LE RECENSEMENT DE LA POPULATION 1990

Innovations méthodologiques
LE RECENSEMENT DE LA POPULATION 1990

Innovations méthodologiques
SOMMAIRE

AVANT-PROPOS ......................................................... 5

1 - PRÉSENTATION
1.1 Le déroulement et l'exploitation du RP .......................... 9
1.2 Quelques aperçus sur l'étranger ................................. 25

2 - LA COLLECTE
2.1 Le questionnaire : Comment traiter la question sur la formation ? .......................... 39
2.2 La préparation et le déroulement de la collecte .......................... 43
2.3 Les erreurs de couverture : l'enquête post-censitaire .................. 61
2.4 Éléments sur les taux de non réponses ............................. 149
2.5 Rapprochement enquête emploi - RP ............................. 155

3 - L’EXPLOITATION
3.1 Le dénombrement du RP 1990 (Bilan chiffré) ..................... 207
3.2 Échantillonnage pour le contrôle de qualité du RP 1990 ............. 215
3.3 La saisie de l’exhaustif .......................................... 237
3.4 Le contrôle de qualité de l’exhaustif léger .......................... 253
3.5 le chiffrement automatique du RP DOM .......................... 269

4 - L’ÉCHANTILLONNAGE
Présentation .............................................................. 299
4.1 La problématique générale .......................................... 301
4.2 Le tirage au 1/20 du RP comme un échantillon de districts .......... 311
4.3 Résultats obtenus dans l’expérimentation du tirage au 1/20 du RP comme un échantillon de districts (note 1222/470 du 28 avril 1986) ........ 317
4.4 Échantillonnage contrôlé de districts pour le tirage du 1/20 du recensement .......................................... 335

Le recensement de la population 1990 3
La réalisation d’un recensement de la population est à chaque fois une nouvelle aventure et, sous l’apparente continuité, se cachent bien des innovations. Ce numéro d’INSEE-Méthodes décrit certaines des améliorations projetées ou réalisées à l’occasion du recensement de 1990.

Si ce n’est dans le premier document où Jean-Paul BARAILLE présente les grandes lignes du recensement, le lecteur aura sans doute une impression de désordre ou de décousu ; cela tient du fait que chaque innovation est le plus souvent indépendante des autres et porte sur un champ limité.

Après une comparaison avec les autres pays européens où Chantal ROUGERIE nous donne une idée de la grande diversité à laquelle Eurostat est confronté, le document est organisé en un plan grossièrement chronologique allant de la collecte à l’exploitation.

Un court texte de Geneviève LEBOUTEUX rappelle les hésitations récurrentes d’un recensement à l’autre sur la façon d’interroger la population sur le thème de la formation : plusieurs tests successifs ont conduit à simplifier sensiblement la question sur le diplôme par rapport à 1982 et à rejeter la formulation en termes de niveau de formation.

Pierre-Alain AUDIRAC présente ensuite un très vivant rapport sur la collecte où là mieux qu’ailleurs peut-être, apparaissent les améliorations "de détail" qui ont fait du RP de 1990 un bon RP aux yeux de tous les partenaires. Riche en suggestions, cet article nourrira sûrement les réflexions sur le futur recensement.

Les articles suivants traitent plutôt de la qualité que des procédures. Nicole COEFFIC décrit l’enquête post-censitaire ; le sous-dénombrement au RP 1990 serait resté au même niveau qu’en 1962, dernier RP pour lequel une mesure avait été établie. Des sous-populations à problèmes sont identifiées (ménages d’une seule personne, agglomération parisienne, étrangers, jeunes, chômeurs, étudiants,...) et le risque de sous ou sur-dénombrement évalué.

Jean-Michel DURR nous rappelle que les bulletins collectés sont parfois incomplets et doivent alors faire l’objet d’imputations.

Dominique ROUAULT a rapproché les questionnaires de l’enquête emploi menée en janvier 1990 et les bulletins individuels correspondants au recensement de mars 1990. La comparaison éclaire principalement le désaccord entre les deux sources sur le...
décompte des chômeurs. L'analyse est menée d'abord à un niveau macro puis au niveau des données individuelles. La réconciliation des résultats apparaît difficile sans prendre en considération les différences dans les protocoles d'enquête.

Le premier objectif du recensement est d'abord de compter les habitants. Chantal MADINIER vient nous rappeler qu'il est bon de compter deux fois plutôt qu'une. Pour la première fois, au recensement de 1990, on a pu confronter par commune le dénombrement traditionnel aux comptages issus de la saisie exhaustive des bulletins et rechercher pour les corriger les principales différences. Le RP 1990 aura donc présenté des résultats bien plus cohérents que ses prédécesseurs.

Raccourcir les délais de saisie des bulletins tout en assurant un contrôle sévère de la qualité était l'objectif pour 1990 puisque avait été retenu le principe d'un dépouillement exhaustif (utilisable localement) avant celui plus sophistiqué par sondage au 1/4. Après une présentation par Jean-Claude DEVILLE de la méthodologie statistique sous-jacente aux contrôles de qualité mis en œuvre pour le contrôle de la saisie exhaustive et du chiffrement du 1/4, Pierre DARMON évoque de nombreux aménagements des procédures qui nous rappellent encore une fois que la réussite d'un RP fait aussi appel au système D à côté de la pure technique statistique. Gérard BADEYAN et Serge FRABOUL reviennent sur la procédure suivie en 1982 pour préciser le niveau de qualité élevé effectivement obtenu.

La codification automatique des libellés est une voie prometteuse d'économies et sans doute de qualité. Pour les régions d'Outre-mer de taille réduite au point qu'un sondage au 1/4 n'apporte pas une précision suffisante, l'exploitation a été menée à l'exhaustif et a constitué la première expérience de codification automatique dans le cadre d'un recensement incluant notamment le chiffrement de la catégorie sociale (CS en 42 positions) et de l'activité économique. Chantal MADINIER évoque les différentes méthodes utilisées (parfois plusieurs pour un même code) et leurs résultats. Elle donne aussi une idée des gains réalisés.

 Traditionnellement, certaines variables du RP sont seulement exploitées pour un sondage au 1/4. Un sous échantillon au 1/20e est prélevé pour disposer de premiers résultats et permettre la réalisation moins coûteuse d'études ne nécessitant pas cet échantillon au 1/4. Le tirage de ces échantillons directement parmi les ménages impose la manipulation de l'ensemble des bulletins ; cette procédure est inévitable pour le sondage au 1/4 mais questionnable dans un souci d'économies pour un sondage au 1/20e : dans ce cas, la sélection d'1/5 des districts de recensement permettrait d'atteindre à coûts bien plus faible un échantillon de la taille souhaitée. Comment alors minimiser les effets de grappe inévitables ? Jean-Claude DEVILLE et Nicole ROTH décrivent la procédure de tirage équilibrée des districts permettant de contrebalancer l'essentiel de la perte de précision.
Partie 1

PRÉSENTATION
1.1- LE DÉROULEMENT ET L’EXPLOITATION DU RP 90 EN FRANCE

Jean-Paul Baraille

En mars-avril 1990 le 32e recensement de la population de la France a été réalisé simultanément sur le territoire métropolitain, dans les départements d’outre-mer et à Saint-Pierre-et-Miquelon (encadré 1). Semblable aux recensements précédents, aussi bien en ce qui concerne les modalités de collecte de l’information et le contenu des questionnaires que du point de vue des méthodes de traitement informatique, le recensement de 1990 se caractérise par un effort important de l’Insee pour mettre à la disposition des utilisateurs des résultats de qualité dans des délais aussi brefs que possible. Pour y parvenir, l’équipe centrale responsable du projet et les directions régionales de l’Insee ont :

- rénové profondément le dispositif de collecte ;
- renforcé la communication entre les multiples acteurs ;
- défini un plan d’exploitation efficace, plus "ramassé" qu’en 1982 ;
- conçu et développé toute une panoplie de "produits de diffusion".

