E-commerce from HES data is presented in two ways (1) expenditure on items through E-commerce, out of total E-commerce consumption and (2) items in which consumption of them through E-commerce is substantial. These can be divided into diary expenditures and questionnaire expenditures. It should be stated that many of the goods and services may be reported in both methods; however, due to the preliminary nature of the data, they have yet to be aggregated. Another advantage (or disadvantage) of the present data is the high resolution of the estimates as they have been taken directly from the HES database, before aggregation into upper-level consumption groups.

5 Expenditure on E-commerce as reported by households. In the survey, households are required to state place of purchase, both in diary and questionnaire. Purchases via catalogues, television, duty-free, etc. are reported separately by households.
Figure 2 below presents the expenditure on items through E-commerce out of total E-commerce questionnaire expenditure:

![Graph showing expenditure on various items through E-commerce.]  

Figure 3 below presents the expenditure on items through E-commerce out of total E-commerce diary expenditure:

![Graph showing expenditure on various diary items.]  

These data are consistent with the private research firm sources that reported higher distribution of expenditure (out of total E-commerce expenditure) on electronic equipment, computers and books. The inclusion of books from the diary expenditures is also consistent with the findings in section 3.1 (figure 1) above.

Now we present additional findings from HES on E-commerce consumption of specific items, out of the total expenditure for that item. In other words, although total expenditure on E-commerce may be a fragment of total expenditure (which seems not to be the case so far), expenditure on specific items may be significant and may negate the possibility of estimating the price changes through "traditional" outlets alone.
Figure 4 below presents items with significant expenditure through E-commerce (out of total expenditure for that item) from the HES questionnaires:

As can be seen in figure 4 above and figure 5 below, electronic equipment and computer peripherals are amongst the items that are purchased, with a relatively high percentage, by E-commerce.

In Figure 5 below we present items with significant expenditure through E-commerce (out of total expenditure for that item) from the HES diaries:

We may now establish that market shares for E-commerce are significant for electronic equipment, computers and computer peripherals, books and other household items. These are evident from private research firm surveys, data from internet service providers and especially from HES conducted by ICBS. Based on these data, we can assign relative consumption weights for E-commerce in those items with a significant portion of E-commerce. However,
are we able to construct a workable sample of outlets and prices to compute price changes over time? These will be determined in the following section.

4. E-commerce – price measurement

Anticipating an exponential increase in E-commerce expenditure, ICBS began price collection of several internet based goods and services in January 2001. During the 2 year period of January 2001-December 2002, over 1,700 prices were collected from 97 Israeli based internet sites (extracted from private firm surveys and IT journals), covering 72 consumption sub-groups in the CPI. The number of prices collected from these sites varied between single prices for certain sites and between 50-200 prices in others. Five of the 97 sites (books and CD's) accounted for more than 40% of the prices. Five other sites (electronic equipment) accounted for another 20% of the prices collected by ICBS. These data are consistent with the expenditure data from HES and private research firms. The distribution of prices collected according to the 10 major consumption groups are:

Table 4

<table>
<thead>
<tr>
<th>Main consumption groups</th>
<th>CPI weight</th>
<th>% of e-commerce prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>13.48</td>
<td>0.2</td>
</tr>
<tr>
<td>Fruits and Vegetables</td>
<td>3.51</td>
<td>0</td>
</tr>
<tr>
<td>Housing</td>
<td>23.16</td>
<td>0.2</td>
</tr>
<tr>
<td>Household maintenance</td>
<td>9.75</td>
<td>0</td>
</tr>
<tr>
<td>Furniture &amp; Household Equipment</td>
<td>4.75</td>
<td>25.1</td>
</tr>
<tr>
<td>Clothing and Footwear</td>
<td>2.9</td>
<td>0.6</td>
</tr>
<tr>
<td>Health services</td>
<td>4.85</td>
<td>0.8</td>
</tr>
<tr>
<td>Education, culture &amp; entertainment</td>
<td>12.93</td>
<td>62.3</td>
</tr>
<tr>
<td>Transportation &amp; communications</td>
<td>20.26</td>
<td>3.2</td>
</tr>
<tr>
<td>Miscellaneous goods &amp; services</td>
<td>4.41</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Electronic equipment, computers and computer peripherals, books and CD's are classified in the Furniture & Household equipment and education, culture & entertainment consumption groups; therefore having the highest percentage of E-commerce prices.

We encountered the following difficulties while attempting to compute price indices for specific items in the pilot E-commerce CPI:

- Inability to collect enough price observations for narrowly defined items in the regular CPI. Therefore we computed indices for sub-aggregates like refrigerators, washing machines, books, cosmetics, etc.

- We did not have enough price observations for each and every month during January 2001-December 2002. Therefore price indices were not necessarily computed at monthly intervals.

- Indices in the E-commerce indices may have varying number of observations between months. This is inconsistent with the regular CPI computation methods.
• Certain items may have enough observations, but are from different sites. Therefore price indices were computed according to changes in average prices and not the average change in prices.

• Most of the internet sites for large electronic equipment were auction sales. We have yet to devise suitable methodology to measure price changes for auction prices (electronic equipment). However we did manage to compute price indices for washing machines and dryers.

Taking into consideration all of these limitations, we succeeded, at this stage, to compute fairly robust price indices, during January 2001-December 2002, for the following four consumption groups: books, CD's, cosmetics and washing machines and dryers.

In figures 6-9 below we present these indices compared to the indices computed in the actual CPI.

**Figure 6: Comparison of E-commerce and actual CPI for Books**

![Graph of books comparison](image)

**Figure 7: Comparison of E-commerce and actual CPI for CD's**

![Graph of compact discs comparison](image)
The variance of price changes in the e-commerce series is much larger than the actual CPI. While some of this variance can be explained by the limitations stated above, it seems that both conditions to justify special (additional) indices for E-commerce have been fulfilled. Consumption weights for e-commerce are significant for several items in the index; trends of

---

6 The base period for the comparison of the actual and e-commerce indices for cosmetics and washers and dryers is June 2001. While the period is shorter, the same trend of variability in the e-commerce index vs. the actual index is evident in this series as in the other indices presented in this paper.
price changes may be substantially different and can not be approximated by the actual CPI. The following table compares the standard deviations of price changes for the two series.

### Table 5

<table>
<thead>
<tr>
<th>STD of price changes</th>
<th>E-commerce</th>
<th>Actual CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books</td>
<td>0.096</td>
<td>0.022</td>
</tr>
<tr>
<td>Compact disks</td>
<td>0.127</td>
<td>0.021</td>
</tr>
<tr>
<td>Cosmetics</td>
<td>0.349</td>
<td>0.041</td>
</tr>
<tr>
<td>Washers &amp; dryers</td>
<td>0.242</td>
<td>0.018</td>
</tr>
</tbody>
</table>

While presenting an exhaustive discussion of E-commerce, we have yet to recount the measurement of internet services (monthly connection fees to ISP's). These are usually included in national CPI's and seem to be straightforward. While this is outside the scope of this paper, in section 5 we present some thoughts on the complications of internet services.

## 5. Internet Services

Internet services are becoming more difficult to measure in a CPI framework. In the beginning, a consumer was required to pay a monthly fee for unlimited connection or various hourly fees. With the introduction of wideband services and fierce competition among internet service providers (ISP's) which led to sinking relative prices, they are now seeking ways to differentiate themselves from each other. A similar process occurred among cellular phone companies and cable television providers in previous years. In order to "capture" the consumer and encourage loyalty to a specific company, cellular and cable operators devised complicated pricing schemes with many value added services.

### 5.1 Value added services provided by ISP's

- **Enlargement of e-mail**: all providers offer a standard 5 MB for regular modem connections and 10 MB for wideband internet. The providers are willing to enlarge e-mail boxes to 50 and 100 MB at extra cost.

- **Private e-mail boxes for each family member**: for an added cost, companies may provide up to 3 boxes and others an unlimited amount for extra monthly fees.

- **Content services**: these are split into three kinds (1) free entrance or minimum fees for sites with admission fees (2) free subscription to printed magazines and (3) newspapers and newsletters that are sent directly to the e-mail of the user.

- **Other services** are faxes, call mail, screening services of pornographic sites and anti-viruses.

Index compilers will have to deal with pricing and quality adjustment issues of these added value services. They may be similar to those of other telecommunication services, discussed in papers on measurement of cellular phones and cable services (estimation of consumer profiles, sample of consumers from companies, etc.).
ICBS launched a multi-year project to compute price indices for goods and services consumed by E-commerce. Like all indices in the CPI, we must define the E-commerce domain, calculate consumption weights, construct survey samples (outlets, items, prices) and measure price changes on a current basis. Although we encountered several difficulties in performing these tasks, we report partial success and have come to the following conclusions:

Data received from private sector sources and Household Expenditure Surveys (usually conducted by national statistical agencies) are sufficient for deriving consumption weights of E-commerce.

In Israel, market shares of E-commerce were calculated in two ways: (1) percentage of e-commerce consumption for goods and services in the CPI basket; (2) percentage of goods and services consumed in E-commerce. The findings are that significant market shares "of" E-commerce and "by" E-commerce were reported for electronic equipment, computers and peripherals, books and CD – both from the private sector sources and reiterated from the "more scientific" HES.

Measurement of e-commerce price changes was attempted for these and many other consumption groups, however, fairly robust indices were computed for only four groups: e-commerce consumption of books, CD's, cosmetics and washers and dryers. These showed large variance in the e-commerce series, compared to the indices computed in the actual CPI.

During 2003-2004, ICBS will continue price collection and measurement of price changes in order to improve the execution of pilot indices and expand the realm of consumption groups that are covered; these shall be introduced into the official CPI at its next major update, in January 2005.

Many methodological and processional problems are to be solved in the future: international standards for definition and classification of e-commerce, constructing of outlet (e-commerce sites) registers, measurement of prices and price changes for auction sales, questions of disappearing items and quality adjustment for e-commerce, etc.

Sharing experiences with other national statistical agencies will be helpful in devising an improved methodology in months to come.

---

7 For example, some of the variance may be explained by the small amount of price observations for each good and service; enlargement of price sample may reduce this.
E-commerce and consumer behaviour

Robin Lowe¹

Statistics Canada

1. Introduction

One of the opportunities the growth of the Internet has provided is the ability for users to browse commercial products shown on the Internet, and in some cases to order them over from a computer. The act of ordering is called e-commerce,² and its growth has been a matter of much speculation. In this paper we look at its impact on consumer behaviour, and by extension, on price measurement in the CPI.

Because of some obvious advantages in convenience to the purchaser, expected cost advantages to the supplier in not having to maintain such a visible inventory and the opening up of more diverse markets for suppliers, e-commerce was predicted to grow quickly. In general it has grown quickly in percentage terms, though not nearly so fast as most predictions, but the latest figures show it to still be a very small percentage of consumer spending, and of business spending as well. In this paper we will concentrate on consumer spending: although business-to-business sales are about five times as large, they are still a very small proportion of business sales, and our main interest here is the impact that e-commerce has for measuring the CPI.

In 2001, e-commerce by private households in Canada was estimated at almost $2 billion, more than a 50% rise from the previous year. Nevertheless, this amounts to only 0.4% of household expenditure. Even if we make an adjustment to express it as a percentage of discretionary spending (for most purchases over the Internet could be so described) it only rises to two or three percent. If e-commerce were one unitary commodity, its weight would be large enough to merit surveying in a CPI, but Internet purchases are distributed among a number of products, and the activity is still a small share of those as a whole, and most products individually. Information from retail trade surveys confirm this: in 2000 the percentage of retail goods sold on-line, or by mail order was about 1% of total retail sales, and even for those commodities most mentioned in e-commerce, books and CDs, the ratio was only about 10%.

For the present, therefore, the question is whether there are a few products whose trade over the internet is already sufficiently important to be included in our price measurement.

¹ The opinions expressed are those of the author only, and not those of Statistics Canada.
² Statistics Canada follows the OECD definition of e-commerce viz: "...sale or purchase of goods or services, whether between businesses, households, individuals, governments and other public or private organisations, conducted over computer mediated networks...". This definition includes orders placed over a network, regardless of whether or not the payment and delivery took place on-line.
Obviously there cannot be many, but these potential cases are discussed later in the paper. A more general question, is how is e-commerce likely to grow in the future, and will it soon be important enough to be measured directly in several areas.

Associated with these questions are the concerns about whether the existence of e-commerce and the way products are offered for sale might lead to changes in the way we measure prices, even if the extent of e-commerce is still not very large.

In this paper we will start by tracking the growth of internet activity and e-business activity in Canada and provide some assessment of its potential further growth. Then we will look in more detail into some of the more popular products that are bought on the web. Finally we will look at some other internet activities related to the delivery of products, and examine these implications for how we measure price change for them.

2. Overall growth in internet activity and e-business

Most of the following information comes from the annual Survey of Household Internet Use (HIUS), carried out by the Science, Innovation and Electronic Information Division (SIEID) at Statistics Canada. It began with the 1997 reference year and has increased detail since. Questions covering e-business and e-commerce date from 1999. The most recent available data is for 2001.

The amount of spending by households depends on three factors – the percentage of households that have access to the internet, the degree to which, having access, households choose to use it for buying particular products, and the range of products available.

Table 1 shows the percentage of households that have access to the internet, either at home or elsewhere, and some percentages that are internet shoppers3, either by window-shopping or engaging in e-commerce, and who pay for e-commerce electronically over the net. “Internet shoppers” include those who research product information without buying, “window-shopping”. We do not know whether household members listed as window-shoppers only ever intended to research product choices or whether they started with the intention of possibly buying over the net, but did not carry through with it. We do know that many people have concerns about security, and about reliability of delivery of products bought over the net. Most companies, knowing about peoples’ security concerns, make arrangements for payments for products purchased over the net to be made by other than electronic means.

Any analysis of e-commerce that spans the years 2000 and 2001 is compromised by the change in the question in 2001. Table 1, for example, shows that in 2000, 21.9% of all households engaged in Internet shopping form home. In 2001, 33.1% of all households engaged in Internet shopping from some location. We cannot know what proportion of that 33.1% did Internet shopping from home, and what proportion only from elsewhere. This

3 Internet shoppers: Regular-use households that engaged either in window shopping or electronic commerce.
Window shoppers (only): Households where all members reported only to have browsed for goods or services using the internet.
Electronic commerce: Households where at least one member was reported to have ordered and/or paid for goods or services using the internet, for personal or household consumption. These households may have window-shopped as well.
Electronic payment: Internet households in e-commerce households who made an online payment for at least one of their transactions.
discontinuity makes it impossible to make any comparison between the two years, and Statistics Canada does not do so. Any observations about internet shopping and e-commerce in this paper are speculations based on guesses about the missing activity, or about patterns of behaviour between home- and other- access users.

Table 1: Percentage of households accessing the Internet

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>All households</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>From any location</td>
<td>41.8</td>
<td>51.3</td>
<td>60.2</td>
</tr>
<tr>
<td>From home</td>
<td>28.7</td>
<td>40.1</td>
<td>48.7</td>
</tr>
<tr>
<td>Internet shopper</td>
<td>15.3</td>
<td>21.9</td>
<td>33.1</td>
</tr>
<tr>
<td>Window shopper (only)</td>
<td>8.3</td>
<td>9.6</td>
<td>14.4</td>
</tr>
<tr>
<td>Electronic commerce</td>
<td>6.9</td>
<td>12.3</td>
<td>18.7</td>
</tr>
<tr>
<td>Electronic payment</td>
<td>5.3</td>
<td>9.7</td>
<td>14.8</td>
</tr>
</tbody>
</table>

Note: In 2001 the survey was redesigned to capture Internet shopping by households that regularly used the Internet from any location. In previous years it only captured shopping conducted from home. Thus there is a break in the series preventing comparisons between 2001 and earlier years.

2.1 Growth of access to the Internet

We know the volume of e-commerce has clearly been growing quickly, partly as more and more people get access to the Internet. Let as look at the potential for further growth as access increases.

Table 2: Growth of Internet access rates, and rate of access and internet shopping in 2001, by income quartile

<table>
<thead>
<tr>
<th></th>
<th>Percentage growth in households with Internet access</th>
<th>Percentage of households in 2001 with access to the Internet</th>
<th>Percentage of Internet users in 2001 who engaged in e-commerce</th>
</tr>
</thead>
<tbody>
<tr>
<td>All households</td>
<td>23.8</td>
<td>16.4</td>
<td>22.7</td>
</tr>
<tr>
<td>Lowest quartile</td>
<td>7.3</td>
<td>43.5</td>
<td>27.2</td>
</tr>
<tr>
<td>Second quartile</td>
<td>32.2</td>
<td>22.7</td>
<td>46.5</td>
</tr>
<tr>
<td>Third quartile</td>
<td>28.7</td>
<td>15.6</td>
<td>26.0</td>
</tr>
<tr>
<td>Highest quartile</td>
<td>21.4</td>
<td>9.5</td>
<td>9.4</td>
</tr>
</tbody>
</table>

Table 2 shows the percentage of households with Internet access in 2001, by income quartile, and the year-to-year growth of access for each income quartile. While the rate of Internet use has risen considerably in each year there was evidence that the rate of growth was beginning to ease off in 2001 and the limit to access is in sight. Three more years of 15% annual growth would put overall access over 90% . Furthermore, the potential for increased access is largest among the lower income households, whose use of e-commerce, when they do have access, has been lower than higher income households. While the percentage of internet users in the highest income quartile in 2001 engaging in e-commerce was 41.6, it fell through 30.1 and 22.4 to 18.0 for the lowest income quartile. If all households had the same access rate as the
highest quartile did in 2001 (87.3%) the number of households engaged in e-commerce, on 2001 rates, would only be 30% higher than in 2001, and the increase in value of expenditures would be rather less, as expenditures by lower income households would naturally be lower. It seems unlikely that there can be more than another 20% or so of growth in e-commerce in total due to increased access alone. Spread out over the next few years it suggests a low rate of growth per year.

2.2 Increased use by owners

The second impetus to growth in e-commerce could be increased use by households who have access. This may show as an higher percentage of users engaging in e-commerce, or more activity by those who are already buying. Table 3 shows how internet shopping and e-commerce rates have changed from 1999 to 2000 among households who have internet access at home, and rates in 2001 for all those who have internet access, from any location.

Table 3: Percentage of regular-use households with internet access engaging in internet shopping

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>From any location</td>
<td>N.A.</td>
<td>N.A.</td>
<td>100.0</td>
</tr>
<tr>
<td>From home</td>
<td>100.0</td>
<td>100.0</td>
<td>N.A.</td>
</tr>
<tr>
<td>Internet shopper</td>
<td>53.2</td>
<td>54.6</td>
<td>55.0</td>
</tr>
<tr>
<td>Window shopper (only)</td>
<td>29.1</td>
<td>23.9</td>
<td>24.0</td>
</tr>
<tr>
<td>Electronic commerce</td>
<td>24.1</td>
<td>30.7</td>
<td>31.0</td>
</tr>
<tr>
<td>Electronic payment</td>
<td>18.4</td>
<td>24.3</td>
<td>24.6</td>
</tr>
</tbody>
</table>

Because of the break between 2000 and 2001 we cannot make exact comparisons but if we make the extreme assumption that in 2000 no-one engaged in e-commerce away from home the figure for 2000, of 30.7% home-access households engaging in e-commerce, would only have dropped from to 24% if expressed as the percentage of households with any kind of access (51.3% of households instead of 40.1%, from table 1). Although the incidence of Internet users using it for in e-commerce is rising, the increase has only been moderate, and for those engaging in window shopping only, it has risen barely, if at all.