1 - Les objectifs et comment les atteindre

Le recensement vise deux objectifs principaux :

- déterminer la population légale des circonscriptions administratives du pays ;
- connaître les structures sociodémographiques et professionnelles de la population ainsi que les caractéristiques du parc de logements, pour toute zone géographique, quelle que soit sa population.

Les habitants ont donc eu à remplir deux types de questionnaires (feuille de logement et bulletins individuels) qui leur étaient remis par des agents recenseurs. La collecte

(1) Une enquête Famille a été réalisée, comme en 1982, sur un échantillon aréolaire au 1/50. Environ 340 000 femmes âgées de 18 à 64 ans ont eu ainsi à remplir un questionnaire spécifique en plus des bulletins du recensement.
Le recensement général de la population des DOM s’est déroulé pour la deuxième fois en même temps qu’en métropole et pour la première fois en utilisant les mêmes méthodes. Néanmoins, une adaptation a été nécessaire pour mieux appréhender la réalité locale.

Un pré-recensement, réalisé à l’automne 1989, a permis de repérer les habitations et d’établir la cartographie. La date de référence du recensement proprement dit était le 15 mars 1990. Les agents recenseurs sont ici des enquêteurs qui remplissent eux-mêmes les bulletins. Mais l’organisation de la collecte est identique à celle de métropole. La campagne de communication a été adaptée aux DOM : spot TV tourné à la Réunion, affiche d’une petite fille réunionnaise...

Les questionnaires de métropole ont été repris pour l’essentiel. Le bulletin individuel est le même ce qui facilitera les comparaisons. La feuille de logement ne comprend pas de questions sur le chauffage, mais sur l’équipement du ménage, l’alimentation en eau, en électricité, l’existence d’une cuisine...

Les résultats ont été publiés en deux temps : une exploitation rapide légère et une exploitation lourde, toutes deux exhaustives, c’est-à-dire portant sur la totalité des bulletins collectés. Les variables de l’exploitation rapide légère sont celles de l’exhaustif léger de métropole.


de ces questionnaires a été réalisée sous la responsabilité des maires. Puis, d’avril à novembre 1990, ils ont été vérifiés et dénombrés par les directions régionales de l’Insee, en vue d’établir la population légale des communes, cantons, arrondissements, départements et régions.

Pourquoi un tel "plan d'exploitation" ? Tout simplement parce que la saisie très rapide de l'ensemble des informations collectées requiert des moyens dépassant largement les possibilités de l'Insee. Le plan adopté pour le recensement de 1990 donne la priorité à l'obtention de données simples, mais essentielles, pour toutes les zones géographiques, même peu peuplées (exploitation exhaustive). Les données plus complexes, telles que la profession, la nationalité, le secteur d'activité économique, seront fournies dans un deuxième temps au niveau de zones géographiques comprenant, en général, au moins 2 000 habitants (sondage au 1/4).

Un sondage au 1/20, saisi en tête du 1/4, devait en plus permettre d'obtenir des résultats nationaux et régionaux sur des thèmes tels que profession, formation, structure des ménages, non abordés dans l'exhaustif, et ce avant la disponibilité du sondage au 1/4 pour l'ensemble du pays.

2 - Les améliorations apportées à la collecte

Des questionnaires plus lisibles, adaptés à l'évolution des problèmes de société

Il est important d'assurer la comparabilité des statistiques de base recueillies puis élaborées dans le cadre des recensements. Le contenu des principaux questionnaires ne peut donc évoluer qu'à la marge ; il faut aussi tenir compte du fait que ce sont les habitants eux-mêmes qui remplissent les questionnaires, et non des enquêteurs dûment formés. Les questions doivent donc être claires, faciles à comprendre et bien présentées. Certains thèmes, trop complexes ou ne concernant qu'une infime partie de la population, n'ont pas lieu d'apparaître sur des questionnaires de recensement. Les principales modifications apportées entre 1982 et 1990 à la feuille de logement (FL) et au bulletin individuel (BI) sont les suivantes :

Feuille de logement

Trois questions ont été ajoutées suite à l'abandon du bordereau de maison utilisé en 1982 :
- année d'achèvement de la construction (FL p. 4, 2) ;
- appartenance à un organisme HLM (FL p. 4, 5) ;
- existence d'une exploitation agricole (FL p. 4, 11).

En contrepartie, on a supprimé cinq questions de la FL 82 :
- cuisine (existence et taille) ;
- pièces à usage exclusivement professionnel ;
- emprunt éventuel relatif au logement ;
- alimentation en eau ;
- téléphone.

Plusieurs questions ont par ailleurs été modifiées ; on a notamment introduit une catégorie "logement utilisé occasionnellement" de façon à mieux cerner la notion de résidence.

**Bulletin individuel**

La présentation a été améliorée (aiguillage en fonction de l’exercice d’une activité professionnelle, structure plus linéaire du questionnement...). Les questions portant sur l’éducation et les diplômes, d’une part, et sur le chômage et l’activité professionnelle, d’autre part, ont été largement revues et reformulées.

Les premières ont été élaguées ou supprimées en raison du faible intérêt ou de la qualité médiocre des réponses obtenues ; ainsi trois des questions posées en 1982 ont disparu :
- l’enfant va-t-il à l’école ?
- êtes-vous actuellement élève ou étudiant, stagiaire rémunéré, dans une autre situation ?
- à quel âge avez-vous cessé de suivre régulièrement les cours d’un établissement scolaire... ou universitaire ?

Les secondes, en revanche, sont développées pour tenir compte de l’évolution du chômage et de l’apparition de formes particulières d’emploi :
- ancienneté de la recherche d’un emploi (BI 11) ;
- temps partiel (BI 17) ;
- situations intermédiaires (apprenti, stagiaire, SIVP, etc.) (BI 19).

Enfin, les questions 20 et 21 ont été détaillées de façon à mieux caractériser la profession et la position des actifs occupant un emploi salarié.

**3 - Une collecte dynamique et soigneusement préparée**


Un décret du Premier ministre (n° 89-274 du 26 avril 1989) a fixé la date et les conditions d’exécution du recensement général de la population de 1990¹. Il y est


*INSEE Méthodes n° 52 - 53*
notamment précisé que les maires, hormis les communautés, sont responsables du recensement, de la population et des logements. La population des communautés (militaires en caserne, élèves internes, travailleurs en foyer, étudiants en cité universitaire, personnes âgées en maison de retraite...) est, quant à elle, recensée par l’Insee. Ainsi, le recensement, seule opération statistique concernant tous les habitants et nécessitant leur concours, est une entreprise qui, dans sa phase de collecte, fait intervenir en parallèle deux "pyramides" ayant une base commune :

La Direction Générale des Collectivités Locales

L'Institut National de la Statistique et des Études Économiques (Direction Générale et 24 Directions Régionales)

96 Préfets

450 Conseillers techniques

36 551 Maîtres

3 100 Délégués

110 000 Agents recenseurs

La collecte s’est déroulée en deux phases : recensement des communautés au mois de février, puis, en mars-avril, recensement des logements d’habitation et des personnes y résidant. La distinction très nette entre ces deux phases, sur les plans de la responsabilité, de l’organisation, des personnels de collecte et du calendrier, a eu un effet très positif sur le déroulement de l’ensemble de ce recensement.

Une vaste campagne d’information du public a été lancée deux semaines avant le jour J (5 mars 1990), date de référence du recensement et début du passage des agents recenseurs dans les logements. Le cœur de cette campagne était constitué de spots télévisés diffusés aux heures de grande audience. Un affichage abondant sur les panneaux municipaux, dans les bureaux de poste et certaines agences bancaires, ainsi que dans les halls d’immeuble, a pris le relais de la télévision. Un slogan simple "Toi, mo, lui, tout le monde compte" accompagnait l’ensemble des messages relatifs au recensement, quelle que fût leur forme.

Deux conférences de presse, puis de nombreuses interviews du Directeur général de l’Insee et des principaux responsables de l’opération ont été l’occasion pour les journalistes de la presse écrite ou audiovisuelle de décrypter le recensement dans toutes ses dimensions. Environ 10 000 articles, décrivant aussi bien les méthodes de collecte et d’exploitation que les résultats attendus et leur utilité, ont paru dans la presse nationale, régionale ou locale, de février à mai 1990.

Présentation 13
Essai de novembre 1985

Portant sur plus de 100 000 logements, l'opération visait à évaluer l'intérêt et le coût d'une collecte par voie postale (analogique à ce qui se fait aux États-Unis ou en Suède) s'appuyant sur un répertoire de logements extrait du fichier de la taxe d'habitation. Outre le coût, jugé prohibitif, et les problèmes de mises à jour du fichier utilisé pour l'envoi des questionnaires, le taux de réponse (57 % après relance) a été considéré comme insuffisant.

Essai de novembre 1986

Analogue au précédent, il a partiellement utilisé, sur recommandation de la CNIL (Commission Nationale de l'Informatique et des Libertés), un fichier des abonnés au téléphone. L'analyse des résultats a conforté l'idée qu'une collecte traditionnelle, par "dépôt-retrait" des questionnaires devait être mise en œuvre en 1990.

Essais de janvier et juin 1987

Ces essais ont porté sur la forme et le contenu des questionnaires principaux (feuille de logement et bulletin individuel). Des comparaisons ont été faites avec l'enquête emploi du printemps 1987, de manière à préciser la formulation des questions portant sur l'emploi, la formation et le chômage. Le dossier d'immeuble collectif (rempli pour tout immeuble comprenant au moins deux logements d'habitation) a remplacé le traditionnel bordereau de maison qui était également à remplir pour les maisons individuelles.


Les caractéristiques techniques envisagées par l’Insee pour le recensement de 1990 (méthode de collecte, questionnaires, plan d'exploitation, diffusion des résultats) ont fait l'objet d'une consultation de plusieurs formations du CNIS en novembre 1987. De très légers ajustements ont été faits à la suite de cette consultation. Un an plus tard, le 16 novembre 1988, le CNIS a rendu un avis stipulant que le recensement de la population devait être inscrit au programme annuel d'enquêtes statistiques publiques pour 1990.
Répétition générale de mai-juin 1988

Un ultime essai de recensement, de grande ampleur (350 000 personnes), a été réalisé en 1988. L'objectif était notamment de roder les mécanismes de la collecte à une échelle suffisante pour détecter et résoudre les problèmes d'organisation et de logistique, tant internes à l'Insee qu'externes. La mise au point des manuels de référence, du carnet de tournée de l'agent recenseur, des outils de formation, et des procédures de suivi et de contrôle d'exhaustivité, était aussi un des buts de l'opération. Enfin, l'essai a permis de préciser les modalités du recensement des communautés et de disposer, pour l'ensemble de la population concernée, de questionnaires remplis pouvant servir à tester les chaînes d'exploitation (saisie, redressements, codification). Sur tous ces points, la "répétition générale" de 1988 a fourni des enseignements précieux. Cette innovation s’est inspirée, dans son principe, des opérations analogues menées aux États-Unis deux ans avant le recensement général. Le bon déroulement de la collecte, l’efficacité de la saisie exhaustive et la qualité technique de COLIBRI 90 s’expliquent sans aucun doute largement par la réalisation de cette "répétition générale" qui a de plus permis à de nombreux agents de l’Insee de s’immerger progressivement dans le recensement.


Une première délibération, n° 89-10 du 14 février 1989, a précisé la position de la CNIL sur les différentes phases de l’opération. Bien qu’émettant un avis favorable, la CNIL a posé plusieurs conditions, voire restrictions, en ce qui concerne notamment les cessions de données du recensement :
- interdiction faite aux maires de saisir des informations figurant sur les questionnaires (cela avait été autorisé en 1982) ;
- interdiction d’enrichir des fichiers nominatifs par l’exploitation des fichiers detail (et réciproquement...) ;
- interdiction d’agréger les données issues du recensement à un niveau inférieur à celui de la commune (pour les communes dont la population est inférieure à 5 000 habitants) et, pour les communes dont la population est supérieure à 5 000 habitants, à un niveau inférieur à ce chiffre (la possibilité de dérogation à cette règle est néanmoins prévue).
Ces interdictions, et plus particulièrement la dernière, ont provoqué de très nombreuses protestations d’utilisateurs traditionnels des résultats du recensement : collectivités locales, agences d’urbanisme, etc.

Après un examen approfondi, mené conjointement, des besoins réels de quelques collectivités locales, la CNIL et l’Insee ont établi un contrat-type servant de cadre à la cession de données infra-communales relativement détaillées à des niveaux géographiques fins. C’est l’objet de la deuxième délibération (n° 90-23 du 20 février 1990) de la CNIL, relative au recensement de 1990. Les bénéficiaires en sont, en plus des communes, les collectivités territoriales et les établissements publics poursuivant, comme les communes, des missions de création d’équipements et de services publics. Les données cessibles sont de trois types :
- une vingtaine de tableaux standard à l’îlot (l’îlot compte en moyenne 150 habitants);
- un "fichier détail" îlot comportant pour chaque enregistrement ("logement" ou "individu") une dizaine de variables essentielles;
- un "fichier détail" 2000 un peu plus riche que le précédent, mais dont les identifiants géographiques les plus fins repèrent des zones infra-communales d’au moins 2 000 habitants.

Décret n° 89-274 du 26 avril 1989

Ce décret fixe la date et les conditions dans lesquelles sera exécuté le recensement général de la population de 1990, il a été publié au Journal officiel du 4 mai 1989.

Par ailleurs, une information précoce et adaptée a été fournie aux maires de telle sorte qu’ils puissent réunir les moyens nécessaires pour faciliter le travail des agents recenseurs et des délégués, personnels recrutés spécifiquement pour la collecte.

Des documents pédagogiques ont été élaborés à l’intention des enseignants des classes de CM1 et CM2 des écoles et des classes de 6e et 5e des collèges, afin qu’ils informent et sensibilisent leurs élèves aux thèmes population, emploi et logement, ainsi qu’à l’opération de recensement elle-même (technique d’enquête, traitement statistique, nature et usage des résultats obtenus...).