Although there may not be a higher proportion of internet users using it for e-commerce, perhaps those who are using it are using it more extensively. Table 4 shows the average size and number of orders. Again we have the problem of comparing between 2000 and 2001 as we do not know the difference in the number and types of order placed from home compared to other locations. The pattern of orders and the frequency of orders may be different between those who place their orders from home and those who place them elsewhere.

Table 4: Average expenditure and number of orders placed

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Expenditure by households engaged in e-commerce ($)</td>
<td>517</td>
<td>757</td>
<td>880</td>
</tr>
<tr>
<td>Average number of orders</td>
<td>4.1</td>
<td>6.2</td>
<td>5.9</td>
</tr>
<tr>
<td>Average expenditure per order ($)</td>
<td>125</td>
<td>121</td>
<td>148</td>
</tr>
</tbody>
</table>

(1999 and 2000 orders placed from home; 2001 orders placed from anywhere)
From 1999 to 2000 there was a sharp increase in e-commerce users’ frequency of use. This coincides with the increase in the percentage of users in 2000. But the increase in frequency of use was not sustained through 2001. It is quite reasonable to guess that some decline in the average number of orders is due to the inclusion of those who do not do their internet shopping from home, as it is probably less convenient, but even if we assumed there was no e-commerce except from home the average size of order would have risen only to 7.3, which we can take as an upper bound. The main cause of increased expenditure in 2001 was an increase in the average value of order, which may reflect a different mix of products being bought.

Taking the information from tables 3 and 4 together there does not seem to be a breakthrough in general usage by existing Internet users. We need next to look at the types of products offered and the way they are offered. One disincentive, security concerns, appears not to have changed in this period, and while it may deter potential buyers from even considering purchasing in the first place, three-quarters of those that do make purchases do it my electronic payment. (compare each year the bottom two lines of table 1).

The change in average order size in 2001 suggests that a different mix of products might be attracting orders than in earlier years. We can see this from Tables 5 and 6 that list the kinds of products that interest Internet shoppers. We have no data on expenditures by category – that would require an extensive and time consuming survey for respondents. These data are based on respondents who do conduct e-commerce saying that they either buy or window-shop in the various listed product categories. We have again the problem of non-comparability between 2000 and 2001. There were also some changes in product categories in successive years, which affects the content of “other” particularly.

2.3 Diversity of products

The make-up of what is bought has changed considerably between 1999 and 2001. The products are listed by their interest of the highest number of people in 2001. The products notionally associated with e-commerce – books, computer software and music – were the most important in 1999, but they have all fallen in importance since.

Table 5: Percentage of internet shoppers buying the following (e-commerce)
(Only those with 5% or more respondents in 2001)

<table>
<thead>
<tr>
<th>Category</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books, magazines, newspapers</td>
<td>37.5</td>
<td>36.6</td>
<td>28.1</td>
</tr>
<tr>
<td>Clothing, jewellery etc.</td>
<td>12.3</td>
<td>17.4</td>
<td>18.2</td>
</tr>
<tr>
<td>Travel arrangements</td>
<td>12.4</td>
<td>12.2</td>
<td>16.2</td>
</tr>
<tr>
<td>Other</td>
<td>12.4</td>
<td>21.4</td>
<td>14.8 (but see new categories)</td>
</tr>
<tr>
<td>Computer software</td>
<td>24.8</td>
<td>16.0</td>
<td>13.8</td>
</tr>
<tr>
<td>Music (CDs, etc.)</td>
<td>18.5</td>
<td>15.4</td>
<td>11.8</td>
</tr>
<tr>
<td>Other entertainment (tickets, etc.)</td>
<td>9.3</td>
<td>8.0</td>
<td>10.6</td>
</tr>
<tr>
<td>Consumer electronics</td>
<td>4.9</td>
<td>6.8</td>
<td>6.7</td>
</tr>
<tr>
<td>Computer hardware</td>
<td>9.0</td>
<td>8.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Toys and games</td>
<td>2.3</td>
<td>6.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Housewares</td>
<td>3.3</td>
<td>5.5</td>
<td>5.6</td>
</tr>
<tr>
<td>Sports equipment</td>
<td></td>
<td></td>
<td>5.5</td>
</tr>
<tr>
<td>Health, beauty, vitamins</td>
<td></td>
<td></td>
<td>5.1</td>
</tr>
<tr>
<td>Videos, video discs</td>
<td>5.7</td>
<td>5.4</td>
<td>5.0</td>
</tr>
</tbody>
</table>
To a large extent this is due to the greater diversity of products available. New users have not been so particularly interested in those products that were most important in 1999. The main products which have grown in importance are clothing, travel arrangements and, to some extent, tickets for entertainment.

What has happened to window-shopping has been quite different.

Table 6: Percentage of window-shoppers examining the following (Only those with 5% or more interest in 2001)

<table>
<thead>
<tr>
<th>Category</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books, magazines, newspapers</td>
<td>32.8</td>
<td>19.8</td>
<td>15.9</td>
</tr>
<tr>
<td>Clothing, jewellery etc.</td>
<td>23.3</td>
<td>28.9</td>
<td>25.6</td>
</tr>
<tr>
<td>Travel arrangements</td>
<td>30.9</td>
<td>17.0</td>
<td>16.3</td>
</tr>
<tr>
<td>Other</td>
<td>7.2</td>
<td>18.1</td>
<td>11.1 (but see new categories)</td>
</tr>
<tr>
<td>Computer software</td>
<td>26.6</td>
<td>13.5</td>
<td>10.7</td>
</tr>
<tr>
<td>Music (CDs, etc.)</td>
<td>22.6</td>
<td>11.9</td>
<td>10.7</td>
</tr>
<tr>
<td>Other entertainment (tickets, etc.)</td>
<td>14.0</td>
<td>4.6</td>
<td>5.1</td>
</tr>
<tr>
<td>Consumer electronics</td>
<td>19.7</td>
<td>18.9</td>
<td>21.6</td>
</tr>
<tr>
<td>Computer hardware</td>
<td>23.4</td>
<td>14.2</td>
<td>12.2</td>
</tr>
<tr>
<td>Toys and games</td>
<td>3.0</td>
<td>9.7</td>
<td>8.1</td>
</tr>
<tr>
<td>Housewares</td>
<td>13.9</td>
<td>17.0</td>
<td>26.4</td>
</tr>
<tr>
<td>Sports equipment</td>
<td></td>
<td></td>
<td>9.0</td>
</tr>
<tr>
<td>Health, beauty, vitamins</td>
<td></td>
<td></td>
<td>5.7</td>
</tr>
<tr>
<td>Videos, video discs</td>
<td>9.5</td>
<td>6.1</td>
<td>6.3</td>
</tr>
<tr>
<td>Automotive products</td>
<td>28.6</td>
<td>23.8</td>
<td>21.2</td>
</tr>
</tbody>
</table>

Here we find that (except clothing) the most important products for e-commerce have dropped dramatically in importance for window-shopping, more than can be explained just by diversification. It appears that for these products users are either prepared to buy, or not interested. On the other hand, there are some product categories that are very popular among window shoppers, but they do not want to make purchases on the basis of computer images or descriptions. (The categories are listed in the same order as in the preceding table, in descending order by importance in e-commerce in 2001). These products are clothing (to some extent), consumer electronics, housewares and most notably automotive products. But except for clothing, the increase in interest in window-shopping has not converted to internet sales.

2.4 Summary of overall activity

Growth in e-commerce in recent years has been driven more by increased numbers gaining access than by an increased propensity to use it by existing internet users. The growth in access is likely to slow considerably, and increases in usage have been only moderate, despite an increasing range of products for which e-commerce purchases have been made. Although there may be some exceptions, e-commerce generally is not likely to account for a large enough share of purchases for it to be necessary to cover it in price surveys; unless a seismic shift in households' perception of the value of internet shopping occurs.
3. Synopsis for specific products

Let us look at the more popular product categories in more detail. The following tables show the percentage of internet users who are window shopping or buying these products. It is an indicator of propensity only; the actual numbers of buyers has been rising as more and more households have had access. Except for clothing there has been a marked decline in the percentage of users window-shopping this was most noticeable between 1999 and 2000 when the questions were unchanged. There appear to be two reasons for this: new users are less inclined to use the internet, and existing users either increase their amount of e-commerce or stop using the internet for those products. As for the trends in e-commerce, there may be specific reasons for the changes in each product, but there are only two categories — travel arrangements and clothing, which are clearly growing.

Table 7: Percentage of internet users shopping on the internet for books, magazines and newspapers

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>9.5</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>E-commerce</td>
<td>9.0</td>
<td>11.2</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Table 8: Percentage of internet users shopping on the internet for computer software

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>7.7</td>
<td>3.2</td>
<td>2.6</td>
</tr>
<tr>
<td>E-commerce</td>
<td>6.0</td>
<td>4.9</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Table 9: Percentage of internet users shopping on the internet for music

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>6.5</td>
<td>2.8</td>
<td>3.3</td>
</tr>
<tr>
<td>E-commerce</td>
<td>4.4</td>
<td>4.7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

These are the products that were the most important for e-commerce in 1999. In all cases (books less definitely because of the discontinuity problem in 2001) the likelihood of an internet user buying these products using the internet has fallen between 1999 and 2001. The drop in window-shopping has been much sharper. The number of households with the internet has grown, and it may well be that most of those who were most interested in these products had got access by 1999. However, current trends do not suggest any significant growth in e-commerce for these products. One caveat that past and current use may be underreported, especially for music and software, is that in the Survey of Internet Use, one household member reports on behalf of everyone in the household. That one person may not know of activities by other members of the household. The reporting in this survey will change from a household to an individual survey for reference year 2004, funding permitting.

Is there a chance of a sudden change in usage? In Canada until 2001 there were two Canadian chains selling books and some CDs over the internet and Amazon.com in the U.S. was also a potential supplier. In the year ending March 31 2001 Chapters reported online sales
to be 7.5% of total retail sales, up from 5.9% the previous year. The dollar increase in online sales and retail stores sales was virtually identical. The two Canadian chains merged into Chapters/Indigo in 2001; in the year ending March 2002 Indigo reported online sales were 5% of total revenues. In 2002 Amazon opened a Canadian branch, eliminating the problems of conversion to US currency and of delays or extra charges through customs. It also provided more direct competition, which has been met by matching prices and promotions in most cases. It will be interesting to see if internet purchases have grown since mid-2002, but there is no evidence of online sales becoming significant anytime soon.

Implications for price measurement: In the Canadian CPI book prices are surveyed in a non-standard way. The top five best-sellers (with some restrictions) are priced in various outlets and the sum become the price for that month to be compared to the sum for the previous month, which may be, and usually is, for a different parcel of books. It would not be difficult to include internet outlets in the sample, and an estimate could be made of their share of the market. They tend to have lower prices than the same books in store, though not necessarily for best-sellers which are often discounted in stores too. In addition, the price of shipping has to be included; this varies according to the size of the order, so a view must be taken of the likely size of the purchase - a question which will be returned to later. This method, treating each outlet independently, does not make any comparison of price levels between internet outlets and other outlets, so that any effective price drop, if any after comparing the qualities of services offered, is not reflected in the index.

### 3.1 Travel arrangements

#### Table 10: Percentage of internet users shopping on the internet for travel arrangements

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>8.9</td>
<td>4.1</td>
<td>3.9</td>
</tr>
<tr>
<td>E-commerce</td>
<td>3.0</td>
<td>3.7</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Philip Wolf, a US analyst of the travel industry, described it a year or so ago as “the oasis in the e-com desert”. It is the only product category that shows a marked increase in the percentage of users, although that percentage is still low. This area shows growth potential. Although booking of hotels and car rentals is included in this category by far the largest portion of purchases in this category is airline tickets. The airline industry, dominated and led in Canada by Air Canada, is trying to increase the amount of direct purchase over the Internet, either by travel agents or by households. At the end of 2000 it was 1% of bookings with Air Canada; a year ago they stated they hoped to raise this to 20% by 2005. Among predictions of e-commerce growth this is among the more conservative, and they are presently ahead of their targets. They are encouraging growth in e-commerce in several ways: increasing the use of electronic tickets (the only option for domestic flights now), advertising sale prices first on the Internet, giving modest incentives for booking that way, and making their pricing more transparent and straightforward. This is already in place with their discount arm Tango which can only be booked over the internet or directly from an Air Canada call centre. In recent months direct purchases by the consumer sector are averaging 7% of all ticket revenue, and

---

4 From Chapters annual report 2000-2001  
5 From Indigo Books and Music, Inc. annual report 2001-2002  
6 Information based on conversation with Air Canada representative, April 2003.
closer to 10% of all bookings, which implies a higher percentage of consumer bookings. They expect these percentages to at least double within the next twelve months.

Implication for price measurement: This is another product where our method of measuring prices does not depend on a retail survey, but on being provided by the carrier with prices under specified conditions for specific routes on certain dates. Although still a small fraction at present, internet sales look likely to rise, especially in the household sector. Internet sales can have a marked impact of the prices that should be used in the CPI, as pricing methods change. Air fares have been notoriously complicated with price discrimination and frequent changes in fares as the departure time approaches, so that it has been difficult to claim that a moderate sample survey captures prices adequately. Air Canada’s CEO recently stated that this complicated system cannot be maintained with prices posted on the internet, so there will be a simplification of the rate schedules that will make it easier to cover the product better. This does not mean constant prices; the pricing policy on Tango combines two new factors. First, flights are priced individually, so that a return trip can be planned by buying any combination of flights. Second, price will vary with demand. If a flight is selling quickly, its prices will rise, and if not, not. So someone who is flexible about their dates can get lower prices. Although the present method of collecting prices is cheap and easy, if we were replacing a retail survey it would be easy to survey the internet at certain times and take the prices from there. This means that we would need to specify the terms of sales more precisely, making some assumptions about consumer behaviour. A potential customer who can plan well ahead and is reasonably flexible about exactly which day to fly will find different prices from one that has to fly at virtually no notice.

3.2 Clothing and automotive products

Table 11: Percentage of internet users shopping on the internet for clothing

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>6.7</td>
<td>6.9</td>
<td>6.1</td>
</tr>
<tr>
<td>E-commerce</td>
<td>3.0</td>
<td>5.3</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Table 12: Percentage of internet users shopping on the internet for automotive products

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window shoppers</td>
<td>8.3</td>
<td>5.7</td>
<td>4.4</td>
</tr>
<tr>
<td>E-commerce</td>
<td>0.6</td>
<td>0.6</td>
<td>0.8</td>
</tr>
</tbody>
</table>

These are two product categories that are most popular for window-shopping. However, for clothing the internet is becoming steadily more popular. It is also an area where some retailers have lost a lot of money and withdrawn from the business. The companies that have been most successful with online selling are those that had successful catalogue sales before. It really provides an alternative to catalogue; and they never got large enough for them to include in our survey.

Implications for price measurement: Because many companies in effect post their catalogue on the internet there is the possibility of using these prices as proxies for field survey prices. The advantages are cost, and the knowledge that these are truly the price at which the good sells if bought over the internet. One disadvantage is that it may be difficult to evaluate
quality change, particularly for clothing, and prices may not be the same in the stores as online. There has been no study of this yet for the CPI, but we do know that instore prices do vary in different locations. The other difficulty is that it is one thing to check online prices against the field selected ones, where possible, and another to use the online prices to design the survey. We would be lacking the information in individual stores that enables the price collector to choose a representative seller. The best use of online prices may be as an editing tool to checking against the field collected prices.

3.3 Digital products

With the exception of clothing, most of the popular or growing products on the internet are products that can be delivered over the internet, or whose delivery depends little on retail service that can vary from store to store. For those products, it seems that consumers still prefer to see the products in person.

Some products, however, sometimes called digital products, are delivered directly to the customer’s computer. Examples include music, gameware, computer software, and services such as courses taken over the internet. All these are included in e-commerce, but according to the survey none of them are yet large enough to demand direct pricing. There are some other products that owe their existence to the Internet and they will be discussed here, as they account for expenditure that is too large to ignore, and whether or not there are paid for over the internet aspects of the product are delivered over the net. These are internet access, banking and phone service.

3.4 Other products

3.4.1 Internet access

Most people with internet access from home purchases it\(^7\); in 2001 it accounted for 0.30% of consumer expenditure in the CPI, and a price index was added to the CPI starting in January 2003. For price measurement it is immaterial whether the service is purchased via e-commerce or not. All providers in the CPI price survey offer their plans via e-commerce, and their prices are collected from the net. The price index is calculated conventionally averaging the percentage changes in price as they occur, with adjustments for quality change in existing products.

3.4.2 Electronic banking

In 2001 44.4% of households with internet access at home used it to do electronic banking. This compares to 26% using it for purchasing goods or services. All the major banks in Canada offer the ability to bank on-line. Direct fees for banking services were included for the first time in the CPI in January 2001, with a weight of 0.51%.

Bank services in Canada can be paid for individually, or in a package of common services. Banks offer many plans, “tailored” to different customers needs. Some of these plan include electronic banking. Most people subscribe to one of these plans, paying a monthly fee, which is sometimes augmented if they use more than is included in the plan. For example if the plan includes forty debit card transactions a month and the customer makes 42, there will be a charge for each of the extra ones. Given the array of different plans it is difficult to select which ones to include in the sample without knowing the customer’s preferences.

\(^7\) A significant proportion have access provided by their employer.
So this index is produced by defining some customer profiles, and calculating, from among the plans each bank offers, the cheapest combination to provide the required services. For these profiles, it is assumed the customer will stay with a particular bank, but will switch plans efficiently. This means that if plans change their content, a different plan may well replace the previous one for a particular profile. Given unchanging requirements, this will be a straight price change. However, maintaining the index will also require that the profiles are reviewed regularly, and then a change is made in the requirements in a profile the value of those changes must be evaluated. For example, if a plan that is the cheapest for one profile adds internet banking access at no extra cost, but the profile does not include it, there is no change in quality or price. However, if a profile is changed to reflect the need for electronic banking, the value of that in any plan that is chosen must be assessed as a quality change.

3.4.3 Phones, particularly cellular phones

The growth of cellular phone use (43% of households reported expenditures in 2001, for 0.50% of consumer expenditure) has also led to a growth of plans, most of which are available over the internet. There is a similar range of plans for long-distance service on landline phones, but only cellular phone service is discussed here. For cellular phones there are hundreds available, and at least one company will offer a custom plan to match any a particular customer may have found elsewhere. There are two ways of paying for cellular phone service, pre-paid and post-paid. Prepaid service is bought by buying credit, using it, then buying more when, or preferably before, it is used up. How fast that happens depends on the type as well as the volume of calls made. The standard way to do this used to be by a physical card, but now credit can be bought over the internet from companies’ websites. There are also a variety of additional services which can also be charged to that credit. Postpaid plans work similarly, except there is a monthly charge for services, which is paid whether they are used or not. These plans do not have to be bought over the internet, but companies are trying to make it easier to do so. Again, this aspect of e-commerce does not show in the e-commerce surveys, but appears to be growing fast.