La collecte proprement dite peut être comparée à une gigantesque mécanique qui met en œuvre un certain nombre d’ingrédients :
- près d’une centaine d’imprimés différents (questionnaires, manuels de référence, documents de gestion administrative, affiches, etc.) ;
- plus de 110 000 personnes effectuant les tâches de base (en majorité les agents recenseurs de la population habitant des logements "ordinaires", mais aussi 1 400 agents recenseurs des communautés et 850 assistants-délégués) ;
- les personnels de mairie impliqués dans l'opération (en appui des agents recenseurs, pour sensibiliser les habitants, pour résoudre les problèmes logistiques, etc.) ;
- l'"encadrement Insee" constitué par les responsables régionaux, les conseillers techniques et les délégués ;
- la formation de l'ensemble des personnels de collecte ;
- la communication entre les différents acteurs et avec les responsables nationaux ;
- le contrôle de l'exhaustivité de la collecte ;
- le suivi de la réalisation des différentes tâches.

Sur ces différents points, les principales avancées du recensement de 1990 sont les suivantes :
- une nette amélioration de la qualité des questionnaires et des manuels de référence (Instructions aux maires, Manuel du délégué, Manuel de l'agent recenseur...) tant pour le contenu et la structure que pour la mise en pages et la qualité de présentation ; à cet égard, le rôle essentiel de la répétition générale de 1988 et la qualité des prestations de l’Imprimerie nationale doivent être soulignés ;
- l’introduction d’une notice départementale d’information donnant à chaque ménage les principales caractéristiques du recensement ainsi qu’un ensemble de résultats et d’analyses simples provenant des recensements précédents ;
- la formation des personnels s’est pour la première fois appuyée sur des modules vidéo conçus et réalisés par l’Insee ("rôle du délégué", "collecte", "numérotation") qui ont agrémenté les nécessaires séances de formation pour lesquelles des affiches et des guides de formation ont comme d’habitude été réalisés ;
- la communication entre les personnels de collecte a bénéficié d’une messagerie Minitel qui, du 1er janvier au 31 mai 1990, a permis aux 4 000 personnes dirigeant et supervisant la collecte d’être en contact à toute heure du jour et de la nuit, de signaler des problèmes, de poser des questions ou de fournir des réponses ; cette "première", pour l’Insee, a été extrêmement appréciée ;
- l’exhaustivité de la collecte est un objectif primordial qu’on a cherché à réaliser en comparant le recensement des logements – effectué par les agents recenseurs– aux nombres fournis par une exploitation des fichiers de la taxe d’habitation ; ce contrôle, fait dès les premiers jours de la collecte, a permis de détecter très vite des omissions et d’y remédier ;
- la mesure de la qualité de la collecte, en termes d’exhaustivité et d’exactitude quant à la catégorie de logement et au lieu de recensement, a par ailleurs fait l’objet d’une enquête post-censitaire sur un échantillon d’environ 30 000 logements (la dernière enquête de ce type remontait à 1962) ;
- la volonté de l’Insee de bien structurer l’opération de collecte et de la réaliser tambour battant et, très généralement, la bonne collaboration des instances communales, ont permis d’achever rapidement cette phase essentielle du recensement : le
calendrier de retour des documents dans les directions régionales de l’Insee, beaucoup plus serré qu’en 1982, a été beaucoup mieux respecté par les mairies.

1.4 - Du dénombrement des bulletins à la population légale

Le simple décompte des logements et des personnes est riche d’enseignements car les évolutions sur une période de 8 ans sont importantes, surtout pour les petites zones géographiques. Mais de lourdes vérifications (de l’ordre de 750 000 heures de travail) sont nécessaires, consistant notamment à "recoller" le recensement des ménages et celui des communautés, et à contrôler la cohérence des chiffres établis par les agents recenseurs et les maires, pour chacun des quelques 320 000 districts de collecte et pour chacune des 36 551 communes. Les résultats du comptage des bulletins individuels sont enfin authentifiés par un décret, qui leur donne une valeur officielle. Plus de 200 textes législatifs ou réglementaires se réfèrent à la population légale, en particulier dans les domaines suivants :
- subventions de l’État aux collectivités locales ;
- modes de scrutin aux élections municipales, nombre de conseillers, nombre des emplois communaux ;
- traitements et indemnités des fonctionnaires des collectivités locales ;
- règles d’adjudication des marchés ;
- plans et travaux d’urbanisme ;
- taux de certaines taxes locales ;
- législation des loyers ;
- création de pharmacies, etc.


1.5 - Une exploitation exhaustive, rapide et solide

Réalisée selon des méthodes classiques, l’exploitation exhaustive 90 constitue néanmoins un grand pas en avant par rapport à celle du recensement de 1982 :
- son contenu a pratiquement doublé ; toutes les informations sur l’immeuble et le logement et la plupart des réponses précédées du bulletin individuel ont été saisies, faisant en réalité du traditionnel exhaustif "léger" un exhaustif "mi-lourd" ;
Encadré 3

Des premières estimations à la population légale

Fin juin 1990, la synthèse des résultats du recensement par commune fournissait des estimations de la population au 5 mars 1990 pour la France, les régions et les départements (voir Insee première, numéros 81, 82 et 83).

De septembre à novembre 1990, les résultats provisoires par commune et les évolutions démographiques 1975-1982-1990 sont publiés dans les fascicules départementaux de la série orange qui permettent une première analyse des mouvements intercensitaires entre villes et campagnes, en faisant la part du solde naturel (différence entre naissances et décès) et du solde migratoire. La croissance des grandes villes a repris. La population des banlieues augmente par excédent naturel tandis que les communes rurales périurbaines bénéficient d’un apport migratoire considérable. En revanche, au cours des quinze dernières années, le chiffre de population est resté stable dans les villes-centres et dans les communes du rural profond (voir Insee première, numéros 111 et 119).


La population dite totale d’une commune peut comporter des doubles comptes : par exemple, les militaires et les élèves internes vivant dans un établissement d’une commune et ayant leur résidence personnelle dans une autre commune. Pour calculer la population d’un ensemble de communes (région, département, canton, unité urbaine…), on ne compte les personnes qu’au lieu de résidence personnelle : c’est le concept de population sans doubles comptes.

Le décret lui-même (texte paru au J.O.) ne comporte que les chiffres de population légale de la France, des régions et des départements. Les tableaux 2 et 3 "annexés au décret" sont de fait publiés dans les fascicules départementaux de la série bleue, qui donne la population des arrondissements, cantons et communes. Il faut également noter que le caractère "légal" de la population d’une commune s’étend à ses composantes : population municipale, population comptée à part (dont doubles comptes) et, par voie de conséquence, à la population sans doubles comptes. En prélude à la série bleue (fascicules départementaux et volume "Population de la France"), l’Insee première n° 116 (janvier 1991) donne la population des régions, des départements et des communes de plus de 20 000 habitants.
les délais d'exploitation ont été très fortement raccourcis, les résultats ayant été mis à disposition à la fin du 1er semestre 1991 alors que ceux de l'exhaustif 82 s'étaient échelonnés de juin 1984 à juin 1986.

Sept entreprises ont procédé, de mai à décembre 1990, à la saisie d'environ 1,5 milliards de caractères correspondant à 26,2 millions de logements et 55,36 millions de personnes y résidant habituellement. La population "hors ménages" (soit 1,26 million) a été traitée dans les directions régionales de l’Insee.

L’exploitation exhaustive comporte, outre les caractéristiques du logement, les données suivantes issues des bulletins individuels :
- sexe ;
- état matrimonial légal ;
- date de naissance ;
- lieu de naissance (2 modalités : en métropole/hors métropole) ;
- indicateur de nationalité (Français de naissance, devenu Français, étranger de la CEE étranger d’un pays n’appartenant pas à la CEE) ;
- type d’activité (actif ayant un emploi, élève ou étudiant, chômeur, retraité, femme au foyer...) ;
- ancienneté de recherche d’un emploi ;
- lieu de travail (département et commune) ;
- statut (salaré, à son compte) ;
- situation particulière d’emploi (apprenti, stagiaire...) ;
- position professionnelle.

Dans les directions régionales de l’Insee, après un contrôle de réception sur échantillon, les fichiers de saisie subissent un certain nombre de traitements : vérification des identifiants géographiques, contrôle de cohérence entre données, détermination du code de la commune de lieu de travail (à partir du libellé saisi), calcul de variables dérivées des données brutes (par exemple : nombre de personnes du ménage et indicateur de confort du logement), etc. Les données, stockées sur disque, permettent d’obtenir un grand nombre de tableaux structurels sur la population et les logements. En simplifiant, les principaux résultats de l’exhaustif diffusés par les Services études et diffusion des directions régionales de l’Insee et Insee Info Service à Paris, se présentent comme suit :
- trois séries de tableaux normalisés (détailés, standard, résumés), disponibles pour les zones géographiques les plus variées (communes ou ensemble de communes et, sous certaines conditions, pour des zones infra-communales telles que quartiers ou îlots) ; ces tableaux sont disponibles sur papier, disquette ou bande magnétique ;
- des fascicules départementaux (série jaune, intitulée "Logements-population-emploi : évolutions 1975-1982-1990") qui donnent les résultats sous forme de tableaux normalisés ou indicateurs pour les communes, les cantons, les arrondissements, les unités urbaines et les zones d’emploi du département ;
- des fichiers de données individuelles anonymes, dits "fichiers détail", extraits de la base de données et portant sur des communes entières ou sur des zones infra-com-
munales d'au moins 5 000 habitants ; sous certaines conditions, ces fichiers peuvent être fournis pour des zones de moins de 5 000 habitants.

Sur le plan technique et organisationnel, signalons que la concomitance partielle entre l'établissement de la population légale et la saisie exhaustive légère a conduit à faire bénéficier chaque opération l'une de l'autre : la vérification de la population légale a pris en compte ceux des comptages déjà réalisés de la saisie exhaustive et en échange la préparation des lots de bulletins adressés aux façonniers a profité des travaux menés dans le cadre de la "population légale" (cf. articles "le dénombrement du RP90" et "la saisie de l'exhaustif léger").

1.6 - Les sondages au 1/4 et au 1/20 (saisie colibri)

Sur la base du fichier exhaustif, est tiré informatiquement un logement sur quatre. Les bulletins individuels correspondant à ce logement font l'objet d'une saisie (complémentaire) de toutes les données ne figurant pas dans l'exhaustif. La saisie est "assistée par ordinateur" en ce sens que certains libellés, saisis en clair, sont analysés à l'aide d'algorithmes permettant, par comparaison avec des fichiers de référence, de déterminer les codes relatifs à ces libellés. La profession, le pays de naissance, la nationalité, l'activité économique de l'établissement, figurent parmi les variables traitées de cette manière, dénommée "CODification en Ligne des Bulletins du Recensement des Individus" ou COLIBRI. Le recensement de 1990 est le troisième, après ceux de 1975 et 1982\(^1\), à mettre en œuvre un système de ce type.

Pour accélérer l'obtention de résultats nationaux, régionaux et départementaux, un sondage au 1/20 a été extrait du sondage au 1/4 par tirage d'un district (aire élémentaire de collecte) sur 5. Ces deux sondages fournissent des informations plus fouillées que celles tirées de l'exhaustif, notamment sur l'emploi, les professions et catégories socioprofessionnelles, le type d'activité, les structures familiales et la composition des ménages, les migrations et la nationalité. Les résultats sont présentés sous la même forme, quel que soit le sondage. Ceux du 1/20 étaient disponibles au 1er trimestre 1992, ceux du 1/41 étaient à la fin du 1er semestre 1992 ; ces derniers permettent de réaliser les études concernant des communes ou des zones géographiques de toutes dimensions, à partir de 2 000 habitants. Les résultats du sondage au 1/20 ont permis d'effectuer plus tôt des études au niveau national et régional.

(1) COLIBRI 82 a fait l'objet d'un article détaillé dans le numéro 30 du Courrier des Statistiques (avril 1984).
1.7 - La diffusion des résultats

D’un recensement à l’autre, les besoins changent et les techniques évoluent. Un certain nombre de principes demeurent. En particulier, il faut savoir concilier rapidité et qualité. Les résultats du recensement servent en effet de base à l’application de multiples règlements et à de très nombreuses études, ou encore de calage pour certains outils statistiques (enquêtes, indices, etc.). Ils sont donc très attendus mais leur rôle de référence justifie le soin et le temps mis à les élaborer. Par ailleurs, la demande présente deux caractéristiques quelque peu contradictoires : les utilisateurs des résultats de recensement souhaitent, d’une part disposer de produits "standard" (cohérents d’un recensement à l’autre et identiques, ou facilement comparables, pour toutes les zones géographiques) et, d’autre part, ils ont des besoins très spécifiques qui ne sont pas toujours satisfaits par l’offre ex ante de l’Insee.

Ces constatations nous ont d’abord amenés à reconduire en 1990 les produits "standard" (fascicules bleus, fascicules orange, tableaux normalisés, etc.) et à publier des résultats provisoires (par exemple, les fascicules orange 1990 qui ont été publiés de septembre à novembre 1990 alors que ceux de 1982 s’étaient échelonnés jusqu’en septembre 1983). De plus, un système informatique très souple, CIRCE, permet d’obtenir :
- des tableaux prédéfinis aux échelons géographiques standard,
- des tableaux prédéfinis agrégés sur des zones spécifiques,
- des tableaux prédéfinis pour des sous-populations,
- des tableaux spécifiques (pour répondre à des demandes particulières),
- des extractions de tableaux ou de données individuelles anonymes ("fichiers détail"), etc.

Les supports de sortie sont, au choix de l’utilisateur, le papier, la bande ou la cassette, la disquette.

Les résultats peuvent également être fournis sous forme de cartes, réalisées à partir de fonds de cartes de l’Institut Géographique National : par exemple, France par départements, département par cantons ou communes.
**Point provisoire**

Un peu plus de deux ans après le début des opérations de collecte, le recensement de mars-avril 1990 vient de sortir des phases d’exploitation et entre dans celle d’analyse des données et de poursuite de la diffusion des résultats. Cette période d’utilisation intense des données se poursuivra en 1993 ; le prochain recensement commencera alors à prendre forme...

Soulignons que le recensement engage des moyens importants (un budget de l’ordre de 950 millions de francs et plus de 2 000 000 d’heures de travail à l’Insee, une importante participation des mairies...) et qu’il mobilise de nombreuses personnes de l’Insee sur un projet dont la réussite a des retombées économiques et sociales importantes pour le pays.