As with banking, it is difficult to select a representative products, or even a representative bundle of products from each supplier. The range of services each offers is virtually identical, and can be parcelled in almost any way. It is necessary therefor to focus on the consumer, and what is a representative package of preferences. This package can then be priced by the various suppliers.

How the prices are combined into the index depends on the profile. With prepaid service it is very easy to switch from one plan, or one supplier to another as the value of a card is used up. For the customer who does that it is appropriate to select as the price, the lowest price offered by any supplier. Pre-paid has been growing at the expense of post-paid plans, and suppliers are trying to reverse that trend. As they do not want to compete on the direct price of service, they have been introducing “loyalty” incentives, a lower phone price if you commit to purchasing enough credit over the next specified period, for example. Loyalty incentives already exist for most post-paid plans. So what range of prices, from what suppliers depends on what is specified in the profile concerning the willingness to switch plans.

4. Summary

What is common to these products, to travel, and to some extent to books and CDs, is that we need to reverse the standard for of field survey. Instead of selecting representative commodities from the range of different products whilst indifferent to the purchaser, we have
to select a representative purchaser of a range of undifferentiated products. This means that we might compare directly products from different supplies in compiling an index.

For phone service, where the cost and inconvenience of switching is small, all price offers might be considered and the lowest one used. (This is already the practice for contracted construction work). For the supply of books or CDs over the internet, where the products are identical, the lowest price would be selected, assuming the customer is indifferent between suppliers. These would depend on the perception of differences in the quality of service different suppliers provide, and ancillary costs. For digital products these differences may be small, but even for other products this approach may be suitable.

For example, there are three alternative carriers within Canada on many domestic routes, Air Canada, its discount arm, Tango, and Westjet. They do not provide the same services, and many users might judge the difference in services not justified by a difference in price. (Though we have never adjusted for any difference in service over time.) However, it is quite possible that a consumer will be quite indifferent between suppliers for a small difference in price, and the profile should reflect that. The difficulty, of course, is to devise representative profiles.

This reversal of focus arises because the point of purchase has moved. The shopping model is no longer of choosing one retail store over others and selecting from its products, but the computer. When different products have different characteristics this does not matter, but for relatively homogeneous products it does. Furthermore, most of these products are a bundle of different services which can be bought in many combinations. As the products are identical among suppliers, or, if different, one supplier can be easily substituted for another it is the consumer’s preference, not the range of products available that is most important. So instead of monitoring the price of a representative product and adjusting for quality change we should monitor the price of a representative consumer’s bundle of requirements, and adjust for quality change when we change the representative consumer.

This situation is not unique to digital products, it applies to the class of products that are delivered by a common system – conventional phones, natural gas, electricity – but the growth of digital products on the internet has highlighted it.

References

Household Internet Use Survey, Science, Innovation and Electronic Information Division, Statistics Canada, www.statcan.ca, Cansim II, tables 358-0002 to 358-0006

Quarterly Retail Commodity Survey, Distributive Trades Division, Statistics Canada, www.statcan.ca, Cansim II, tables 080-0010

Retail Non-store Survey, Distributive Trades Division, Statistics Canada.
The Present Situation of E-Commerce and a Pilot Survey on E-Commerce

Masato Okamoto\textsuperscript{a} and Eiji Nitta\textsuperscript{b} 1

\textsuperscript{a}Statistics Center and \textsuperscript{b}Statistics Bureau, Ministry of Public Management, Home Affairs, Posts and Telecommunications (MPHPT), Japan

1. Introduction

Following inclusion of index for PCs as a part of the 2000 revision, Japanese CPI was further enhanced with respect to IT-related categories by addition of indices for PC printers, digital still cameras and internet access fees in the 2003 minor revision. At present, plans for the next 2005 revision are discussed. One of the focal points related to the diffusion of information technology among consumers is the possibility of the expansion of e-commerce. In this context, the present situation of e-commerce has been analyzed based on the existing statistics, and a pilot survey was conducted in November of the last year in order to clarify the impact of e-commerce on price statistics. Price comparison between direct sales by manufacturers and sales by retailers was also conducted for PCs as a supplement to the pilot survey, taking it consideration that all PC makers have begun direct sales via the internet recently, and one of the makers specializing in direct sales has made a success of its business.

2. Present situation

As shown in Table 1, “order-from-remote-place” such as mail-order, teleshopping and “order-via-the-internet” was estimated to account for 1.8 percent of household spending for goods, and 0.8 percent of the total household consumption in 1999 (bank transfers, stock trade and reservation of purchases such as ticket reservation and hotel reservation by phone/fax or via the internet are excluded here.). Thus, e-commerce was minor in the consumer market just before the turn of the century (even if including ticket reservation and other services). Probably, e-commerce has been expanding its share with a rapid increase of internet users since the beginning of this century while the traditional “order-from-remote-place” business has slowed down. At present, we do not know the exact market share of e-commerce, which is to be obtained from a new survey soon. However, several known statistics seem to support a view that it still remains minor in the market:

a) About 10 percents of households make purchases using the internet in a month as shown in Table 3.

b) Frequency of purchase via the internet is only once in two months on average. Several surveys are available on this matter. All of them are about the same.

\textsuperscript{1} The opinions expressed in this paper are those of the authors and do not represent official views of either Statistics Bureau or Statistics Center.
c) Spending for purchases via the internet is about 67,000 yen in a year on average according to a web survey for visitors at the website of some direct marketing traders association.

Based on a) and c), e-commerce is estimated to account for 0.1 – 0.2 percent of household expenditure. It indicates the share of e-commerce is still very small although it should be noted that a sample of such a type of survey cited in c) tend to be biased.

| Table 1: Percentage of "order-from-remote-place" |  
| --- | --- | --- | --- | --- | --- | 
|  | in the total consumption | in spending for goods |  
| PCs | 4.2 % | 4.7 % | 2.6 % | 4.0 % | 10.5 % | 15.1 % | 3.8 % | 3.4 % | 17.1 % | 28.7 % |  
| Books & magazines | 4.7 % | 4.0 | 4.7 | 4.0 | 4.7 | 4.0 | 4.7 | 4.0 | 4.7 | 4.0 |  
| Audio & video disks | 10.5 % | 15.1 % | 10.5 % | 15.1 % | 10.5 % | 15.1 % |  
| Clothes & footwear | 3.8 % | 3.4 % | 3.8 % | 3.4 % | 3.8 % | 3.4 % |  
| Health fortification | 17.1 % | 28.7 % | 17.1 % | 28.7 % | 17.1 % | 28.7 % |  

Source: The National Survey of Family Income and Expenditure, Statistics Bureau, MPHPT

| Table 3: Purposes of use of the Internet (monthly average in 2002) |  
| --- | --- |  
| There is a household member who used the Internet in the latest month | 37.0 % |  
| E-mail | 29.0 |  
| Gathering information | 32.3 |  
| Purchases (order, reservation) of goods and services via the Internet | 8.8 |  
| Participation in auction | 3.1 |  
| Bank transfer | 2.5 |  
| Transaction in financial assets such as stocks, insurance | 1.6 |  
| Others | 1.9 |  

Source: The Survey of Household Economy, Statistics Bureau, MPHPT

| Table 4: Goods and services purchased by household heads via the Internet (Nov., 2001) |  
| --- | --- |  
| Ticket booking | 40.3 % |  
| Books & CDs | 36.9 |  
| PCs & related items | 32.6 |  
| Miscellaneous goods for hobbies | 29.1 |  
| Accessories for a dress | 24.6 |  
| Food | 16.0 |  
| Tour reservation | 13.6 |  
| Furniture, electrical appliances & household articles | 12.7 |  

Source: The Communications Usage Trend Survey, Information & Communications Policy Bureau, MPHPT

The above table shows a ratio of a number of household heads purchased the relevant commodity via the Internet to a number of all household heads made purchases via the Internet in the past year.
Table 5: Goods and services purchased via the Internet (Mar., 2001)

<table>
<thead>
<tr>
<th>Item</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>books &amp; magazines</td>
<td>36.9%</td>
</tr>
<tr>
<td>music CDs &amp; videos</td>
<td>19.7%</td>
</tr>
<tr>
<td>PCs</td>
<td>33.2%</td>
</tr>
<tr>
<td>PC &amp; game software</td>
<td>28.7%</td>
</tr>
<tr>
<td>food, drink &amp; liquor</td>
<td>34.3%</td>
</tr>
<tr>
<td>Clothing, footwear &amp; accessories for a dress</td>
<td>32.2%</td>
</tr>
<tr>
<td>hotel &amp; tour reservation</td>
<td>28.7%</td>
</tr>
<tr>
<td>Cosmetics &amp; goods for beauty treatment</td>
<td>16.4%</td>
</tr>
</tbody>
</table>

Source: The E-mail Survey on Internet Users, Fujitsu Research Institute.
The above table shows a ratio of a number of persons purchased the relevant commodity via the Internet to a number of all persons made purchases via the Internet in the past 6 months.

As a whole, this new channel of distribution is considered to be a matter of marginal importance to the CPI compilation at present.

For all that, attention should be paid to the impact of e-commerce on price statistics because e-commerce has possibly already obtained or being expected to shortly obtain a certain portion of household spending for some specific categories. Several statistics indicate books, magazines and CDs are the most popular items in "order-via-the-internet". However, as such items are sold at uniform prices due to the resale price maintenance system, the accuracy of price index has not been affected by prosperity of online bookstores or CD shops so far. This fact indicates that consumers prefer "order-via-the-internet" for the reason of convenience in some cases; they do not always search online shops for lower prices. On the other hand, as for other items possibly promising in the net-world such as PCs, users' motivation to choose the online market has not yet been clarified.

3. Pilot survey for price comparison between sales on the internet and sales at retail shops

In order to clarify price differences between e-commerce and traditional retail at shops, a survey for price comparison between both channels of distribution was conducted in November of the last year as a part of a quinquennial cross-sectional price survey named the "National Survey of Prices". Because of practical reasons such that there is no reliable register of e-commerce business, and it is difficult to identify equivalent products regarded as representative in both new and traditional channels of distribution in many cases, the price comparison was conducted as a pilot survey, focused on prices of six kinds of electrical products: PCs, PC printers, digital still cameras, digital video cameras, DVD players and handy audio players. These products were selected based on the fact that PCs and related products such as PC printers are the most popular category in "order-via-the-internet" after books, magazines and CDs, and the rest four items also seem to be purchased via the internet.

The National Survey of Prices aims at investigating characteristics of retail stores having influence on price level such as location, size and type of stores. The regional difference index of consumer prices is compiled from the results of this survey.

by preference in addition to an advantage in that it is easy to identify specific products in these items.

About 300 traders in e-commerce were searched at portal sites or picked up from a list of members of direct marketing traders association. Questionnaires were mailed to them, and mailed back to our office after filling out questionnaires. In addition to a report of prices of products specified on the questionnaire, we asked them to provide basic characteristics of their business such as type of organization and size, and also give answers whether they put prices different from those of sales at stores in the case they owned retail shops. Since the main results of the National Survey of Prices need to be tabulated and released first, results of the pilot survey is not yet available.

4. Price comparison between direct sales by makers and sales at electrical stores

Besides the pilot survey, price comparison between direct sales by makers and sales at electrical stores were conducted for PCs as a supplement to price comparison between online retailers and traditional retailers by the pilot survey. Four major PC makers not specializing in direct sales and a major PC maker specializing in direct sales were selected. Prices at their website for direct sales were compared with prices of the same products at six major electric stores in Tokyo as to the former four companies.

As shown in Chart, results vary from company to company. In the case of company A, prices at direct sales are identical with those of major stores except old models disappearing from the market. Old models tend to be priced much lower by major retailers. Company C is in a similar situation. However, there is a difference in that old models have already disappeared in the case of company C. One of the new models is priced much lower by major retailers instead. This model is a handy type, and considered to be special. In the case of company B, all models are priced slightly lower by major retailers. The Price gaps are constant to 3,000 yen in all models.

The cases of company A, B and C imply there is a strong relationship between prices at makers’ direct sales and major retailers, and prices are not so different at the time new models appear in the market although we speculate that retailers reduce prices afterward while makers do not change the prices. To the contrary, company D offers lower prices at its website than major retailers. One notable difference from the other three companies in its sales policy is that model numbers assigned for direct sales are different from those of sales at shops. It is possible for consumers to identify an equivalent product and compare the prices based on information accessible at the website but it is not easy work and takes time for consumers not acquainted with PCs. This fact may form the sharp contrast with the other three companies.

As for makers specializing in direct sales, it is not possible to make price comparison by matched models. Thus, we chose basic models manufactured by a major maker specializing in direct sales (company E), and estimate prices of equivalent products at electric stores using hedonic regression models used for the regular CPI compilation derived from scanner data collected at major electric stores across the country. It should be noted that sometimes a significant difference between the actual price and the estimated price comes out due to differences in characteristics which are not included in the regression model explicitly. For example, almost all note PCs made by other makers are equipped with CD-RW & DVD-ROM combo drive. Therefore, types of optical disk drive are excluded from explanatory variables in our regression model. However, basic models of the company E are equipped
with CD-ROM drive, and such a difference is not ignorable for price comparison. For this reason, we added some adjustment to estimated prices derived from the hedonic regression model. Our final estimation indicates prices are about the same with equivalent products manufactured by other companies as shown in Table 6.

**Chart: Price difference between major electric stores – direct sales by manufacturers (PCs, Feb. 9, 2003)**

Source: Prices at major electric stores were obtained from Nikkei PC Price Survey. Prices at directly sales by manufacturers were picked up from their website. Prices at 6 major electric stores in Tokyo were collected. An average is taken for each model after excluding the lowest and the highest prices, from which the corresponding price at direct sales by the maker is subtracted.

Taking all these results into consideration, as for PCs, we rather lean towards a view that price differences between direct sales by makers and sales by retailers are not so large that CPI compilers need to be worried about, with the proviso that further observation should be made particularly about companies putting a great deal of effort into direct sales such as company D and E.
5. Notes at the end

Since results of the pilot survey are not yet available, we feel it is not appropriate to draw the conclusion at this moment. Instead, we would like to mention some points. First, goods and services purchased via the internet — i.e. market basket for e-commerce — are expected to change gradually with an increase of female and elderly internet users. It is known that female users tend to prefer purchasing food and clothes via the internet while male users tend to prefer purchasing books, magazines and PCs. Therefore, price comparison covering a wider range of categories will be possibly required in the future. Next, in our experience, it is not difficult to find lower prices on the internet than major electrical stores as for PCs but usually sellers offering the good prices are not well known. As far as we know, there is no report that major electrical stores are defeated by online stores in price competition. Our impression is that direct sales by makers rather attract considerable attention. Several surveys revealed that many internet users feel insecure to spend a large amount of money for online shopping. Therefore, we may need to consider a possibility that lower prices offered by almost-unknown online shops do not necessarily have much effect on price index. The change in the market share should also be watched to evaluate the impact on price index.

Table 6: Actual price and estimated price of PCs sold by some maker specializing in direct sales (company E)

<table>
<thead>
<tr>
<th></th>
<th>Desktop</th>
<th></th>
<th>Note</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Hedonic</td>
<td>Actual</td>
<td>Hedonic</td>
</tr>
<tr>
<td>Oct. 2001</td>
<td>122,000</td>
<td>111,578</td>
<td>166,925</td>
<td>151,121</td>
</tr>
<tr>
<td>Oct. 2002</td>
<td>164,050</td>
<td>182,461</td>
<td>177,300</td>
<td>179,000</td>
</tr>
</tbody>
</table>

1) Average of 4 basic models is taken. Price of 15 inch TFT-LCD display is added in the case display is optional.
2) The corresponding prices at electric stores are estimated using hedonic regression parameters for the official CPI compilation.
3) Price of optional I/O devices such as modem, stereo speakers are added. (Other makers' PCs are usually equipped with such devices, so explanatory variables for those devices are omitted from hedonic regression model.)
4) After quality adjustment of CD-ROM drive relative to CD-RW & DVD-ROM combo drive. (All basic models are equipped with CD-ROM drive while other makers' note-PCs are usually equipped with CD-RW & DVD-ROM combo drive, so explanatory variables for various types of optical disk drive are omitted from hedonic regression.)
E-commerce and Data Capture Opportunities for Price Indexes

Andrea Grcman and Keith Woolford

Australian Bureau of Statistics

The authors would like to acknowledge the assistance provided by Neale Austen, General Manager eCatalogue Services, EAN Australia in describing their operations and providing temporary access to the public portion of their database for evaluation purposes. Any errors in interpretation are, however, our own.

Abstract: National Statistical Offices are constantly seeking ways to improve the cost effectiveness of their operations while reducing the burden on data suppliers. This paper describes some work currently being undertaken by the Australian Bureau of Statistics (ABS) to assess the opportunities for tapping into the infrastructure supporting B2B e-commerce for the purpose of obtaining data for use in constructing price indexes. The dominant system in use in Australia (if not globally) is described and the opportunities of particular interest to prices statisticians identified. Future areas for investigation at both the national and international level are outlined.

1. Introduction

In common with other National Statistical Offices (NSO's), the Australian Bureau of Statistics (ABS) is continually seeking to improve the cost effectiveness of data collection and processing while reducing the burden on data suppliers.

The ABS first investigated the possibility of using product bar-coding to assist in data collection for the Consumer Price Index (CPI) in the early 1990's. At that time the conclusion reached was that the use of bar-codes in the direct collection of data from retailers did not offer any gains over traditional price collection procedures. Although the option of obtaining aggregate scanner data\(^2\) appeared to be a promising alternative, the data costs were prohibitive. In any event, it was discovered that the use of bar-codes for product identification was not as reliable as expected due to the imperfect adherence of producers to the "standards"\(^3\).

Following these earlier setbacks, the ABS has been maintaining a watching brief on developments in this field and is now of the view that the technology and standards supporting e-commerce have matured to the point that a further detailed investigation is warranted. The current investigation is specifically targeted at assessing the opportunities for tapping into the infrastructure supporting business to business (B2B) e-commerce for the

\(^1\) The views expressed in this paper are those of the authors and do not necessarily represent the position or policies of the Australian Bureau of Statistics (ABS).

\(^2\) Derived from retailers bar-code based point of sale systems and collated by market research organisations.

\(^3\) A number of instances of changes in the quality of products not being accompanied by a change in product number were observed.
purpose of obtaining data for use in constructing price indexes (consumer and producer price indexes and cross-country Purchasing Power Parities).

The key elements of the B2B e-commerce infrastructure that are of interest are:

- Standardised numbering systems for the identification of goods, services, shipments, assets and locations;
- Standardised data carriers (e.g. bar codes) capable of presenting the standard numbering systems in a machine readable format;
- E-messaging standards to transmit data between trading partners; and
- Standardised product classifications.

Although these components of the e-commerce infrastructure have been evolving independently since the late 19th century it has only been in very recent years that they have become sufficiently integrated to support a truly viable e-commerce ideology. While the prospects for adoption of this technology across all business sectors look promising, take-up has been greatest by businesses in the retail and grocery sectors and by their downstream providers.

This paper is in four parts. The first part (section 2) describes the development/evolution of the key elements of the infrastructure; the second (section 3) describes the EANnet system in use in Australia; the third (section 4) identifies specific opportunities for statistical agencies; and the last part (section 5) briefly outlines three future streams of work.

2. A potted history

2.1 Standard data carriers

The first articulation of what we might recognise as a data carrier system for consumer goods was put forward by Wallace Flint in a master's thesis in 1932. Flint devised a scheme that would enhance the useability of punch cards in business processes. He describes a supermarket in which the consumer would perforate cards to mark selections and insert them into a reader at the checkout. This would then activate machinery to bring the purchases to them on a conveyor belt. Store managers would make a record of what was being purchased. At the time Flint’s scheme was economically unfeasible, however it largely describes today's data carrier (bar code) systems.

In the late 1940's, Bernard Silver and Norman Woodland set about devising a method for automatically capturing product information at the checkout. Woodland had many ideas ranging from ink that was sensitive to ultraviolet light, drawing dots and dashes in the sand to simulate Morse Code, through to replacing the lines with concentric circles (which became known as “bulls eye code”). On 20 October 1949 Woodland and Silver filed a patent application for the “Classifying Apparatus and Method”, describing their invention as “article classification...through the medium of identifying patterns”. In 1962 the patent was sold to Philco and then finally to RCA in 1971.

5 Bar Codes, www.inventors.about.com/library/inventors/blbar_code.htm
Throughout the 1960's David Collins was developing the use of bar code technology to keep track of rail freight information. He developed a group of orange and blue stripes made of reflective material, which could be arranged to represent the digits 0 through to 9.

Collins foresaw applications for automatic coding far beyond the railroads and in 1967 he presented the idea to his bosses. The company refused to fund further research so Collins quit the organisation and founded the Computer Identics Corporation.

Collins went on to develop "little black-and-white-line equivalents" that would be used for conveyor control and anything else that moves. By 1969, laser beam technology had become available and affordable. Collins used this technology to replace bulkier devices. The device he developed was a success and Computer Identics flourished proving the potential for bar codes in industrial settings. General Motors, Michigan was the first company to utilise a true bar coding system.

In early 1971 RCA demonstrated the use of the "bulls eye bar code" system at a grocery industry meeting. Businesses could foresee the potential of such a system and, following a series of trade association meetings, numerous technology companies were approached to develop an inter industry product coding system and associated symbolism.

IBM immediately set up a business unit to develop such a system with Woodland, the inventor of the bar code, playing a prominent role on the project.

Woodland and a fellow employee George Laurer went on to develop the Universal Product Code (UPC) incorporating an eleven-digit numbering system with symbolism based on Woodland's original two-dimensional Morse code approach with thin lines replacing the dots and thick lines replacing the dashes.

IBM's UPC symbolism worked well as any extra ink flowed out the top or bottom of the 'bars' so no information was lost. RCA continued to push the bulls eye code, however, due to printing problems and scanning difficulties, the code proved to be less effective.

As computer systems have advanced, bar codes have become more prevalent in society. There are now other bar code symbolisms designed to meet the specific needs of particular applications or industries. As a data carrier, the future of the original bar code lies in the amount of information required by the multitude of industries that are keeping track of their various business transactions. With advances in technology it is likely that the barcode will be replaced by another symbolism capable of carrying a much richer range of data which will in turn promote a rapid evolution in the broader infrastructure supporting B2B e-commerce.

### 2.2 Standard item numbering

With the adoption of the UPC as the industry standard in April 1973 the Uniform Code Council (UCC) was created to standardise the numbering system for all items that would be

---

6 A Brief History of Bar Codes, [http://www.geocities.com/SiliconValley/Campus/8351.htm](http://www.geocities.com/SiliconValley/Campus/8351.htm)
7 History of Bar Codes, [http://www.swlamall.com/WebTronics/barcodeHistory.htm](http://www.swlamall.com/WebTronics/barcodeHistory.htm)
8 Bar Codes Technology, [http://www.geocities.com/SiliconValley/Campus/8351.htm](http://www.geocities.com/SiliconValley/Campus/8351.htm)
9 History of Bar Codes, [http://www.swlamall.com/WebTronics/barcodeHistory.htm](http://www.swlamall.com/WebTronics/barcodeHistory.htm)
10 A Brief History of Bar Codes, [http://www.geocities.com/SiliconValley/Campus/8351.htm](http://www.geocities.com/SiliconValley/Campus/8351.htm)
11 Tan Jin Soon, Singapore Article Numbering Council, *An Introduction to Bar Coding*
bar coded. On June 26, 1974, at Marsh Supermarket in Troy Ohio, the first item, a packet of Wrigley’s chewing gum was scanned\(^12\).

Other countries began developing modified versions of the UPC. The most successful being the European Article Numbering (EAN) system, developed by a council made up of 12 European countries in 1974 and administered by EAN International.

The system administered by EAN spread rapidly through European countries and then to their trading partners throughout the world. The advantages of a single global standard were quickly recognised and EAN International and the UCC formed a partnership. The agreed aim of the partnership is “to develop compatible standards that will increase business efficiency by providing a common global language for trade”\(^13\).

In joint cooperation, they administer the EAN.UCC System through the management of and by providing standards for, the unique identification and communication of products, transport units, assets and locations. They aim to provide business tools that will optimise supply chain management for their members. Today EAN International and the UCC have Member Organisations (MOs) in 128 countries. Each MO administers the EAN.UCC numbering, bar coding and electronic messaging system locally. There are approximately 900,000 member organisations from a wide range of industries.

A GTIN (Global Trade Item Number) is used for the unique identification of trade items world wide within the EAN.UCC System. A GTIN has a 14 digit structure however its bar code may contain 13 digits (EAN-13), 12 digits (UPC-12) or 8 digits (EAN-8). The GTIN is defined as a 14 digit number to accommodate the different structures in use. EAN International administers the EAN numbers, while the UCC administers the allocation to organisations in North America (United States and Canada).

A major step towards facilitating global trade was the recent announcement by the UCC that all North American companies who presently scan the 12-digit UPC symbol need to be capable of scanning a 14-digit GTIN symbol by January 1, 2005\(^14\).

2.3 E-messaging

The ability to exchange information electronically is central to the effective conduct of e-commerce. In the late 1960’s and early 1970’s various e-messaging formats began to emerge in different industries and countries. It became clear to the business world that if Electronic Data Interchange (EDI) was to meet the requirements of the international community then it was essential to develop an international standard for EDI.

Early in 1980 the United Kingdom presented to the United Nations Economic Commission for Europe (UN/ECE) a standard for electronic communication between trading partners, known as the Trade Data Interchange (TDI) Standard.

At the same time the UN/ECE recognised the need to coordinate the development of one international standard and organised an international working party. In 1986 this working party commenced the challenge of combining the increasing number of European and United

---

\(^{12}\) History of Bar Codes, http://www.swlamall.com/WebTronics/barcodeHistory.htm

\(^{13}\) EAN Australia, http://www.ean.com.au

\(^{14}\) http://www.uc-council.org/2005sunrise on 08/02/03
States standards with those from the rest of the world. The following year the International Organisation for Standardization (ISO) was presented with and subsequently endorsed, the internationally accepted e-commerce syntax implementation and message design guidelines, which became known as UN/EDIFACT (United Nations Electronic Data Interchange for Administration, Commerce and Transport).

Today the UN/EDIFACT standard is still maintained and developed under the auspices of the United Nations by the Centre for the Facilitation of Procedures and Practices for Administration, Commerce and Transport (CEFACT). There are over 60 countries and numerous international organisations represented in the CEFACT, including EAN International and the UCC.

Industry support for the management of e-commerce across borders is imperative. The Global Commerce Initiative (GCI) is a global user group, created in 1999, to “improve the performance of the international supply chain for consumer goods through collaborative development and endorsement of recommended standards and key business processes”\(^\text{15}\). EAN International and the UCC are a part of this user group, supporting the GCI to better manage standards around the world.

### 2.4 Product classifications

While the use of unique product codes or GTINs, is sufficient to support bilateral e-commerce between established business partners, product codes themselves are not sufficient to facilitate e-commerce between potential new business partners. In order to locate potential suppliers, businesses require a means of searching for available products. For this purpose a product classification is required.

There would appear to be two ‘standard’ product classifications being used; the United Nations Standard Product and Services Code (UN/SPSC) and; the Universal Standard Products and Services Classification (UNSPSC). Although the authors have not been able to ascertain the extent of any material differences between these two classifications, plans are in place for unification\(^\text{16}\).

However, it is unfortunately the case that neither is an international reference classification that has been approved by the United Nations Statistical Commission (UNSC) or other competent intergovernmental board (despite the impression gained from their titles). Nor are they derived classifications based on any reference classification. They are cited as being open, non-proprietary, global standards for classifying products and services for use throughout the global market place. Users of the EANnet system in Australia classify products according to the UNSPSC.

### 2.5 Global data synchronisation

EAN International and the UCC, in conjunction with the Global Commerce Initiative (GCI), have been working to develop standards for the establishment of a Global Data Synchronisation Network (GDSN). The GDSN will enable the regional ‘master data’ synchronisation pools to be linked to provide a single global interconnected network. Master data, in this instance, refers to information that is specific to a particular item (e.g. product description, classification code, pricing, size, unique number).

\(^{15}\) [http://www.globalcommerceinitiative.org](http://www.globalcommerceinitiative.org) on 28/01/2003

\(^{16}\) UNSPSC Unification Project, [http://www.unspsc.org/](http://www.unspsc.org/)
Data synchronisation is defined as being the continuous and automated exchange of master data between trading partners. Master data sharing between trading partners is both complex and fundamental to all supply chain processes. Integrity and timeliness of master data is critical to the flow of goods, services and information throughout the chain. Sharing data effectively and efficiently relies upon access to common data definitions, data accuracy and agreement on the process used to exchange the data.

The fundamental rationale for global data synchronisation is that it enables each part of the global supply chain to be notified immediately of any changes to the agreed-upon data. The system adds a layer of visibility, by linking the data pools via a Global Registry. The Global Registry is a global service for the registration and validation of items and provides global search capability.

Global Data Synchronisation allows continuous improvement in e-commerce supply chain management practices by providing international standards for item identification, data capture, electronic messaging and the process for data exchange.

Up until recently, development work on data pools has been mainly focused on regional or national requirements. For global data synchronisation to become a reality, data pool interconnection and interoperability is essential. In October 2002, EAN International and the UCC announced that they have agreed to endorse the GLOBAL registry™ service being delivered by UCCnet. The Global Commerce Initiative (GCI) also supports this initiative.

3. The EANnet system in Australia

EAN Australia has developed a national data synchronisation and product registry service called EANnet. Within the context of the Global Data Synchronisation Network, EANnet is best described as a national data pool. This service provides an infrastructure for organisations to exchange common information while at the same time keeping commercially sensitive information maintained and secure from competitors. Within the near future, EANnet will also link to the Global Registry.

Registration for use of the EANnet service is available to financial members of EAN Australia and EAN New Zealand. There are three levels of access available: Vendor/Supplier, Buyer/Retailer, and Search and Download Only.

3.1 Vendor/supplier access (Data Source)

This type of user provides a community of trading partners with master data and includes, but is not limited to, manufacturers, importers, distributors, and wholesalers or brokers who are responsible for the supply of goods and services to their trading partners. EANnet provides them with a tool to publish a central catalogue of master data including item, price and promotional information about their GTINs. There is also the option to automatically send out changes made to their EANnet catalogue to nominated trading partners.

Suppliers are effectively able to create multiple catalogue views, each tailored to a specific trading partner. This enables trading partner specific information, for example prices, to be securely published. Suppliers are then able to grant catalogue access to their trading partners. Once access is granted an EANnet trading partnership exists between the two parties. The
trading partner is then able to view the supplier’s catalogue, including the commercially sensitive data such as purchaser specific prices and promotional information.

When a new supplier joins the network, the EANnet Client Services Team validate all data prior to it being loaded to EANnet to ensure adherence to standards such as correct use of GTIN, commodity classification, description, trading partner requirements etc. Once a supplier has been accredited, the supplier is responsible for the quality of all data maintained on EANnet and is able to add new products and update existing products when required.

3.2 Buyer/retailer (Data Recipient)

This type of user is authorised to view, use and download a set of master data, from EANnet, provided by a data source. For example: retailers, convenience stores, government departments, distributors, wholesalers, e-tailors, and electronic market places are typical data recipients. Users could also include organisations offering market research, data processing and information services.

EANnet provides buyers/retailers with the facility to browse a vendor’s product catalogue. They can also opt to automatically receive updates of changes made to the supplier’s catalogue. Data stored on the catalogue (such as GTIN, descriptions, photographs etc) can be downloaded or automatically “pushed” out to the buyers for incorporation in their internal systems (e.g. merchandising, stock management, or point of sale systems).

Buyer/retailers do not store any data on EANnet nor do they use EANnet to conduct commercial transactions.

3.3 Search and download only

This type of user consists of organisations that are neither a vendor/supplier nor a buyer/retailer but have an interest in or a demand for publicly available information contained in EANnet. Commercially sensitive information such as pricing and promotional information is not available for viewing in this domain.

3.4 System architecture

The system can be thought of as operating in both a “push” mode and a “pull” mode. The push mode is invoked by suppliers populating and maintaining their master data (pushing data in) which in turn triggers pushing updated data out to trading partners. The pull mode is invoked by buyers/retailers and the search and download community undertaking catalogue trawling activities. The strength of the system currently lies with the push mode. As use of the service grows, particularly through data synchronisation, its usefulness as a repository from which users can pull data can be expected to increase.

The following diagram illustrates the relationships between the various user communities and the system. EANnet itself is Internet enabled which supports the relationships. EANnet also supports machine to machine exchange of master data using UN/EDIFACT messaging standards.
The EANnet Registry Service currently contains information on approximately 380,000 fast moving consumer goods. However, coverage is growing rapidly as companies migrate.

### 3.5 Neutral and relationship dependent data

To protect commercially sensitive data, information contained within EANnet is classified as being either neutral or relationship dependent. There are in excess of 230 neutral and 70 relationship dependent fields.

Neutral data is generally shared between multiple trading partners and can be viewed in both the EANnet Data Synchronisation Service ("Trade" environment) and Product Registry Service ("Public" environment). Relationship dependent data is only available to nominated trading partners and can only be viewed in the EANnet Data Synchronisation Service ("Trade" environment). Some examples of neutral and relationship dependent data are provided in Table 1 below.

It is also worth noting that the catalogue supports parent-child relationships between GTINs. These relationships exist when a supplier produces different ‘bundles’ of the same product at different prices. For example, a single can of tomatoes will be assigned a unique GTIN and, if regularly supplied in packs of (say) 12 cans, the 12 can pack will also be assigned its own GTIN with the catalogue maintaining a link to its single can 'child'. This can be repeated if, for example, the 12 can packs are in turn regularly supplied in (say) 10 pack pallets. These child-parent relationships are only recognised if they are standard ordering units (produced as such with their own pricing schedules) and should not be confused with logistical units created when a customer simply orders multiple units of a particular GTIN.\(^{17}\)

---

\(^{17}\) Though a separate system of bar codes exists to manage the creation and movement of these logistical units.
Table 1: Examples of neutral and relationship dependent data

<table>
<thead>
<tr>
<th>Neutral data</th>
<th>Relationship dependent data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Identifiers</td>
<td>List Price</td>
</tr>
<tr>
<td>Product Classifications</td>
<td>Invoice Price</td>
</tr>
<tr>
<td>Dates</td>
<td>Temporary Price Reduction</td>
</tr>
<tr>
<td>Descriptions</td>
<td>Allowances</td>
</tr>
<tr>
<td>Packaging and Dimensions</td>
<td>Charges</td>
</tr>
<tr>
<td>Pallet Specifications</td>
<td>Promotional Pricing</td>
</tr>
<tr>
<td>Hazardous Information</td>
<td></td>
</tr>
<tr>
<td>Taxes</td>
<td></td>
</tr>
<tr>
<td>Pharmaceutical Information</td>
<td></td>
</tr>
<tr>
<td>Free Text Fields</td>
<td></td>
</tr>
<tr>
<td>Digital Product Images (low resolution)</td>
<td></td>
</tr>
<tr>
<td>Scan Verification Reports</td>
<td></td>
</tr>
<tr>
<td>Material Safety Data Sheets</td>
<td></td>
</tr>
</tbody>
</table>

4. Opportunities for statistical agencies

The notion of a single electronic catalogue containing detailed descriptions of all goods traded in an economy with an attached commodity classification, together with information about the supplying businesses, is the stuff of statisticians' dreams. While it would be misleading to imply that this is what exists today, it would be equally naïve to dismiss the evolutionary potential particularly given the pressure being exerted by (at least) the large retail chains in insisting that all of their suppliers are part of this system.

The EAN.UCC System has the potential to change the way that Statistical agencies collect data for the compilation of Price Indexes. There are three main benefits that this system provides:

- an infrastructure that identifies items with an internationally unique and unambiguous number;
- the ability to group like items together using a standard product classification understood by and used by businesses;
- a way of communicating to businesses using EDI; and
- a way of obtaining price information automatically through the use of data synchronisation catalogues.

Although this paper is only concerned with the potential uses for compilers of price indexes, it is not difficult to envisage additional applications in the field of business statistics. The following sections look at how this system might be used in compiling price indexes by national statistical organisations if they were to become members of the search and download community or the buyer/retailer community.

International Working Group on Price Indices - Seventh Meeting
4.1 Producer price indexes

In the short to medium term, the potential use of this system is most promising for the construction of producer price indexes (PPIs).

At the most basic level of EANnet access (search and download) the UNSPSC can be used to identify all the unique products and their suppliers for a specific commodity 'class' for use in sample design and maintenance. Although the catalogue does not contain aggregate sales data (nor indeed price data at this level of access), use of the UNSPSC by businesses provides a natural level of aggregation for businesses to provide such data directly to the statistical office. The adoption of the GTIN as an integral component of the product description for sampled items can then serve to remove any potential for ambiguity in the products for which prices are to be reported by businesses.

While the above application can serve to improve the effectiveness of sample design and maintenance operations and improve communications with data providers, it offers only marginal prospects for reducing the reporting burden on respondents.

Stepping up the level of EANnet access to buyer/retailer, offers more exciting prospects. With this level of access it should be possible to negotiate with suppliers to be granted access to their relationship dependent data (including the different catalogues furnished to different customers). Once these arrangements are negotiated, the statistical organisation could be notified of any changes to the catalogue (new goods, disappearing goods, changes in prices etc) utilising the inbuilt "push" technology at zero cost to the respondent. Direct collection of prices from the respondent would cease, while at the same time the timeliness and quality of reported price data would improve. The improvement in data quality would come about as the reliance on (normally) junior staff to interpret forms and source correct price information would be replaced by the contractually enforceable prices loaded to the catalogue.

Although this more sophisticated use of the catalogue by NSOs would require the development of additional processing systems, with the data coming from the central catalogue in a standard format for all respondents rather than from individual business systems, the benefits should outweigh the costs.

4.2 Consumer price index

The catalogue does not store any retailer specific data (items stocked or prices charged). Retailers either have data pushed out to them by the catalogue or they search the catalogue for information. As a consequence the catalogue does not offer the same prospects for automatic data collection for CPIs as it does for PPIs.