---

(1) Cet article a été rédigé en 1992.
1.2 - LES RECENSEMENTS DANS LA COMMUNAUTÉ EUROPÉENNE

Chantal Rougerie

1 - Les recensements européens :
un héritage plus que séculaire

De tout temps, les États ont éprouvé le besoin de dénombrer leur population et de connaître ses principales caractéristiques. Dès le milieu du XIXème siècle, tous les pays de la Communauté européenne ont procédé à un recensement de la population. C’est notamment avec la pratique régulière du recensement national que débute en Europe l’ère statistique, que se créent et se développent les services officiels de statistique (Tableau 1).

Les accidents de l’histoire ont parfois compromis l’organisation des recensements.

La Seconde Guerre mondiale, par exemple, retarda ou annula, dans les pays belligérants, l’exécution des recensements prévus pour 1940/1941. D’autre part, suite à l’annexion ou à la perte de territoire, les recensements de population ne sont pas toujours strictement comparables dans le temps.

La synchronisation et la comparabilité internationale des résultats des recensements a largement motivé l’action d’Eurostat.


---


Présentation
Toutefois, la France et l'Italie avaient obtenu une dérogation. La France, pour des motifs électoraux, a anticipé le mouvement général d’un an (mars 1990), alors que le recensement italien a été réalisé en octobre 91 (graphique 1).

**Graphique 1**

*Le programme communautaire des recensements européens 1990/1991*

<table>
<thead>
<tr>
<th>Pays</th>
<th>Date Commencement</th>
<th>Date Fin</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR</td>
<td>17-03</td>
<td>21-04</td>
</tr>
<tr>
<td>L</td>
<td>01-03</td>
<td>21-04</td>
</tr>
<tr>
<td>E</td>
<td>01-03</td>
<td>15-04</td>
</tr>
<tr>
<td>F</td>
<td>05-03</td>
<td></td>
</tr>
<tr>
<td>DK</td>
<td>01-03</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>01-03</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mars 90</th>
<th>Janvier 91</th>
<th>Mars 91</th>
<th>Avril 91</th>
<th>Mi-91</th>
<th>Octobre 91</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Enfin, l’Allemagne, le Danemark et les Pays-Bas doivent fournir des informations comparables, extraites de registres administratifs et/ou d’enquêtes spécifiques.

**2 - Les recommandations internationales : le rôle d’Eurostat**

Depuis de nombreuses années, des organismes internationaux, tel qu’Eurostat, émettent un certain nombre de recommandations, afin d’aider les pays à améliorer la planification, l’organisation, la réalisation et le rendement de leur opération censitaire. Plus concrètement, il s’agit de favoriser la synchronisation dans le temps et l’harmonisation du contenu des recensements, dans un souci de comparabilité internationale.


Les recommandations d’Eurostat se sont donc articulées autour de deux exigences :
- la synchronisation dans le temps des recensements ;
- l’élaboration de tableaux statistiques homogènes concernant certaines caractéristiques démographiques, professionnelles et sociales des individus, des ménages et des familles.

Au-delà de ses visées internationales, le recensement de la population reste avant tout une source d’informations indispensables à l’échelle de chaque pays.

Contrairement au programme communautaire précédent, le conseil des ministres de la communauté européenne autorise (article 3 de la directive du 26 mai 1987) le recours à des méthodes de remplacement, pour tous les pays qui ne sont pas en mesure d’organiser un recensement classique exhaustif. Le Danemark, l’Allemagne et les Pays-Bas figurent parmi ceux-ci (tableau 1).

**Tableau 1**

*Analyse comparative des procédures de collecte de données utilisées dans le cadre du programme communautaire de recensement*

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>DK</th>
<th>D</th>
<th>GR</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quelle source d’information ?</strong></td>
<td>recensement</td>
<td>registre</td>
<td>enquête</td>
<td>recensement</td>
<td>recensement</td>
<td>recensement</td>
</tr>
<tr>
<td>Le recensement est-il précédé d’une enquête ou d’un recensement pilote ?</td>
<td>non</td>
<td>sans objet</td>
<td>sans objet</td>
<td>oui</td>
<td>non</td>
<td>oui</td>
</tr>
<tr>
<td>Quelle est la procédure de collecte ?</td>
<td>correspondance</td>
<td>sans objet</td>
<td>sans objet</td>
<td>interview</td>
<td>dépôt</td>
<td>dépôt</td>
</tr>
<tr>
<td>Combien d’agents recenseurs ?</td>
<td>14 000 = 1/700 hab</td>
<td>sans objet</td>
<td>sans objet</td>
<td>130 000 = 1/80 hab</td>
<td>32 000 = 1/1200 hab</td>
<td>110 000 = 1/500 hab.</td>
</tr>
<tr>
<td>Une enquête de contrôle post-censitaire est-elle prévue ?</td>
<td>oui, avec registre national</td>
<td>sans objet</td>
<td>sans objet</td>
<td>oui</td>
<td>oui</td>
<td>oui</td>
</tr>
</tbody>
</table>
Analyse comparative des procédures de collecte de données utilisées dans le cadre du programme communautaire de recensement

<table>
<thead>
<tr>
<th>Quelle source d'information ?</th>
<th>IRL</th>
<th>I</th>
<th>L</th>
<th>NL</th>
<th>P</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>recensement</td>
<td>oui</td>
<td>oui</td>
<td>oui</td>
<td>sans objet</td>
<td>oui</td>
<td>oui</td>
</tr>
<tr>
<td>Le recensement est-il précédé d'une enquête ou d'un recensement pilote ?</td>
<td>oui : presse, radio TV</td>
<td>oui : presse, radio TV</td>
<td>oui : presse, radio TV</td>
<td>sans objet</td>
<td>oui : presse, radio TV</td>
<td>oui : presse, radio TV</td>
</tr>
<tr>
<td>Quelle est la procédure de collecte ?</td>
<td>dépôt</td>
<td>dépôt</td>
<td>dépôt</td>
<td>sans objet</td>
<td>dépôt</td>
<td>dépôt</td>
</tr>
<tr>
<td>Combien d'agents recenseurs ?</td>
<td>3 550 = 1/1 000 ab.</td>
<td>90 000 = 1/650 hab.</td>
<td>1 990 = 1/200 hab.</td>
<td>sans objet</td>
<td>12 000 = 1/850 hab.</td>
<td>117 500 = 1/500 hab.</td>
</tr>
</tbody>
</table>

2.1. Les recensements classiques

Bref aperçu de leur contenu

Neuf pays, la Belgique, la Grèce, l'Espagne, la France, l'Irlande, l'Italie, le Luxembourg, le Portugal et le Royaume-Uni\(^1\) ont choisi la voie classique du recensement exhaustif. Le contenu de leur questionnaire concerne les personnes, les ménages, les noyaux familiaux, les logements et les bâtiments.

Dans une optique de comparabilité internationale, les Nations unies dressent une liste de "caractéristiques fondamentales" à laquelle les pays doivent se conformer. Celles-ci serviront également à remplir les tableaux statistiques élaborés par Eurostat en collaboration avec les instituts nationaux de statistiques. Ces tableaux couvrent des domaines aussi divers que la structure de la population, la population active, l'éducation, les ménages, les noyaux familiaux et les logements (tableau 2).