However, given that retailers are one of the key beneficiaries of the end-to-end business processes supported by the system as a whole, opportunities still exist to improve communications with retailers and to automate data capture.

Retailers are totally dependent on the GTINs to support their ordering, stock management and point of sale (POS) systems. They are loading an increasing number of catalogue data fields to their own systems (witness the move to displaying item pictures when items are scanned at POS). Their ability to source new items is also determined by the reliability of the product classification (UNSPSC) which is also incorporated in their own databases.
Adherence to the UN/EDIFACT standards should facilitate the development of software tools capable of polling retailers POS systems to report prices in respect of designated GTINs. The process of implementing such collection strategies in respect of retailers will simply be more involved (and costly) than for producers due to the more detailed individual agreements/processes involved.

In the short term, the use of the UNSPSC linked to GTINs by retailers is likely to assist in obtaining consistent, reasonably fine level sales data to assist in sample design. The incorporation of GTINs in the descriptions of items sampled should also remove ambiguity and assist in better co-ordinating item samples across the CPI and PPIs\textsuperscript{18}.

4.3 International Comparison Program

The most costly element of the International Comparison Program (ICP) is the construction of Purchasing Power Parities (PPPs). One of the most time consuming tasks in constructing PPPs is ensuring that identical products are being priced in different countries.

Global data synchronisation and the use of GTINs in PPP price collection activities can assist in a number of ways:

- if the items priced in different countries have identical GTINs, then the compilers of PPPs can be assured that the items are comparable;
- where identical GTINs are not able to be priced, the detailed descriptions contained in the catalogue can be used to ascertain whether the differences are material or not (e.g. the difference may simply be due to different power voltages in different countries);
- where the differences are judged to be material, the descriptions may assist in making quality adjustments;
- the commodity classification can be used to assist in selecting representative items; and
- the prices for items in some basic headings (producer prices) may be obtained directly from the catalogue.

5. Future projects

We believe that the potential for using this infrastructure for constructing our domestic price indexes is such that further more detailed feasibility testing is warranted. Our proposals in respect of the PPI and CPI are outlined briefly below.

We are also of the view that there are significant long term opportunities for the international statistical community which could be greatly facilitated if the commodity classification used by the business community was either a UNSC endorsed reference classification or a classification derived from a reference classification. Although this is a task beyond the capabilities of a single NSO it is also described below in the hope that it may be taken up by one of the international organisations such as the United Nations Statistics Division (UNSD).

\textsuperscript{18} For at least some commodity classes better coordinated samples may assist in deriving measures of margins.
5.1 Project 1: To evaluate EANnet as a data source for PPIs

The objective of this project is to assess the feasibility of obtaining price data for the PPIs directly from EANnet.

The first stage will be to evaluate the potential coverage currently offered by EANnet. This will involve:

- identifying any existing PPI respondents lodging master data to the system;
- comparing the number of items currently priced from these respondents with the number of items loaded to EANnet;
- matching the specific items currently priced with those on EANnet; and
- using the UNSPSC code for matched items to identify other potential items and/or suppliers of data.

The second stage will be to approach data suppliers to assess their willingness to supply data via EANnet.

If stages one and two deliver promising outcomes and it is assessed that the cost of access to EANnet can be achieved through savings on traditional data capture methods, a proposal will be put forward for the development of the required system tools.

5.2 Project 2: To evaluate the potential of the EAN.UCC system for CPIs

The objective of this project is to assess the feasibility of utilising the underlying EAN.UCC infrastructure to assist in obtaining data directly from retailers POS systems. This project will require much greater liaison with data suppliers than the PPI project, so it is not possible to be more specific about what will be required until some initial discussions take place. In any event, it is not proposed to commence this project until the PPI investigation is completed.

5.3 Project 3: Development of a derived standard product classification

The ultimate objective of this project would be to produce concordances between the UNSPSC and existing reference classifications such as the CPC and COICOP.

However, as the commercial requirements of the UNSPSC are likely to be such that the classification will be changing much more frequently than any of the reference classifications (at least at the lower levels of the hierarchy), the one-off development of a concordance by the statistical community may not have a long life. A far more effective long term solution is probably best accomplished by forming a strategic alliance with the custodians of the UNSPSC. This is something best undertaken by an international organisation.

Although statisticians can readily appreciate the potential gains to be derived from sharing a common commodity classification with the business community, the advantages will be less obvious to businesses. The current users of the UNSPSC will place a high value on the requirement that the classification be capable of being quickly updated. They will therefore be wary of any proposal, that in their view, would jeopardise this requirement. Establishing a role for the international statistical community in the maintenance of this classification will therefore not be a simple task.
Considering that the UNSPSC had its genesis in a classification developed under the auspices of the UNDP and Dunn and Bradstreet, it is the view of the authors that the UNSD is best placed to represent the international statistical community on this issue.
References


United Nations Development Program, New classification for goods and services can help developing countries compete globally through e-commerce, [Available online <www.undp.org/dpa/pressrelease/releases/P990223E.html> (accessed on 13/03/2003)


Session 7 - ILO Activities on CPIs

Chair: John Greenlees, US Bureau of Labor Statistics

Summary of session

The speaker at this session was Valentina Stoevska of the International Labour Organization (ILO). She began with a brief summary of the situation with respect to the forthcoming ILO Manual on CPIs. The bulk of her presentation consisted of a review of the draft Resolution on CPIs to be approved at the 17th International Conference of Labor Statisticians (ICLS) meeting in late November 2003.

- Manual on CPIs

Work on the Manual was directed by the Inter-secretariat Working Group on Price Indices formed in late 1998, with the assistance and advice of a Technical Expert Group. Both of these Groups contain Ottawa Group members, and the Manual was discussed at the 2001 Canberra meeting. Approximately half of the 23 chapters are theoretical in nature, with the remainder being devoted to practical aspects of CPI construction. Ms. Stoevska indicated that all chapters were expected to be finalized by the end of June 2003, and that it was expected to be printed in early fall. The latest versions of the chapters are on the ILO web site, and this electronic version is planned to be a “living document” that could be reviewed and updated as needed. Ms. Stoevska also noted that an IMF annex on structured product descriptions has been added to the Manual, along with an annex comparing how prices are used in CPIs and PPIs.

- Draft Resolution on CPIs

According to Ms. Stoevska, usual ILO practice would imply that work on a Manual would be preceded by adoption of a Resolution on CPIs at an ICLS. In this instance, however, the Resolution to be approved at the 17th ICLS meeting will benefit from the material prepared for the Manual. The ICLS meeting may also lead to some revisions or additions to the manual contents.

The present Resolution on CPIs dates from 1987. So its updating was recognized as a necessity according to recent methodological and computational developments. The proposed new Resolution was drafted in 2001 and discussed at an ILO Expert Group meeting in the fall of that year. No attempt was made to develop an Ottawa Group position on the Resolution. Rather, the discussion was aimed at providing comments that the ILO could use in developing the final form of the Resolution.

The Resolution has several purposes, notably to specify the elements of best practice in producing CPIs and to reduce incompatibilities among the series produced by different statistical agencies. It is useful to provide guidelines, particularly to developing countries. As presented at this Ottawa Group session, the Resolution has 17 sections, comprising 84 paragraphs.
The sections are:

- Preamble
- The nature and meaning of a consumer price index (CPI)
- The uses of a consumer price index
- Scope of the index
- Acquisition, use or payment
- Basket and weights
- Sampling for price collection
- Index calculation
- Elementary aggregates
- Upper level indices
- Price observations
- Collection
- Replacements
- Quality changes
- Accuracy
- Dissemination
- Consultations and integrity

In addition, there are four annexes, on Terminology and definitions; Quality adjustment methods; Types of errors; and the COICOP classification structure.

The session response to the Resolution was generally positive. Several participants congratulated the ILO and commented favorably on the content and wording. Nevertheless, a lively discussion ensued throughout the allotted time period, with numerous comments and questions presented on specific elements of the Resolution.

The Chair led with several comments. Both he and another participant argued that paragraph 15 of the Resolution should not attempt to match the acquisition, use, and payment measurement approaches to the primary purposes of the index. For example, the Chair did not agree that the “acquisition” approach is necessarily the most appropriate for an index intended to be used as macroeconomic indicator.

There were several other comments concerning the section on acquisition, use and payment in the paragraph 17 context of owner-occupied housing. Two participants pointed out that for the acquisitions approach, the relevant weight is net acquisitions, not the value of new dwellings acquired, and that the value of alterations and additions should be included. Another argued that the rental equivalence technique should be mentioned specifically.

The Chair also mentioned the issue of index revisions and corrections, which the Resolution argues against. He noted that revisions are inherent in the new U.S. superlative index. He further noted that many users would prefer to see a CPI that is revised whenever any error is found in previous data and also whenever methodological changes are made in the index.

Two participants argued that paragraph 26 should not imply that five years is the optimal interval for revising expenditure weights in a CPI. Rather, the paragraph should state that weights should be revised “at least” every five years. The manual should not be seen as arguing against more frequent review or updating of weights.
Several participants suggested that the Resolution, for example in paragraphs 2, 3, and 43, is too limiting when it contrasts a fixed-weight approach with a cost-of-living index (COLI) approach. It was argued that reflecting consumer substitution in the index is not just a COLI concept. A Divisia index could also be justified without adopting a COLI framework, and this will become more feasible as more price and quantity data become available for index construction.

Among the other specific comments, one participant noted a possible inconsistency between the draft Manual, which does not recommend the inclusion of own consumption in the scope of the index, and paragraph 16 of the Resolution. Another argued for inclusion of a discussion of metadata on quality adjustment, and for a paragraph discussing chaining and hedonic methods as alternatives to matched-model approaches in high-turnover sectors of the index. Another participant argued that sample imprecision is, in fact, recognized as a type of error in statistical science, contrary to the wording in the Resolution’s glossary. An apparent inconsistency in paragraph 43 was also noted with respect to whether elementary aggregates could be weighted. Finally, one participant argued that the wording for the selection of quality adjustment methods should be revised to include the direct comparison approach and to avoid support for explicit but subjective approaches (paragraph 67 and annex 2).

In addition to the discussion concerning the contents of the Resolution, the session also generated a number of general comments and questions. Two participants suggested that the Manual should be mentioned in the Resolution. Another, while approving of the plan to update the Manual over time as a living document, asked for further explanation of how the ILO would handle the distinctions between the paper and electronic versions of the Manual. It was also argued by one participant that the Resolution needed less technical language, since in contrast to the Manual it should have more of a layman’s orientation.

One participant, although indicating that the Resolution was very good, said that it failed to give clear guidance on several issues, notably on sources of weights (paragraph 23), treatment of seasonal items (paragraph 30), and weighting within elementary aggregates (paragraph 43). Similarly, another argued that the most difficult methodological issue concerns the treatment of services and durable goods, and that the Resolution discusses that issue only in theoretical terms. Ms. Stoevska, in her response to the discussion, recognized the difficulty in meeting the needs of both developed and less developed nations, especially given that often there is no single approved method for dealing with specific issues.

There were several stimulating comments about what role the ILO Resolution can play in view of the dynamic, sophisticated state of price measurement theory and practice at the present time, as represented in the wealth of material in the Manual. One participant stated that the Resolution, although excellent, could be thought of as a “monument of a past era” when there was much less literature on price indexes, and that to provide guidance to all countries in one Resolution was impossible. Another agreed that the Resolution should give less emphasis to the numerous complex methodological issues facing practitioners. Rather, the focus should be on regulating relationships between governments and statisticians, in particular providing a strong defense of the independence of statistical agencies.
Proposals For A Draft Resolution Concerning Consumer Price Indices

International Labour Office, Bureau of Statistics
ILO, Geneva

1. Preamble

Recalling the resolution adopted by the Fourteenth International Conference of Labour Statisticians concerning consumer price indices and recognizing the continuing validity of the basic principles recommended therein and, in particular, the fact that the consumer price index (CPI) is designed primarily to measure the changes over time in the general level of prices of goods and services that a reference population acquires, uses or pays for,

Recognizing the need to modify and broaden the existing standards in the light of recent methodological and computational developments to enhance the usefulness of the international standards in the provision of technical guidelines to all countries and particularly those with less advanced statistical infrastructure,

Recognizing the usefulness of such standards in enhancing the international comparability of the statistics,

Recognising that the consumer price index is used for a wide variety of purposes, governments should be encouraged to identify the (priority) purposes a CPI is to serve, to provide adequate resources for its compilation, and to guarantee the professional independence of its compilers, Recognising that the (priority) objectives and uses of CPI differ among countries and that, therefore, a single standard could not be applied universally,

Recognizing that the CPI needs to be credible to observers and users, both national and international, and that better understanding of the principles and procedures used to compile the index will enhance the users' confidence in the index,

Agrees that the principles and methods used in constructing a CPI should be based on the guidelines and methods that are generally accepted as constituting good statistical practices

Adopts, this ... day of ..., the following resolution which replaces the previous one adopted in 1987.

---

2. The nature and meaning of a consumer price index (CPI)

1. The CPI is a current economic indicator that is constructed to measure changes over time in the general level of prices of consumer goods and services that households acquire, use or pay for consumption.

2. The objective of the index may be to measure the change over time in the cost of purchasing a fixed basket of consumer goods and services of constant quality and characteristics, with the items in the basket being selected to be representative of households' expenditure during a year or other period. Such an index is called a fixed basket price index and may be calculated by using the Laspeyres-type index number formula. This is the type of index most commonly produced by national statistical agencies.

3. The objective of the index may also be to measure the effects of price changes on the cost of achieving a standard of living (i.e.: utility or welfare) corresponding to that achieved during a year or other specified period. Such an index is called a cost-of-living index (COLI). Operationally, a COLI cannot be directly calculated but may be approximated.

3. The uses of a consumer price index

4. The consumer price index is used for a wide variety of purposes, the two most common ones being (i) to adjust wages and government and social security benefits to compensate, partly or completely, for changes in the cost of living or in consumer prices, and (ii) to provide an average measure of price inflation for the household sector as a whole. CPI sub-indices are also used to deflate components of household final consumption expenditure in the national accounts and the value of retail sales to obtain estimates of changes in their volume.

5. CPIs are also used for purposes less directly linked to the price experiences of households. These include monitoring the overall rate of price inflation for all sectors of the economy, the adjustment of government fees and charges, and the adjustment of payments in commercial contracts. In these types of cases, the CPI is used as more appropriate measures simply do not exist, or because other characteristics of the CPI (e.g. high profile, wide acceptance, predictable publication schedule, etc) are seen to outweigh any conceptual or technical deficiencies.

6. Given that the CPI may be used for many purposes, it is unlikely to perform equally satisfactorily in all applications. It may therefore be appropriate to construct a number of alternative price indices for specific purposes, if the requirements of the users justify the extra expense. Each index should be properly named to avoid confusion and a “headline” CPI measure should be explicitly identified.

7. Where only one index is produced, it is the main use that should determine the type of index produced (fixed basket price index or COLI), the range of goods and services covered, its geographic coverage, the households it relates to, as well as to the concept of price and the formula used. If there are several major uses, it is likely that compromises may have to be made with regard to how the CPI is constructed. Users should be informed of the compromises made and of the limitations of such an index.
4. **Scope of the index**

8. The scope of the index depends on the main use for which it is intended, and should be defined in terms of the type of households, geographic areas, and the types of consumer goods and services acquired, used or paid for by the reference population.

9. If the primary use of the CPI is for adjusting money incomes, a relevant group of households, such as wage and salary earners, may be the appropriate target population. For this use, all consumption expenditures by these households, at home and abroad, should be covered. If the primary use of the CPI is to measure domestically-sourced inflation, it may be appropriate to cover consumption expenditures made within the country, rather than the expenditures of households resident within the country.

10. In general, the reference population for a national index should be defined very widely (to cover consumption expenditure of households resident in the country or consumption expenditure that take place within the country.) If any income groups, types of households or particular geographic areas are excluded, for example, for reasons of practicality or costs, then this should be explicitly stated.

11. The geographic scope refers to the geographic coverage of price collection and of consumption expenditures of the reference population and both should be defined as widely as possible, and preferably consistently. If price collection is restricted to particular areas due to resource constraints, then this should be specified. The geographic coverage of the consumption expenditure may be defined either as covering consumption expenditure of the resident population (resident consumption) or consumption expenditure within the country (domestic consumption).

12. Significant differences in the expenditure patterns and/or price movements between specific identifiable population groups or regions may exist. In such cases, separate indexes for these population groups or regions may be computed if there is sufficient demand to justify the additional cost.

13. The CPI should relate to all types of consumer goods and services of significance to the reference households, without any omission of those that may be regarded as non-essential or undesirable. Goods and services not legally available should also be included if and when possible, and if the expenditures on such items can be expected to be significant. Special aggregates may be constructed to assist those users who may wish to exclude certain categories of goods and services for particular applications or for purposes of analysis.

14. Goods and services purchased for business purposes, expenditures on assets such as works of art, financial investment (as distinct from financial services), and payments of income taxes, social security contributions and fines are not considered to be consumer goods or services and should be excluded from the coverage of the index.

5. **Acquisition, use or payment**

15. In determining the scope of the index, the time of recording and valuation of consumption, it is important to consider whether the purposes for which the index is used are best
satisfied by defining consumption in terms of "acquisition", "use", or "payment". In general, the "acquisition" approach is regarded as the most appropriate for an index intended to be used as macroeconomic indicator while the "payment" approach is the most appropriate for a compensation (an income adjustment) index. For the purposes of an index intended to measure changes in the cost of living, the "use" approach may be suitable. The decision regarding the approach to follow (or combination of these approaches) for a particular group of items should be based on the purpose of the index, as well as on the costs and the acceptability of the decision to the users who should be informed of the approach followed for different items.

16. The differences between the three approaches are most pronounced in dealing with items such as owner-occupied housing, durable goods, goods and services produced for own-account consumption, items acquired on credit, remuneration in kind and goods and services provided without charge or at prices subsidized by government and non-profit institutions serving households (NPIs).

17. The most complex and important of the items mentioned above is owner-occupied housing. In most countries, a significant proportion of households are owner-occupiers of their housing, with the housing being characterized by a long useful life and a high purchase outlay (price). Under the "acquisition" approach, the value of the new dwellings acquired in the weights reference period is used for deriving the weight (and the full price of the dwelling is included in the CPI at the time of acquisition, regardless of when the consumption is taking place). Under the "payment" approach, the weights reflect the amounts actually paid out for housing (and the prices enter the CPI in the period(s) when the prices are paid), while under the "use" approach it is the value of the flow of housing services consumed during the weights reference period estimated using an implicit or notional cost (and prices or estimated opportunity costs enter the CPI when the consumption is taking place).

18. Own-account consumption, remuneration in kind and/or goods and services provided without charge or subsidised by governments and NPIs may be important in some countries where the purpose of the index is best satisfied by defining consumption in terms of "use" or "acquisition" (under the payment approach these are out of scope). The inclusion of these items will require special valuation and pricing techniques. One option is to impute their prices and expenditures on the basis of the prices of similar goods and services bought and sold on the market. This requires increasing the weight applied to such market prices to include the value of the corresponding non-market transactions. If market prices are not relevant or cannot be reliably observed, or there is no interest in using hypothetically imputed prices, the expenditures and prices paid for the inputs into the production of these goods and services should be used instead. Users should be informed of the methods followed for the different items.