\(^1\) La présentation et le contenu des questionnaires de l'Angleterre, de l'Ecosse, de l'Irlande du Nord et du pays de Galles diffèrent légèrement (tableaux 3 et 4).
Tableau 2

Contenu des tableaux statistiques élaborés par Eurostat

<table>
<thead>
<tr>
<th>Sujet</th>
<th>Intitulé du tableau</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure de la population</strong></td>
<td>- Population totale, par sexe et âge au dernier anniversaire à la date du recensement</td>
</tr>
<tr>
<td></td>
<td>- Population totale, par sexe et année de naissance</td>
</tr>
<tr>
<td></td>
<td>- Population, par âge et état matrimonial</td>
</tr>
<tr>
<td></td>
<td>- Population totale, par sexe et par nationalité</td>
</tr>
<tr>
<td></td>
<td>- Population, par nationalité et par groupe d'âge</td>
</tr>
<tr>
<td><strong>Population active</strong></td>
<td>- Population active (ayant un emploi ou en chômage), par nationalité et par groupe d'âge</td>
</tr>
<tr>
<td></td>
<td>- Femmes actives, par groupe d'âge et état matrimonial</td>
</tr>
<tr>
<td></td>
<td>- Population active, par sexe, groupe d'âge et statut professionnel</td>
</tr>
<tr>
<td></td>
<td>- Personnes ayant un emploi, par nationalité et groupe d'âge (total, hommes et femmes)</td>
</tr>
<tr>
<td></td>
<td>- Salariés, par nationalité et groupe d'âge (total, hommes et femmes)</td>
</tr>
<tr>
<td></td>
<td>- Personnes ayant un emploi, par sexe et par branche d'activité économique</td>
</tr>
<tr>
<td></td>
<td>- Travailleurs étrangers, par nationalité, sexe et activité économique</td>
</tr>
<tr>
<td></td>
<td>- Personnes ayant un emploi, par sexe et par groupe de professions (CITP)</td>
</tr>
<tr>
<td></td>
<td>- Personnes économiquement inactives, par groupe d'âge et par catégorie</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td>- Personnes actives ayant dépassé l'âge minimal de fin de scolarité, par sexe, groupe d'âge et niveau d'instruction atteint</td>
</tr>
<tr>
<td></td>
<td>- Personnes inactives n'étudiant pas à plein temps, ayant dépassé l'âge minimal de fin de scolarité, par sexe, groupe d'âge et niveau d'instruction atteint</td>
</tr>
<tr>
<td><strong>Ménages</strong></td>
<td>- Personnes en ménage privé et en ménage collectif, par type d'institution et par sexe</td>
</tr>
<tr>
<td></td>
<td>- Personnes en ménage privé et en ménage collectif, par type d'institution et par statut</td>
</tr>
<tr>
<td></td>
<td>- Personnes en ménage privé et en ménage collectif, par type d'institution et niveau régional</td>
</tr>
<tr>
<td></td>
<td>- Personnes en ménage privé et en ménage collectif, par type d'institution, par sexe et par activité</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés par type et taille de ménage</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés et personnes en ménages privés, par type de ménage</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés, par nombre total de personnes et nombre de personnes actives dans le ménage</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés, par nombre total de personnes et nombre de personnes âgées de 65 ans et plus dans le ménage</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés, par nombre total de personnes et nombre d'enfants âgés de moins de 15 ans</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés, par nombre total de personnes, de personnes actives et de chômeurs</td>
</tr>
<tr>
<td></td>
<td>- Ménages privés, par nombre total de personnes dans le ménage et nationalité du membre de référence</td>
</tr>
<tr>
<td><strong>Noyaux familiaux</strong></td>
<td>- Noyaux familiaux par type et par nombre d'enfants</td>
</tr>
<tr>
<td></td>
<td>- Noyaux familiaux avec des enfants de moins de 6 ans, par type et par nombre total d'enfants</td>
</tr>
<tr>
<td><strong>Logements</strong></td>
<td>- Locaux d'habitation, ménages privés et occupants, par type de locaux d'habitation</td>
</tr>
<tr>
<td></td>
<td>- Logements classiques occupés, par nombre de personnes dans le logement et nombre de pièces</td>
</tr>
<tr>
<td></td>
<td>- Logements classiques occupés, par statut d'occupation, nombre de logements, de personnes et de pièces</td>
</tr>
<tr>
<td></td>
<td>- Logements classiques, par régime d'occupation et période de construction des logements</td>
</tr>
<tr>
<td></td>
<td>- Logements classiques occupés, selon le régime d'occupation, le système d'adduction d'eau et d'existence de lieux d'aisance et de salles d'eau</td>
</tr>
<tr>
<td></td>
<td>- Logements classiques occupés, selon le régime d'occupation, l'existence de chauffage central et selon le combustible principal utilisé pour le chauffage</td>
</tr>
</tbody>
</table>

D'une manière générale, les questionnaires (des pays) permettent de remplir les tableaux recommandés par Eurostat.
Le nombre et la diversité des questions posées varient d’un pays à l’autre. Pourtant, cela ne compromet pas forcément le contenu des tableaux de résultats qui seront publiés ultérieurement, le conseil des communautés européennes laissant la possibilité de recourir à d’autres sources d’informations (registres, enquêtes...). Les tableaux 3 et 4 qui reçoivent la liste des "caractéristiques fondamentales" et des principales "caractéristiques supplémentaires" dressée par les Nations unies, fournissent une synthèse des principales informations contenues dans les recensements classiques des neuf pays de la communauté européenne.

**Tableau 3**

**Sujets traités dans les recensements de la population et des habitations : programme communautaire de recensement 1990/1991**

*Les caractéristiques fondamentales selon les recommandations des Nations unies*

<table>
<thead>
<tr>
<th>Type de Caractéristiques</th>
<th>Sujets</th>
<th>Pays ayant inclus ce sujet</th>
<th>Pays n’ayant pas inclus ce sujet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Géographie</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Lieu de résidence</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Lieu de résidence à une date de référence antérieure</td>
<td></td>
<td>GR, E, F, IRL, I, L, P, UK</td>
<td>B</td>
</tr>
<tr>
<td>Démographie</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Sexe</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Age</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Situation matrimoniale</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Pays de naissance</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Nationalité</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Économie</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Type d’activité</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Profession</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Branche d’activité économique</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Situation dans la profession</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Lieu de travail</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Instruction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Niveau d’instruction</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Ménage et famille</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Lieu avec la personne de référence du ménage privé</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Modalité de jouissance du logement</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Unités d’habitation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Régime de propriété</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Emplacement du local d’habitation</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Type de local d’habitation</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Régime d’occupation</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Occupation par un ou plusieurs ménages</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Nombre d’occupants</td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Nombre de pièces</td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Cuisine</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Système d’adduction d’eau</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Lieux d’aisances</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Salles d’eau</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Type de chauffage</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Bâtiments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Type de bâtiment</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>- Époque de construction</td>
<td></td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
</tbody>
</table>
**Tableau 4**

**Sujets traités dans les recensements classiques de la population et des habitations : programme communautaire de recensement 1990/1991**

Les principales caractéristiques supplémentaires selon les recommandations des Nations unies