---

2 Annex 1

3 Because of the practical difficulties in uniformly defining consumption and estimating the flow of services provided by other durable goods in terms of "use", it may be necessary to adopt a mixed approach – e.g. "use" for owner-occupied housing and "acquisition" or "payments" basis for other consumer durables.
6. Basket and weights

19. Decisions on the composition of the basket and the weights follow directly from the choice of reference population and geographic coverage, as well as from the choice between the “acquisition”, “use” or “payment” approaches.

20. Once defined, the expenditures that fall within the scope of the index should be grouped into similar categories (group of products that have a common purpose (end-use) or are considered substitutes for each other) in a hierarchical classification system (e.g. divisions/groups/classes) to make the index useful for descriptive and analytical purposes. The classification used for the index compilation should be as consistent as possible with that used for household expenditure statistics and should meet the needs of users for special sub-indices. For the purposes of international comparison, the classification should also conform to, or be reconcilable with, the standard international Classification of Individual Consumption according to Purpose (COICOP), at least at its division level\(^4\).

21. In order to facilitate the analysis and interpretation of the results of the index, it may be desirable to classify goods and services according to various supplementary classifications (i.e. source of origin, durability, seasonality, etc). Calculation of the CPI by using various classifications should generate the same overall results as the original index.

22. The classification should also provide a framework for the allocation of expenditure weights. Expenditures at the lowest level of the classification system (class level), expressed as a proportion of the total expenditure, determine the weights to be used at this level. These weights stay fixed from one period of reweighting to the next. When the weights are to remain fixed for several years, the objective should be to adopt weights that are most likely to be representative of the contemporary household behaviour, rather than to precisely reflect the situation of a particular period that may have been abnormal in some way.

23. The two main sources for deriving the weights are the results from Household Expenditure Surveys (HESs) and National Accounts estimates on household consumption expenditure. The results from a household expenditure survey are more appropriate for an index defined to cover the consumption expenditures of reference population groups resident within the country, while national account estimates are likely to be more suitable for an index defined to cover consumption expenditures within the country. The decision about what source or sources to use and how they should be used depends on an analysis of their respective advantages and disadvantages and on the main purpose of the index.

24. The information from the main source (household expenditure surveys or national accounts) should be supplemented with all other available information on the expenditure pattern. Sources of such information that can be used for disaggregating the expenditures are surveys of sales in retail outlets, point-of-purchase survey, surveys of production, export and import data and administrative sources. Based on these data the weights for certain items may be further disaggregated by region and type of outlet. Where the data obtained from different sources relate to different periods, it is important to ensure, before weights are allocated, that expenditure are adjusted so that they have the same reference period.

\(^4\) Annex 4
25. Where the weight reference period differs from the price reference period, the weights should be price updated to take account of price changes between the weights reference period and price reference period. Where it is likely that price updated weights are less representative of the consumption pattern in the price reference period this procedure may be omitted.

26. Weights should be reviewed and if appropriate revised as often as accurate and reliable data are available for this to be done, preferably once every five years. Periodic revision are important to reduce the impact on the index of item substitutions. Weights at the lowest level of index classification, for at least some categories, may need to be updated more frequently as they are likely to become out of date more quickly than upper level weights.

27. When a new basket (structure and/or weights) replaces the old, a continuous CPI series should be created by linking together the index numbers based on the new basket of goods and services to those based on the earlier basket. The particular procedure used to link index number series will depend on the particular index compilation technique used. The objective is to ensure that the technique used to introduce a new basket does not, of itself, alter the level of the index.

28. Completely new types of goods and services (i.e. goods and services that cannot be classified to any of the existing expenditure classes) should normally be considered for inclusion only during one of the periodic review and re-weighting exercises. A new model or variety of an existing product that can be fitted within an existing expenditure class should be included at the time it is assessed as having a significant and sustainable market share. If a quality change is detected an appropriate quality adjustment should be made.

29. Some items like seasonal items, insurance, second hand goods, expenditure abroad, interest, own production, expenditures on purchase and construction of dwellings, etc. may need special treatment when constructing their weights. There are many alternatives for dealing with these items and the choice should be determined by national circumstances, the main purpose of the index and whether the acquisition, use and payment approach is used.

30. Seasonal items should be included in the basket. It is possible to use: (i) a fixed weight approach which use the same weight for the seasonal item in all months using an imputed price in the out-of-season months, and (ii) a variable weights approach where a changing weight is attached to the item in various months. The decision on the approach should be based on conditions in each country, but also on whether the focus is on short-term (month-to-month) changes or on the long-term index changes.

31. When second hand goods, including houses, are included in the index, then expenditure weights for these goods should be based on the net trade of such goods of the reference population group with others, e.g. dealers of such goods or households outside the reference population. Where more appropriate acquisitions without netting sales may be used for estimating weights for second hand goods.

---

5 See Annex 1
6 See Annex 2
32. When consumption from own production is within the scope of the index, the weights should include the expenditures on quantities purchased plus the value of quantities consumed from own production. Valuation of consumption from own production should be made on basis of prices prevailing on the market, unless there is some reason to conclude that market prices are not appropriate. In this case the expenditures and prices of the inputs of agricultural materials could be used instead. The third option is to valuate it by using quality adjusted market prices.

33. Expenditure classes are normally too broad to be of direct use in selecting representative products for pricing. A finer level of detail needs to be specified and decisions need to be taken concerning the items for which prices should be observed. The factors that need to be taken into account in this process include the relative significance of each item, its expected price behaviour and the practicality of obtaining prices of items at constant quality. In this selection it is important to include items that may be purchased mainly by or for all main groups in the reference population, e.g. as defined by gender, age or geographic location.

7. **Sampling for price collection**

34. A CPI is an estimate based on a sample of households to estimate weights, as well as a samples of zones within regions, a sample of outlets, a sample of goods and services, a sample of varieties and a sample of price observations.

35. The sample size and sample selection methods for both outlets and the goods and services for which prices are to be observed should ensure that the prices collected are representative and sufficient to meet the requirements for the accuracy of the index, but also that the collection process is cost-effective. The sample of price observations should reflect the importance of the goods and services available for purchase to consumers in the reference period, the number, types and geographic spread of outlets that are relevant for each good and service, and the dispersion of prices across outlets.

36. Probability sampling is the preferred method as it permits estimation of sampling variation (errors) and enables the sample size to be optimised. However, it is costly to implement and can result in the selection of items that are very difficult to price to constant quality.

37. In cases where appropriate sampling frames are lacking and it is too costly to obtain them, samples of outlets and items have to be obtained by non-probability methods. Statisticians have to use all available information and apply their best judgement to ensure that representative samples are selected. Particular attention should be given to the possibility of applying a cut-off sampling strategy, especially where the sample size will be small.

38. Efficient sampling, whether random or purposive, requires comprehensive and up to date sampling frames for outlets and products. Statistical business registers, business telephone directories, results from the point-of-purchase surveys or surveys of sales in different types of outlets may be used as sampling frames for selection of outlets. Catalogues or other product lists drawn up by major manufacturers, wholesalers or trade associations, or lists of products that are specific to individual outlets such as large supermarkets might be

7 See Annex 1
used as sampling frame for selection of products. Data scanned by bar-code readers at the cashier's desk (electronic databases) can be particularly helpful in the selection of goods and services.

39. The sample of outlets and of goods and services should be reviewed periodically and updated where necessary to maintain its representativeness.

8. Index calculation

40. CPIs are calculated in two steps. In the first step, the elementary aggregate indices are calculated. In the second step, higher level indices are calculated by aggregating the elementary aggregate indices.

9. Elementary aggregates

41. The elementary aggregate is the smallest and relatively homogeneous set of goods or services for which expenditure data are available and established (used) for CPI purposes. It is the only aggregate for which an index number is constructed without any explicit expenditure weights, although other kind of weights might be explicitly or implicitly introduced into the calculation. The set of goods or services covered by an elementary aggregate should be similar in their end-uses (i.e., close substitutes for each other) and are expected to have similar price movements. They may be defined not only in terms of their characteristics but also in terms of the type of location and outlet in which they are sold. The degree of homogeneity achieved in practice will depend on the availability of corresponding expenditure data.

42. An elementary index is a price index for an elementary aggregate. As expenditure weights usually cannot be attached to the prices/price relatives for the sampled items within the elementary aggregate, an elementary aggregate index is calculated as an un-weighted average of the prices/price relatives.

43. There are many different ways in which the prices, or the price relatives, might be averaged. The three most commonly used formulae are the ratio of arithmetic mean prices (RAP), the geometric mean (GM) and the arithmetic mean of price relatives (APR). The choice of formula depends on the need to reflect substitution within the elementary aggregate (especially with respect to COLI), degree of homogeneity, dispersion of prices and price movements, etc. It is possible to use different formulae for different elementary aggregates within the same CPI. It is recommended that the GM formula be used where possible but particularly for those elementary aggregates for which it is reasonable to assume a high degree of substitutability. In cases where consumers have only a limited ability to substitute among products or outlets or where one of the prices may become zero, it is recommended that the elementary aggregate index be constructed by using the RAP formula. For an index intended to measure the "pure" price change, RAP should be used for all elementary aggregates that are homogeneous. The APR formula should be avoided, particularly in its chained form, as it is known to result in biased estimates of the elementary aggregate indices. Wherever possible the elementary aggregate indices should be calculated as weighted averages, where the weights reflect, for example, the sales of particular items, market shares or regional weights.
44. The elementary index may be computed by using either a chained or direct form of the formula chosen. The use of a chained form generally makes the estimation of missing prices and the introduction of replacement items easier.

10. Upper level indices

45. Price indices at the class/group/division/national levels are constructed as weighted averages of elementary aggregate indices. There are many types of formulae that could be used to average the elementary aggregate indices. In order to produce a timely index, the only practical option is to use a formula that relies only on the weights observed for some past period. One such formula is the Laspeyres-type index, the formula used by most national statistical agencies. (Other alternatives would be the weighted geometric mean index and the constant elasticity of substitution (CES) approach. The latter relies on base period weights and on estimating the coefficient of substitution for all the goods in the basket. The Laspeyres and geometric mean indices are special cases of the CES formula, with zero and unitary substitution elasticities, respectively.)

46. For analytical purposes it may be appropriate to calculate the index retrospectively by using one of the index number formulas that employs both base-period weights and current period weights, such as the Fisher or Walsh index. Comparing the difference between the index of this type and the Laspeyres index can give some indication of the combined impact of income change, preference change and substitution effects over the period in question, which may be important information for producers and users of the CPI. The magnitude of the difference is expected to decline with an increase in the frequency at which the weights in the Laspeyres index are updated.

47. The upper level indices may be computed as a direct or two stage indices. The two stage index is calculated by multiplying the index for the previous period (measuring the change between base period and t-1) by the index measuring the change between t-1 and t. The later should be calculated with the price-updated expenditure weights in which the quantities of the base period are valued at the prices of the previous index period (t-1).

11. Price observations

48. The quality of the price information is the crucial determinant of the reliability of the index. Standard methods for collecting and processing price information should be developed and procedures put in place for collecting them systematically and accurately at regular intervals. Price collectors should be well trained and well supervised, and should be provided with a comprehensive manual explaining the procedures they have to follow.

12. Collection

49. An important consideration is whether the index should relate to monthly (or quarterly) average prices or to prices for a specific period of time (e.g. a single day or week in a month). This decision is related to a number of issues, which include the use of an index, the practicalities of carrying out price collection and the pattern of price movements. When the aim is a point-in-time pricing, prices should be collected over a very small number of days each month (or quarter). The interval between price observations should be uniform for each item. Since the length of the month (or quarter) varies, this uniformity needs to be defined carefully. When the aim is monthly (or quarterly) average
prices, prices, especially of the items whose prices are subject to significant variation over the month or quarter, should be collected more than once during the period.

50. Attention should also be paid to the time of the day selected for price observation. In the case of perishable goods, price observations should not be made just before closing time, as stocks may be low, or dumped to minimise wastage. In these particular cases, prices observed may not be representative.

51. Price collection should in principle be undertaken in all regions within the scope of the index, especially where there are significant differences in price movements between areas. Even when it is assessed that prices are unlikely to move differently in different areas, it may still be necessary to collect prices in all areas to maintain confidence in the index, and to monitor whether prices move in parallel. The number of price observations from each area depends on their relative significance to the CPI.

52. Prices should be collected in all types of outlets that are important, including open air markets and informal markets, and in free markets as well as price-controlled markets. Where more than one type of outlet is important for a particular type of item, an appropriately weighted average should be used in the calculation of the index.

53. Specifications should be provided detailing the variety and size of the items for which price information is to be collected. These should be precise enough to identify all the characteristics that are necessary to ensure that, as far as possible, identical goods and services are priced in successive periods in the same outlet. The specifications should include, for example, make, model, size, terms of payment, conditions of delivery, type of guarantees and type of outlet.

54. Prices to be collected are actual transaction prices, including indirect taxes and non-conditional discounts, that would be paid, agreed or costed (accepted) by the reference population. Where prices are not displayed or have to be negotiated, where quantity units are poorly defined or where actual purchase prices may deviate from listed or fixed prices, it may be necessary for the price collectors to purchase items in order to determine the transaction prices. A budget needs to be provided for any such purchases. When this is not possible, consideration should be given to interviewing customers about the prices actually paid. Tips for services, where compulsory, should be treated as part of the price paid.

55. Exceptional prices charged for stale, shop-soiled, damaged or otherwise imperfect goods sold at clearance prices should be excluded, unless the sale of such products is a permanent and widespread phenomenon. Sale prices, discounts, cut prices and special offers should be included when applicable to all customers on the date of the price observation without there being significant limits to the quantities that can be purchased by each customer.

56. In periods of price control or rationing, where limited supplies are available at prices which are held low by measures such as subsidies to the sellers, government procurement, price control, etc., such prices as well as those charged on any significant unrestricted markets should be collected. The different price observations should be combined in a way that uses the best information available with respect to the actual prices paid and the relative importance of the different types of sales.
57. Where centrally regulated or centrally fixed prices are collected from the regulatory authorities, checks should be made to ascertain whether the goods and services in question are actually sold and whether these prices are in fact paid. For goods and services where the prices paid are determined by combinations of subscription fees and piece rates (e.g. for newspapers, journals, public transport, electricity and telecommunications) care must be taken to ensure that a representative range of price offers are observed. Care must also be taken to ensure that price differences between different types of consumers are observed, e.g. those linked to the age of the purchaser or to memberships of particular associations.

58. For each type of item, different alternatives for collecting prices should be carefully investigated, to ensure that the price observations could be made reliably and effectively. Means of collection could include visits to outlets with paper forms or hand-held devices, interviews with customers, computer assisted telephone interviews, mail-out questionnaires, brochures, price lists provided by large or monopoly suppliers of services (including scanner data) and prices posted on the Internet. For each alternative, the possible cost advantages need to be balanced against an assessment of the reliability and timeliness of each of the alternatives.

59. The collected price information should be reviewed for comparability and consistency with previous observations, the presence of replacements, unusual or large price changes and to ensure that price conversions of goods priced in multiple units or varying quantities are properly calculated. Extremely large or unusual price changes should be examined to determine whether they are genuine price changes or are due to changes in quality. Procedures should be put in place for checking the reliability of all price observations. This could include a program of direct pricing and/or selective re-pricing of some items shortly after the initial observation was made.

60. Consistent procedures should be established for dealing with missing price observations because of, e.g., inability to contact the seller, non-response, observation rejected as unreliable or items temporarily unavailable. Prices of the items that are temporarily unavailable should be estimated until they re-appear or are replaced, by using appropriate estimation procedures, e.g. imputation on the basis of price changes of similar non-missing items. Carrying forward the last observed price should be avoided, especially in periods of high inflation. Treatment of the items that have permanently disappeared is discussed in paragraphs 61-63.

13. Replacements

61. Replacement of an item will be necessary when it disappears permanently from the outlet(s) where its price is observed, and may be necessary also when it is no longer available or sold in significant quantities or under normal sale conditions. Replacement should be made within the first three months (quarters) of the item becoming unavailable. Clear and precise rules should be developed for selecting the replacement item. Depending on the frequency of sampling and the potential for accurate quality adjustment, the alternatives are to select: (i) the most popular variety among those that belong to the same elementary aggregate, (ii) the most similar to the replaced variety and (iii) the variety most likely to be available in the future. Precise procedures should be laid down for price adjustments with respect to the difference in characteristics when replacements
are necessary, so that the impact of changes in quality is excluded from the observed price. Responsibility for making such price adjustments should be clearly stated.

62. Replacement of an outlet may be necessary if prices cannot be obtained e.g. because it has closed permanently or temporarily, because of a decline in representativeness or because the outlet no longer co-operates. Clear rules should be established on when to discontinue price observations from a selected outlet, on the criteria for selecting a replacement, as well on the adjustments that may be required to price observations and/or weights. Such rules should be consistent with the objectives of the index and with the way in which the price collection sample has been determined.

63. Deletion of an entire elementary aggregate will be necessary if all items in that elementary aggregate disappear from most or all outlets and it is not possible to locate a sufficient number of price observations to continue to produce a reliable index for this elementary aggregate. In such situations, it is necessary to redistribute the weight assigned to the elementary aggregate among the other elementary aggregates included in the next level of aggregation.

14. Quality changes

64. As far as possible, the same item should be priced in each period. However, in practice, items that can be observed at different time periods may differ with respect to package sizes, weights, volumes, features and terms of sale as well as other characteristics. Thus it is necessary to monitor the relevant characteristics of the items being priced to ensure that the impact of any changes in quality/utility can be excluded from the observed prices and pure price change can be estimated.

65. Identifying possible changes in quality/utility is relatively more difficult for complex durable goods and services. It is necessary, therefore, to collect a considerable amount of information on the relevant characteristics of the items for which prices are collected. Some of this information can be obtained in the course of collecting prices, but often the most important sources of information on changing characteristics will be producers, importers or wholesalers of the goods included and the study of articles and advertisements in trade publications.

66. When a quality change is detected, an adjustment must be made to the price, so that the index reflects only pure price change. If this is not done, the index will either record a price change that has not taken place or fail to record a price change that did happen. The choice of method for such adjustments will depend on the particular goods and services involved. Great care needs to be exercised because the accuracy of the resulting index depends on the quality of this process. To assume that all price change is a reflection of the change in quality, or, on the other hand, to assume that items with different qualities are essentially equivalent, should be avoided as it could cause bias.

8 Annex 1
67. The methods for estimating quality-adjusted price may be:

- **Explicit (or direct) quality adjustment methods** that directly estimate the value of the quality difference between the old and new item and adjust one of the prices accordingly. Pure price change is then implicitly estimated as the difference in the adjusted prices.

- **Implicit (or indirect) quality adjustment methods** which estimate the pure price change component of the price difference between the old and new items based on the price changes observed for similar items. The difference between the estimate of pure price change and the observed price change is considered as change due to quality difference.

Whenever possible, it is preferable to use one of the explicit quality adjustment methods as the implicit ones are generally considered to be less reliable and may result in biased estimates. As the explicit methods are more complex, difficult and costly to apply, their application should be directed to items with large weights and characteristics that change in ways that are easily described. For items whose characteristics change in ways that are difficult to describe and/or are not observable, indirect methods might be applied. Publishing estimates of the aggregate effects of any quality adjustments made would enhance the transparency of the compilation process.

15. **Accuracy**

68. CPI estimates may be subject to both sampling imprecisions and sampling and non-sampling errors arising from a variety of sources. Compilers of CPIs need to be aware of the possible sources of error, and to take steps during the index construction and compilation processes to minimise their impact.

69. The following are some well-known sources of potential error, either in pricing or in index construction, that over time can lead to significant errors in the overall CPI: incorrect selection of items and incorrect observation and recording of their prices; failure to observe and adjust correctly for quality changes, appearance of new goods and outlets; failure to adjust for item and outlet substitution (in a COLI index) or loss of representativity (in a “pure” price index); the use of inappropriate formula(s) for computing elementary aggregate and upper level indices.

70. In general, regularly updating weights and baskets, employing unbiased elementary aggregate formulae, making appropriate adjustments for quality change, allowing adequately and correctly for new products, and taking proper account of substitution issues (in a COLI index) as well as quality control of the entire production process will minimise the index’s potential for giving a misleading picture.

16. **Dissemination**

71. The CPI estimate should be computed and publicly released as quickly as possible after the end of the period to which it refers, and according to a pre-announced timetable. It should be made available to all users at the same time, in a convenient form, and should be accompanied by a short methodological explanation. Rules relating to its release should be made publicly available and strictly observed. In particular, they should include...
72. The index should be produced and released monthly. Where countries do not have the necessary resources and/or there is no strong user demand for a monthly series, the CPI should be prepared and released with quarterly or semi-annual periodicity.

73. The press release presenting CPI results should show the index level from the index reference period. It is also useful to present derived indices, such as the one that shows changes in the major aggregates between (i) the current month and the previous month, (ii) the current month and the same month of the previous year and (iii) the average of the latest 12 months and the average of the previous 12 months. The indices should be presented in both seasonally adjusted and unadjusted terms, if seasonally adjusted data are available.

74. Sub-indices should also be produced and released. Consideration should be given to producing sub-indices for the divisions and groups of the UN Classification of Individual Consumption according to Purpose (COICOP). Sub-indices for different regions or socio-economic groups, and alternative indices designed for analytical purposes, may be produced and publicly released if they are judged to be reliable and their preparation is cost effective.

75. Comments and interpretation of the index should accompany its publication to assist users. An analysis of the contributions of various items or group of items to the overall change and an explanation of any unusual factors affecting the price changes of the major contributors to the overall change should be included.

76. The index reference period should be changed as frequently as necessary to ensure that the index numbers remain easy to present and understand. The index reference period may be chosen to coincide with the latest weights reference period or it could be established to coincide with the base period of other statistical series.

77. Average prices and price ranges for important and reasonably homogeneous items may be estimated and published in order to support the research and analytical needs of users.

78. Retrospective corrections (for example, as a result of an error in the price observations or in the calculations) of the publicly released indices should only be made when the index estimate previously presented was sufficiently distorted to be of concern to users. Corrections should be made as soon as possible after detection and a press release prepared and released to explain the differences.

79. In order to ensure public confidence in the index, a full description of the data collection procedures and the index methodology should be published and made widely available. The documentation should include an explanation of the main objectives of the index, details of the weights, the index number formulas used, and a discussion of the precision of the index estimates, even if this is based only on subjective assessments. The precise identities of the outlets and goods and services used for price collection should not be revealed, as this could influence the representativeness of the index.
80. Users should be warned in advance of any changes that are going to be made to the scope, weights or methodology used to estimate the CPI.

17. Consultations and integrity

81. The compiling agency should have the professional independence, competence and resources necessary to produce a high quality CPI. The United Nations “Fundamental Principles of Official Statistics”\(^\text{11}\) and the ILO “Guidelines concerning dissemination practices for labour statistics”\(^\text{12}\) should be respected.

82. The agency responsible for the index should consult representatives of users on issues of importance for the CPI, particularly during preparations for any changes to the methodology used in compiling the CPI. One way of organizing such consultations is through the establishment of advisory committee(s) on which social partners as well as other users and independent experts might be represented.

83. Comparing CPI movements across countries is difficult because of the different measurement approaches used by countries of certain items, particularly housing and financial services. The exclusion of housing (actual rents and either imputed rents or acquisition of new houses, and maintenance and repair of dwelling) and financial services from the all-items index will make the resulting estimates of price change for the remaining items more comparable across countries. Therefore, in addition to the all-items index, countries should, if possible, produce and provide for dissemination to the international community an index that excludes housing and financial services. It should be emphasised, though, that even for the remaining items in scope, there are still be difficulties when making international comparisons of changes in consumer prices.

84. Countries should report national CPI results and methodological information to the International Labour Office in the format and at the frequency requested, and as soon as possible after the national release of the corresponding results.

\(^\text{11}\) UN Economic and Social Council, 1994
\(^\text{12}\) 16th International Conference of Labour Statisticians, 1998
Annex 1: Terminology and definitions

- "Consumer goods" are goods or services that are used by households for the direct satisfaction of individual needs or wants.

- "Consumption expenditures" are expenditure on consumer goods and services and can be defined in terms of "acquisition", "use", or "payment":

  - "acquisition" indicates that it is the total value of the goods and services acquired during a given period that should be taken into account, whether through purchase, own account production or as a social transfer in kind received from government or non-profit institutions, irrespective of whether they were wholly paid for or used during the period; The prices enter the CPI in the period when consumers accept or agree prices, as distinct from the time payment is made.

  - "use" indicates that it is the total value of all goods and services actually consumed during a given period that should be taken into account; for durable goods this approach requires valuing the services provided by these goods during the period; The prices (opportunity costs) enter the CPI in the period of consumption.

  - "payment" indicates that it is the total payment made for goods and services during a given period that should be taken into account, without regard to whether they were delivered or used during the period. The prices enter the CPI in the period or periods when the payment is made.

- "Scope of the index" refers to the population groups, geographic areas, items and outlets for which the index is constructed.

- "Coverage" of the index is the set of goods and services whose prices are observed for inclusion in the index. For practical reasons, coverage may have to be less than what corresponds to the defined scope of the index.

- "Reference population" refers to that specific population group for which the index has been constructed.

- "Weights" are the aggregate consumption expenditures on any set of goods and services expressed as a proportion of the total consumption expenditures on all goods and services within the scope of the index in the weight reference period. They are a set of numbers summing to unity.

- "Price updating of weights" is a procedure that is used to bring the expenditure weights in line with the index or price reference period. The price updated weights are calculated by multiplying the weights from the weight reference period by elementary indices measuring the price changes between weight reference and price reference period and rescaling to sum to unity.

- "Index reference period" is the period for which the value of the index is set at 100.0. This may be a year, a quarter or an individual month.
• "Price reference period" is the period whose prices the prices in the current period are compared. The period whose prices appear in the denominators of the price relatives.

• The "weight reference period" is the period, usually a year, whose estimates of the volume of consumption and its components are used to calculate the weights.

• Probability sampling is the selection of a sample of units, such as outlets or products, in such a way that each unit in the universe has a known non-zero probability of selection.

• Cut-off sampling is a sampling procedure in which a predetermined threshold is established with all units in the relevant population at or above the threshold being eligible for inclusion in the sample and all units below the threshold being excluded. The threshold is usually specified in terms of the size of some relevant variable (such as some percentage of total sales), the largest sampling units being included and the rest excluded.

• Imputed expenditures are the expenditures assigned to an item that has not been purchased, such as an item that has been produced by the household for its own consumption (including housing services produced by owner-occupiers), an item received as payment in kind or as a free transfer from government or non-profit institution.

• Imputed price refers to the estimated price of an item whose price during a particular period has not been observed and is therefore missing. It is also the price assigned to an item for which the expenditures have been imputed, see above.

• "Outlet" indicates a shop, market stall, service establishment, internet seller or other place where goods and/or services are sold or provided to consumers for non-business use.

• "Linking" means joining together two consecutive sequences of price observations, or price indices, that overlap in one or more periods, by rescaling one of them so that the value in the overlap period is the same in both sequences, thus combining them into a single continuous series.

• "Price" is defined as the value of one unit of a product, for which the quantities are perfectly homogeneous not only in a physical sense but also in respect of a number of other characteristics.

• "Pure price change" is that change in the price of a good or service which is not due to any change in its quality. When the quality does change, the pure price change is the price change remaining after eliminating the estimated contribution of the change in quality to the observed price change.

• "Quality adjustment" refers to the process of adjusting the observed prices of an item to remove the effect of any changes in the quality of that item over time so that pure price change may be identified.

• "Consumer substitution" occurs when, faced with changes in relative price, consumers buy more of the good that has become relatively cheaper and less of the good that has become relatively more expensive. It may occur between varieties of the same item or between different expenditure categories.
Annex 2: Quality adjustment methods

Implicit quality adjustment methods

- The “overlap” method assumes that the entire price difference at a common point in time between the disappearing item and its replacement is due to a difference in quality.

- The “overall mean imputation” method first calculates the average price change for the elementary aggregate without the disappearing item and its replacement, and then uses that rate of price change to impute a price change for the disappearing item. It assumes that the pure price difference between the disappearing item and its replacement is equal to the average price changes for continuing (non-missing) items.

- The “class mean imputation” method is a variant of the overall mean imputation method. The only difference is in the source of the imputed rate of price change to period t+1 for the disappearing item. Rather than using the average index change for all the non-missing items in the elementary aggregate, the imputed rate of price change is estimated using only those price changes of the items that were judged essentially equivalent or were directly quality-adjusted.

Explicit quality adjustment methods

- The “expert’s adjustment” method relies on the judgement of one or more industry experts, commodity specialists, price statisticians or price collectors on the value of any quality difference between the old and replacement product. None, some, or all of the price difference may be attributed to the improved quality.

- The “differences in production costs” approach relies on the information provided by the manufacturers on the production costs of new features of the replacements (new models), to which retail mark-ups and associated indirect taxes are then added. This approach is most practicable in markets with a relatively small number of producers, with infrequent and predictable model updates. However, it should be used with caution as it is possible for new production techniques to reduce costs while simultaneously improving quality.

- The “quantity adjustment” method is applicable to items for which the replacement item is of a different size to the previously available one. It should only be used if the differences in quantities do not have an impact on the quality of the good.

- The “option cost” method adjusts the price of the replacements for the value of the new observable characteristics. An example of this is the addition of a feature that earlier has been a priced option as standard to a new automobile model.

- A hedonic regression method estimates the price of an item as a function of the characteristics it possesses. The relationship between the prices and all relevant and observable price-determining characteristics is first estimated and then results are used to estimate the effects of changes in these characteristics on prices.
Annex 3: Types of errors

- "Quality change error" is the error that can occur as a result of the index's failure to make proper allowance for changes in the quality of goods and services.

- "New goods error" is the failure to reflect either price changes in new products not yet sampled, or given a COLI objective, the welfare gain to consumers when those products appear.

- "Outlet substitution error" can occur when consumers shift their purchases among outlets for the same item without proper reflection of this shift in the data collection for the index. Most relevant when trying to estimate a COLI.

- "New outlets error" is conceptually identical to new goods error. It arises because of the failure to reflect either price changes in new outlets not yet sampled, or the welfare gain to consumers when the new outlets appear.

- "Upper level substitution error" arises when the index does not reflect consumer substitution among the basic categories of consumption owing to the use of an inappropriate method for aggregating elementary aggregates in the construction of the overall index value. Only relevant to a COLI, although an equivalent (representativity error) may be defined from the perspective of the pure price index.

- "Elementary index error" arises from the use of an inappropriate method for aggregating price quotations at the very lowest level of aggregation. The elementary index error can take two forms: formula error and lower level substitution error. The index suffers from formula error if, as a result of the properties of the formula, the result produced is biased relative to what would have been the result if a pure price change could have been estimated. The index suffers from lower level substitution error if it does not reflect consumer substitution among the items contained in the elementary aggregate. Lower level substitution error is only relevant to a COLI.

- "Sampling imprecision " is not an error, but a consequence of the CPI being based on samples and not on a complete enumeration of the population.

- "Selection error" arises when the sample of price observations is not fully representative of the intended population of outlets and/or items. The first four types of errors listed above can be seen as special cases of this type of error.
Annex 4: Classification of Individual Consumption According to Purpose (COICOP)\(^\text{13}\)

<table>
<thead>
<tr>
<th>01 - 12</th>
<th>Individual Consumption Expenditure of Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.0</td>
<td>Food and non-alcoholic beverages</td>
</tr>
<tr>
<td>01.1.0</td>
<td>Food</td>
</tr>
<tr>
<td>01.2.0</td>
<td>Non-alcoholic Beverages</td>
</tr>
<tr>
<td>02.0</td>
<td>Alcoholic beverages and tobacco</td>
</tr>
<tr>
<td>02.1.0</td>
<td>Alcoholic Beverages</td>
</tr>
<tr>
<td>02.2.0</td>
<td>Tobacco</td>
</tr>
<tr>
<td>02.3.0</td>
<td>Narcotics</td>
</tr>
<tr>
<td>03.0</td>
<td>Clothing and Footwear</td>
</tr>
<tr>
<td>03.1.0</td>
<td>Clothing</td>
</tr>
<tr>
<td>03.2.0</td>
<td>Footwear</td>
</tr>
<tr>
<td>04.0</td>
<td>Housing, Water, Electricity, Gas and Other Fuels</td>
</tr>
<tr>
<td>04.1.0</td>
<td>Actual Rentals for Housing</td>
</tr>
<tr>
<td>04.2.0</td>
<td>Imputed Rentals for Housing</td>
</tr>
<tr>
<td>04.3.0</td>
<td>Maintenance and Repair of the Dwelling</td>
</tr>
<tr>
<td>04.4.0</td>
<td>Water Supply and Miscellaneous Services Related to the Dwelling</td>
</tr>
<tr>
<td>04.5.0</td>
<td>Electricity, gas and other fuels</td>
</tr>
<tr>
<td>05.0</td>
<td>Furniture, Household Equipment and Routine Household Maintenance</td>
</tr>
<tr>
<td>05.1.0</td>
<td>Furniture and Furnishing, Carpets and Other Floor Coverings</td>
</tr>
<tr>
<td>05.2.0</td>
<td>Household Textiles</td>
</tr>
<tr>
<td>05.3.0</td>
<td>Household Appliances</td>
</tr>
<tr>
<td>05.4.0</td>
<td>Glassware, Tableware and Household Utensils</td>
</tr>
<tr>
<td>05.5.0</td>
<td>Tools and Equipment for House and Garden</td>
</tr>
<tr>
<td>05.6.0</td>
<td>Goods and Services for Routine Household Maintenance</td>
</tr>
<tr>
<td>06.0</td>
<td>Health</td>
</tr>
<tr>
<td>06.1.0</td>
<td>Medical products, Appliances and Equipment</td>
</tr>
<tr>
<td>06.2.0</td>
<td>Outpatient Services</td>
</tr>
<tr>
<td>06.3.0</td>
<td>Hospital Services</td>
</tr>
<tr>
<td>07.0</td>
<td>Transport</td>
</tr>
<tr>
<td>07.1.0</td>
<td>Purchase of Vehicles</td>
</tr>
<tr>
<td>07.2.0</td>
<td>Operation of Personal Transport Equipment</td>
</tr>
<tr>
<td>07.3.0</td>
<td>Transport Services</td>
</tr>
<tr>
<td>08.0</td>
<td>Communication</td>
</tr>
<tr>
<td>08.1.0</td>
<td>Postal Services</td>
</tr>
<tr>
<td>08.2.0</td>
<td>Telephone and Telefax Equipment</td>
</tr>
<tr>
<td>08.3.0</td>
<td>Telephone and Telefax Services</td>
</tr>
<tr>
<td>09.0</td>
<td>Recreation and culture</td>
</tr>
<tr>
<td>09.1.0</td>
<td>Recreational Equipment and Accessories</td>
</tr>
<tr>
<td>09.2.0</td>
<td>Recreational and Cultural Services</td>
</tr>
<tr>
<td>09.3.0</td>
<td>Newspapers, Books and Stationery</td>
</tr>
<tr>
<td>09.4.0</td>
<td>Package holidays</td>
</tr>
<tr>
<td>10</td>
<td>Education</td>
</tr>
<tr>
<td>10.1.0</td>
<td>Educational Services</td>
</tr>
<tr>
<td>11.0</td>
<td>Hotel, Cafes and Restaurants</td>
</tr>
<tr>
<td>11.1.0</td>
<td>Catering services</td>
</tr>
<tr>
<td>11.2.0</td>
<td>Accommodation services</td>
</tr>
<tr>
<td>12.0</td>
<td>Miscellaneous Goods and Services</td>
</tr>
<tr>
<td>12.1.0</td>
<td>Personal Care</td>
</tr>
<tr>
<td>12.2.0</td>
<td>Prostitution</td>
</tr>
<tr>
<td>12.3.0</td>
<td>Personal Effects n.e.c.</td>
</tr>
<tr>
<td>12.4.0</td>
<td>Social Protection</td>
</tr>
<tr>
<td>12.5.0</td>
<td>Insurance</td>
</tr>
<tr>
<td>12.6.0</td>
<td>Financial services n.e.c.</td>
</tr>
<tr>
<td>12.7.0</td>
<td>Other services n.e.c</td>
</tr>
</tbody>
</table>

\(^{13}\) Explanatory notes are available on [http://esa.un.org/unsd/cr/registry/reqcst.asp?Cl=5&Lg=1](http://esa.un.org/unsd/cr/registry/reqcst.asp?Cl=5&Lg=1)
Session 8 - Future directions, next meeting

Chair: Timo Koskimäki, Keith Woolford

Summary of the session

The eighth meeting of the Ottawa Group is to be held in Helsinki, Finland 23rd to 25th August, 2004.

A list of potential topics for the next meeting was prepared by the Steering Committee for consideration of the participants. The list included following seven themes:

- **Price indices for services**

  The price index problems relating to services have been on the agenda for both the sixth meeting (Canberra) and the seventh meeting (Paris). Although there is potential for some overlap with the Voorburg Group, the Steering Committee is of the view that price indices for services should remain on the agenda of the Ottawa Group. In particular, research papers relating to financial services, insurance services and gambling services would be considered relevant for the eighth meeting.

- **Relationships between PPI’s and CPI’s**

  The remit of the Ottawa group states, that the focus of the Group is particularly, though not exclusively, in the area of consumer price indices. The development of producer price indices for services will bring the field of producer price indices closer to CPI’s which should provide greater opportunities for integration of methods and practices. The use of indices for components of the CPI and PPI in the deflation of aggregates in the national accounts also suggests a need for similar approaches (the issue of the most appropriate index formula is particularly relevant here).

- **Housing**

  The weight of housing services – owner occupied housing and rental markets – is very significant in most CPI’s. Owner occupied housing is also treated quite differently in different countries due to the existence of alternative conceptual approaches. On a more technical level, the complexity of the housing market presents a number of challenges for index compilation, especially in relation to making appropriate adjustments for changes in quality. This applies both to owner occupied housing and rental markets.

- **Health**

  Prices for health services often include insurance-like components and subsidised or regulated pricing schemes. Also technical innovations are likely to result in quality improvements which are, at best, imperfectly accounted for in most health services price indices. Although price indices for health were on the agenda of the sixth meeting of the Group, more research in this field would be welcome.
• Sampling

The production of CPI’s typically involves making choices about how outlets and items are to be sampled. Different sampling strategies are often used within a single index. Some of the more common approaches include variations of multi-stage probability sampling, quota sampling and judgemental sampling. The evaluation of these alternative approaches is quite challenging and empirical research directed at providing tools for evaluating different sampling schemes would be welcomed.

• Quality assurance of price indices

In less complex statistical systems there often exist some generally accepted measures like non-response rates and statistical variance estimators to monitor bias and statistical accuracy of the statistics. In the case of price indices, quality assurance is more difficult. The results of any work in developing systematic approaches for producing quality measures for price indices would be welcomed.

• Price index data processing – automation of the production process

The compilation of price indices often involves the processing of a considerable number of individual price observations. The decisions relating to data processing and validation are always to some extent automated, i.e. based on some pre-defined algorithms. The degree of process automation is known to vary across countries and across individual price indexes. Increasing the degree of automation has the potential to reduce costs, but may increase the risk of introducing systematic biases. A useful starting point for considering these issues would be to collate a systematic account of existing automated practices and statistical methods for risk assessment related to automated procedures.

During the general discussion the list of topics presented by the Steering Committee was generally welcomed. The following clarifications and additional topics were proposed during the discussion:

• Topic 2 - relationships between PPI and CPI – might be extended to spatial comparisons such as PPP’s as well.

• E-commerce, already on the agenda of the Paris meeting, should be followed up. The topic might be extended to cover more general issues associated with globalisation of commerce like cross-border purchases, relationship to National accounts and statistics on exports and imports.

• How to measure sampling error in CPI’s.

• Treatment of transfer prices between different units of the same companies operating in different countries.

• Quality of CPI weights and sources for weights.

Participants also discussed the general role of the Ottawa Group and the ways in which the group contributes to the international statistical community. Generally, the work of the Ottawa Group was seen as important and good practices presented in the meetings of the
Group have been adopted as standards by statistical agencies and institutions. However, there was a general agreement that the Ottawa Group should, if possible, provide more formal feedback on its activities. It was proposed and accepted by the meeting that, once the new international manuals on price indices have been published, future discussions of the Group should take into account any relevant sections of the manuals. The intention is that the Ottawa Group will provide an ongoing forum for identifying areas of the manuals that require updating.

The meeting also wished to record the following:

- Its appreciation of the new ILO Manual on Consumer Price Indices and its thanks to all that have contributed to the work.

- Its endorsement of the importance of ensuring that the new ILO manual and the forthcoming new ILO resolution on Consumer Price Indices be consistent.
Appendix 1 - agenda of the meeting

Tuesday 27 May 2003

08:30 - 09:10 REGISTRATION

09:10 - 09:45 OPENING EVENTS

- Welcome speech, Jean-Michel Charpin, INSEE
- Opening remarks, Thierry Lacroix, INSEE

09:45 - 10:30 SESSION 1 - Coping with changing to complex pricing schemes

Chair: George Beelen

- Paul Haschka, Statistik Austria, Simple Methods of Explicit QA for Services in Complex Pricing Schemes
- Alina Gluchowska, CSO Poland, Pricing some complex products for the CPI needs (based on the Polish practice)
- François Magnien, INSEE, Measuring the price change of mobile phone services: an arduous task

10:30 - 10:50 COFFEE BREAK

10:50 - 12:15 SESSION 1, continued

12:15 - 13:45 LUNCH

13:45 - 15:15 SESSION 2 - Use of hedonic regression

Chair: Mick Silver

- Jan de Haan, Statistics Netherlands, Time Dummy Approaches to Hedonic Price Measurement
- Erwin Diewert, University of British Columbia, Canada, Hedonic Regressions: A Review of Some Unresolved Issues

15:15 - 15:35 COFFEE BREAK
15:35 - 17:45 SESSION 3 - Elementary aggregation, superlative indexes

Chair: Keith Woolford

- Mick Silver, Cardiff University, *Why Price Index Number Formulae Differ: Economic Theory and Evidence on Price Dispersion*
- John Greenlees, BLS, *Introducing the Chained Consumer Price Index*

18:00 - 19:30 COCKTAIL RECEPTION

- Salons de l’Aveyron

Wednesday 28 May 2003

08:45 - 09:00 COFFEE, PASTRIES

09:00 - 10:20 SESSION 4 - Financial services, including insurances

Chair: Thierry Lacroix

- Lionel Viglino, Eurostat, *Insurance and quality adjustment: excess and option-cost method*
- Martin Ribe, Statistics Sweden, *Financial Services in Swedish Price Indices*

10:20 - 11:00 PHOTO OF THE GROUP AND COFFEE BREAK

11:00 - 12:30 SESSION 5 - New products, substitution between products and outlets

Chair: Bert M. Balk

- Timo Koskimäki, Statistics Finland, *Segmented Markets and CPI Elementary Classifications*
- Rósmundur Guðnason, Statistics Iceland, *How do we measure inflation? Some measurement problems*
- Jean-Claude Roman, Eurostat, *On the treatment of newly significant goods and services in the HICP*

12:30 - 14:00 LUNCH
14:00 - 14:40  SESSION 5, continued

14:40 - 15:10  SESSION 6 - E-commerce

Chair: Rósmundur Guônason

- David Fenwick, ONS, Internet retail channels in Price Indices: the challenges involved in including non-traditional retailers in the UK Retail Prices Index
- Yoel Finkel, Central Bureau of Statistics, Israel, E-commerce in the Israeli CPI
- Robin Lowe, Statistics Canada, E-commerce and consumer behaviour
- Masato Okamoto, National Statistics Center, Japan, The Present Situation of E-Commerce and a Pilot Survey on E-Commerce
- Keith Woolford, Australian Bureau of Statistics, E-commerce and Data Capture Opportunities for Price Indexes

15:10 - 15:30  COFFEE BREAK

15:30 - 16:30  SESSION 6, continued

19:45 - DINNER

Restaurant « L'O à la bouche »

Thursday 29 May 2003

08:45 - 09:00  COFFEE, PASTRIES

09:00 - 10:30  SESSION 6, continued

10:30 - 10:50  COFFEE BREAK

10:50 - 12:30  SESSION 7 - ILO activities on CPIs

Chair: John Greenlees

- ILO Manual: release and updating process, Valentina Stojevska, ILO
- Valentina Stojevska, ILO, Proposals for a draft resolution concerning Consumer price indices
12:30 - 14:00  **LUNCH**

14:00 - 15:20  **SESSION 8 - Future directions, next meeting**

Chair: Timo Koskimäki, Keith Woolford

15:20 - 15:30  **Closing of the meeting**

- Closing remarks, Thierry Lacroix, INSEE
## Appendix 2 - List of participants

### Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Organisation</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALLARD-SAULNIER Marie</td>
<td>Statistics Canada</td>
<td>Jean Talon Building, 13th Floor Ottawa, Ontario K1A 0T6 CANADA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:marie.saulnier@statcan.ca">marie.saulnier@statcan.ca</a></td>
</tr>
<tr>
<td>ASTIN John</td>
<td>Consultant, UK</td>
<td>Broadwater Consulting Ltd 27 St Jacob's Place, Canterbury Kent CT1 3TU</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED KINGDOM <a href="mailto:astinjohn@hotmail.com">astinjohn@hotmail.com</a></td>
</tr>
<tr>
<td>BADJI Yankhoba Jacques</td>
<td>Direction de la Prévision et de la Statistique, Sénégal</td>
<td>BP 116 Dakar Dakar SENEGAL <a href="mailto:jacobbadji@yahoo.com">jacobbadji@yahoo.com</a></td>
</tr>
<tr>
<td>BALK Bert M.</td>
<td>Statistics Netherlands</td>
<td>Prinses Beatrixlaan, 428 P.O. Box 4000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2270 JM Voorburg THE NETHERLANDS <a href="mailto:bblik@cbs.nl">bblik@cbs.nl</a></td>
</tr>
<tr>
<td>BEELEN George</td>
<td>Statistics Canada</td>
<td>Jean Talon Building, 13th Floor Ottawa, Ontario K1A 0T6 CANADA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:George.Beelen@statcan.ca">George.Beelen@statcan.ca</a></td>
</tr>
<tr>
<td>BIGGERI Luigi</td>
<td>ISTAT, Italy</td>
<td>Via Cesare Balbo 18 00184 - Rome ITALY</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:biggeri@istat.it">biggeri@istat.it</a></td>
</tr>
<tr>
<td>BRANCHI Mariagnese</td>
<td>ECB, Frankfurt am Main</td>
<td>Kaisersstrasse, 29 Postfach 160319</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D-60066 Frankfurt am Main GERMANY <a href="mailto:mariagnese.banchi@ecb.int">mariagnese.banchi@ecb.int</a></td>
</tr>
<tr>
<td>CORTI Marcello</td>
<td>Office Fédéral de la Statistique, Suisse</td>
<td>10 Espace de l’Europe CH-2010 Neuchâtel SWITZERLAND</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:marcello.corti@bfs.admin.ch">marcello.corti@bfs.admin.ch</a></td>
</tr>
<tr>
<td>De HAAN Jan</td>
<td>Statistics Netherlands</td>
<td>Prinses Beatrixlaan, 428 P.O. Box 4000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2270 JM Voorburg THE NETHERLANDS <a href="mailto:jhhn@cbs.nl">jhhn@cbs.nl</a></td>
</tr>
<tr>
<td>DIEWERT Erwin</td>
<td>University of British Columbia, Canada</td>
<td>4169 Yuculta Crescent Vancouver, B. C. CANADA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:diewert@econ.ubc.ca">diewert@econ.ubc.ca</a></td>
</tr>
<tr>
<td>Name</td>
<td>Organization</td>
<td>Address/Contact Information</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>EHARA Katsuhisa</td>
<td>National Statistics Center, Japan</td>
<td>19-1 Wakamatsu-cho Shinjuku-ku Tokyo 162-8668 JAPAN <a href="mailto:kehara@nstac.go.jp">kehara@nstac.go.jp</a></td>
</tr>
<tr>
<td>ENGELHARDT Nadin</td>
<td>Destatis, Germany</td>
<td>Gustav-Stresemann Ring 11 D-65189 WIESBADEN GERMANY <a href="mailto:nadin.engelhardt@destatis.de">nadin.engelhardt@destatis.de</a></td>
</tr>
<tr>
<td>FENWICK David</td>
<td>Office for National Statistics, UK</td>
<td>1 Drummond Gate, Room D2-19A London SW1V 2QQ UNITED KINGDOM <a href="mailto:david.fenwick@ons.gov.uk">david.fenwick@ons.gov.uk</a></td>
</tr>
<tr>
<td>FINKEL Yoel</td>
<td>Central Bureau of Statistics, Israel</td>
<td>66, Kanfei Nesharim Street P.O. Box 13015 91130 Jerusalem ISRAEL <a href="mailto:yoel@cbs.gov.il">yoel@cbs.gov.il</a></td>
</tr>
<tr>
<td>GLUCHOWSKA Alina</td>
<td>Central Statistical Office, Poland</td>
<td>Al. Niepodleglosu 208 00-925 Warsaw POLAND <a href="mailto:a.gluchowska@stat.gov.pl">a.gluchowska@stat.gov.pl</a></td>
</tr>
<tr>
<td>GUDNASON Rósmundur</td>
<td>Statistics Iceland</td>
<td>Borgartun 21a 150 Reykjavik ICELAND <a href="mailto:rosmundur.gudnason@hagstofa.is">rosmundur.gudnason@hagstofa.is</a></td>
</tr>
<tr>
<td>GREENLEES John</td>
<td>Bureau of Labor Statistics, USA</td>
<td>2 Massachusetts Avenue NE, Washington, DC, 20212 USA <a href="mailto:Greenlees_J@bls.gov">Greenlees_J@bls.gov</a></td>
</tr>
<tr>
<td>GUEDES Dominique</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:dominique.guedes@insee.fr">dominique.guedes@insee.fr</a></td>
</tr>
<tr>
<td>HASCHKA Paul</td>
<td>Statistik Austria</td>
<td>Hintere Zollamtstrasse 2b A-1033 Vienna AUSTRIA <a href="mailto:Paul.Haschka@statistik.gv.at">Paul.Haschka@statistik.gv.at</a></td>
</tr>
<tr>
<td>KOSKIMÄKI Timo</td>
<td>Statistics Finland</td>
<td>P. O. B. 5A FI-00022 Helsinki FINLAND <a href="mailto:timo.koskimaki@stat.fi">timo.koskimaki@stat.fi</a></td>
</tr>
<tr>
<td>LACROIX Thierry</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:thierry.lacroix@insee.fr">thierry.lacroix@insee.fr</a></td>
</tr>
<tr>
<td>LOWE Robin</td>
<td>Statistics Canada</td>
<td>Jean Talon Building, 13th Floor Ottawa, Ontario K1A 0T6 CANADA <a href="mailto:lowerob@statcan.ca">lowerob@statcan.ca</a></td>
</tr>
<tr>
<td>MAGNIEN François</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:francois.magnien@insee.fr">francois.magnien@insee.fr</a></td>
</tr>
<tr>
<td>Name</td>
<td>Organization</td>
<td>Address</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>OKAMOTO Masata</td>
<td>National Statistics Center, Japan</td>
<td>19-1 Wakamatsu-cho Shinjuku-ku Tokyo 162-8668 JAPAN <a href="mailto:mokamoto2@nstac.go.jp">mokamoto2@nstac.go.jp</a></td>
</tr>
<tr>
<td>PAYNE Lyall</td>
<td>Statistics New Zealand</td>
<td>PO Box 2922 Wellington NEW ZEALAND <a href="mailto:Lyall.Payne@stats.govt.nz">Lyall.Payne@stats.govt.nz</a></td>
</tr>
<tr>
<td>RIBE Martin</td>
<td>Statistics Sweden</td>
<td>Box 24300 SE-104 51 Stockholm SWEDEN <a href="mailto:martin.ribe@scb.se">martin.ribe@scb.se</a></td>
</tr>
<tr>
<td>ROMAN Jean-Claude</td>
<td>Eurostat, Luxembourg</td>
<td>Bâtiment Jean Monnet Rue Alcide de Gasperi L-2920 Luxembourg - Kirchberg LUXEMBOURG <a href="mailto:jeanclau.de.roman@cec.eu.int">jeanclau.de.roman@cec.eu.int</a></td>
</tr>
<tr>
<td>SILVER Mick</td>
<td>Cardiff University, UK</td>
<td>Cardiff Business School Colum Drive Cardiff CF10 3EU UNITED KINGDOM <a href="mailto:silver@cardiff.ac.uk">silver@cardiff.ac.uk</a></td>
</tr>
<tr>
<td>STOEVSKA Valentina</td>
<td>International Labour Office, Geneva</td>
<td>4, Route des Morillons CH-1211 GENEVA SWITZERLAND <a href="mailto:stoevska@ilo.org">stoevska@ilo.org</a></td>
</tr>
<tr>
<td>TURVEY Ralph</td>
<td>London School of Economics</td>
<td>30 Sloane Gardens London SW1W 8DJ UNITED KINGDOM <a href="mailto:ralph@turvey.demon.co.uk">ralph@turvey.demon.co.uk</a></td>
</tr>
<tr>
<td>VARTIA Yrjö</td>
<td>University of Helsinki</td>
<td>PO Box, 54, 000140 Helsingin yliopisto Helsinki FINLAND <a href="mailto:yrjo.vartia@helsinki.fi">yrjo.vartia@helsinki.fi</a></td>
</tr>
<tr>
<td>VIGLINO Lionel</td>
<td>Eurostat, Luxembourg</td>
<td>Bâtiment Jean Monnet Rue Alcide de Gasperi L-2920 Luxembourg - Kirchberg LUXEMBOURG <a href="mailto:lionel.viglino@cec.eu.int">lionel.viglino@cec.eu.int</a></td>
</tr>
<tr>
<td>WOOLFORD Keith</td>
<td>Australian Bureau of Statistics</td>
<td>Room 5N107, Locked Bag 10 Belconnen ACT 2616 AUSTRALIA <a href="mailto:keith.woolford@abs.gov.au">keith.woolford@abs.gov.au</a></td>
</tr>
</tbody>
</table>
### Observers

<table>
<thead>
<tr>
<th>Name</th>
<th>Organisation</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRUDIEU Germaine</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:germaine.brudieu@insee.fr">germaine.brudieu@insee.fr</a></td>
</tr>
<tr>
<td>GUIRONNET Nicolas</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:nicolas.guironnet@insee.fr">nicolas.guironnet@insee.fr</a></td>
</tr>
<tr>
<td>MAUREL Françoise</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:francoise.maurel@insee.fr">francoise.maurel@insee.fr</a></td>
</tr>
<tr>
<td>SEMECURBE François</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:francois.semecurbe@insee.fr">francois.semecurbe@insee.fr</a></td>
</tr>
<tr>
<td>VITTEK Gérard</td>
<td>INSEE, France</td>
<td>18, Bd Adolphe Pinard 75675 PARIS CEDEX 14 FRANCE <a href="mailto:gerard.vittek@insee.fr">gerard.vittek@insee.fr</a></td>
</tr>
</tbody>
</table>
La collection « Regards à l’étranger » :

n° 1 La statistique au Royaume-Uni, Véronique ALEXANDRE, 1998
n° 2 La statistique en Allemagne, Cécile LEFEVRE, 1998
n° 3 Salaires et coûts salariaux en France et en Allemagne, 1999
n° 4 Groupe de Paris. Emploi et Rémunération. Labour and Compensation
   Rapport sur la session 1998 [Londres 6-7 juillet], 1999
n° 5 Groupe de Paris. Emploi et Rémunération. Labour and Compensation
n° 6 La statistique en Italie, René HALLU, 2001
n° 7 Registres de population et traitement des données personnelles en Italie
   René HALLU, 2001
n° 8 La statistique en Chine, Véronique ALEXANDRE et Luc BROUARD, 2001
n° 9 Séminaire Insee - Eurostat sur les recensements après 2001
n° 9A Insee - Eurostat Seminar on censuses after 2001
n° 10 International Working Group on Price Indices (Ottawa Group)
International Working Group on Price Indices:
Proceedings of the Seventh Meeting

The International Working Group on Price Indices (Ottawa Group) is a "city group" created in 1994 with the approval of the United Nations. It provides a forum for specialists to share their experiences and discuss research on crucial problems of measuring price change, particularly in the area of consumer price indices. Participants are researchers and practitioners who work for, or are advisors to, statistical agencies in different countries, or international organisations. The seventh meeting of the group was hosted by the Institut National de la Statistique et des Études Économiques (INSEE) of France in Paris from 27 to 29 May 2003. This volume contains the proceedings of the meeting, including presented papers and reports of the eight sessions.

La série Regards à l'étranger présente des dossiers sur les statistiques publiques de pays étrangers, comparées à la statistique française. Ces dossiers portent soit sur des systèmes statistiques nationaux et leur insertion dans l'administration et la société du pays, soit sur des questions thématiques précises.

ISSN 1291-1887
ISBN 2-11-069228-0