<table>
<thead>
<tr>
<th>Type de Caractéristiques</th>
<th>Sujets</th>
<th>Pays ayant inclus ce sujet</th>
<th>Pays n'ayant pas inclus ce sujet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Géographie</td>
<td>- Lieu de présence au moment du recensement</td>
<td>GR, E, F, IRL, I, L, P, UK</td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
</tr>
<tr>
<td></td>
<td>- Durée de résidence</td>
<td>B, E, I, L</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Année d'immigration dans le pays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Démographie</td>
<td>- Lieu de naissance</td>
<td>B, GR, E, F, IRL, I, P, UK</td>
<td>L, UK</td>
</tr>
<tr>
<td></td>
<td>- Groupe national et/ou ethnique</td>
<td>IRL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Langue</td>
<td>IRL, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Religion</td>
<td>B, GR, E, L, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Nombre total d’enfants nés vivants</td>
<td>B, GR, E, I, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Date du premier mariage et/ou du mariage actuel de la femme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Démographie</td>
<td>- Temps de travail</td>
<td>B, GR, F, I, L, P, UK</td>
<td>E, IRL</td>
</tr>
<tr>
<td></td>
<td>- Principal moyen d’existence</td>
<td>B, L, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Secteur d’emploi</td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Lieu de l’école, de l’université</td>
<td>B, IRL, I, L, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Trajet jusqu’au lieu de travail</td>
<td>B, GR, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td>Économie</td>
<td>- Diplômes obtenus</td>
<td>B, E, IRL, P, UK</td>
<td>GR, F, I, L</td>
</tr>
<tr>
<td></td>
<td>- Fréquentation scolaire</td>
<td>B, GR, E, F, IRL, I, L, P, UK</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Aptitude à lire et à écrire</td>
<td>GR, E, I, P</td>
<td></td>
</tr>
<tr>
<td>Ménage et famille</td>
<td>- Loyer</td>
<td>IRL, L, P</td>
<td>B, GR, E, F, I, UK</td>
</tr>
<tr>
<td></td>
<td>- Nombre de voitures par ménage</td>
<td>B, F, IRL, I, UK</td>
<td>GR, E, I, P</td>
</tr>
<tr>
<td>Unités d’habitation</td>
<td>- Surface utile et/ou habitable</td>
<td>B, E, I, L</td>
<td>GR, F, IRL, P, UK</td>
</tr>
<tr>
<td></td>
<td>- Eau chaude</td>
<td>E, I, L</td>
<td>B, GR, F, IRL, P, UK</td>
</tr>
<tr>
<td></td>
<td>- Type de système d'évacuation des eaux usées</td>
<td>B, GR, IRL, L, P</td>
<td>E, F, I, UK</td>
</tr>
<tr>
<td></td>
<td>- Principale source d'énergie pour le chauffage</td>
<td>B, E, F, IRL, I, L</td>
<td>GR, P, UK</td>
</tr>
<tr>
<td></td>
<td>- Isolation</td>
<td>GR, E, P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Electricité</td>
<td>B, GR, E, I, L</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Téléphone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bâtiments</td>
<td>- Matériaux de construction de parties déterminées du bâtiment</td>
<td>I, P</td>
<td>B, GR, E, F, IRL, L, UK</td>
</tr>
</tbody>
</table>

(i) Exclue en Irlande du Nord.
(2) Incluse en Écosse, en Irlande du Nord et au Pays de Galles.
(3) Incluse en Irlande du Nord.
Si les questionnaires nationaux se conforment aux recommandations d’Eurostat, ils restent avant tout adaptés aux besoins spécifiques de chaque pays.

Ainsi, le questionnaire belge pose cinq nouvelles questions relatives au domaine de l’informatique, tandis que certaines questions démographiques – lieu de résidence antérieure, date du mariage actuel et nombre d’enfants encore en vie – ont été supprimées.

Le Royaume-Uni est le seul pays à poser une question sur la race et le groupe ethnique. L’Irlande et le Portugal se distinguent également en abordant le sujet de la religion. Toujours au chapitre des particularités, l’Irlande, pays d’émigration par excellence, est le seul pays à s’intéresser à ses ressortissants établis à l’étranger depuis plus d’un an. Enfin la Grèce consacre une rubrique entière de son questionnaire à l’artisanat et au travail à domicile.

**Les limites du recensement classique et la concurrence d’autres méthodes de collecte**

Au cours de ces vingt ou trente dernières années, des instruments statistiques qui fournissent des informations a priori identiques à celles du recensement, se sont développés et affirmés. Il s’agit essentiellement des enquêtes par sondage et des registres administratifs.

Comme toutes les méthodes d’observation, le recensement est entaché d’erreurs et d’omissions. La qualité du recensement peut dépendre de la compétence et de la motivation des agents recenseurs, du sérieux avec lequel les personnes interrogées remplissent le questionnaire, de la formulation et du type de questions...

Les questions du recensement portent exclusivement sur des faits. En d’autres termes, le recensement ne permet pas d’apprécier les motivations, les attitudes, les intentions ou les aspirations des personnes interrogées face à un problème précis. En outre, ces questions doivent être acceptées par la population. Il est impératif qu’elles ne soient pas perçues comme une gêne, comme une intrusion dans la vie privée.

Le recensement de la population fournit avant tout une image statique de la société. L’introduction de quelques questions rétrospectives (par exemple, les naissances et les décès survenus au cours des douze ou vingt-quatre derniers mois, le lieu de résidence il y a cinq ou dix ans) n’autorise qu’une reconstitution très partielle de la dynamique d’une population.

Reste la possibilité de comparer les résultats de plusieurs recensements. Or un intervalle intercensitaire de dix ans occulte les modifications rapides de la société. Un intervalle plus
court - de cinq ans, comme c’est le cas en Irlande ou en Espagne avec le Padron\(^1\) - serait une solution mais se heurterait à un obstacle de taille, le coût de l’opération.

Le recensement de la population est une opération coûteuse. En 1981/82, les restrictions budgétaires ont compromis, avec une intensité variable, les programmes élaborés par les pays de la communauté européenne. Certains renoncèrent à utiliser la radio et la télévision comme agents publicitaires. D’autres furent incapables d’évaluer la couverture et la qualité de l’opération. On relèvera également une sous-exploitation des informations brutes fournies par les recensements ainsi que la lenteur avec laquelle les résultats furent disponibles et publiés.

La tendance actuelle est, quoi qu’il en soit, à la compression des coûts du recensement et, par extension, à la réduction du volume du questionnaire.

### 2.2. Le cas du Danemark :

*un recensement basé sur les registres administratifs*

Le Danemark est le seul pays de la communauté européenne à collecter les informations requises dans le cadre du programme communautaire de recensement, par le biais de registres administratifs. Le dernier recensement classique au Danemark remonte à 1970.


Parmi les avantages de cette méthode, relevons :
- une périodicité *a priori* plus courte, sachant que les informations sont directement prélevées des registres administratifs ;
- les recensements basés sur les registres autorisant les études rétrospectives ;
- la plupart des données étant collectées dans le cadre du fonctionnement normal de l’administration : cette méthode n’impose en théorie aucune gêne au public ;
- le coût, nettement inférieur à celui d’un recensement classique, se limite à l’accès aux données, à leur saisie, à l’élaboration de tableaux statistiques et à la publication des résultats.

---

\(^{1}\) Le Padron espagnol, réalisé tous les cinq ans, est un dénombrement de la population et de ses principales caractéristiques démographiques.
Au chapitre des inconvénients, il faut bien admettre que le contenu du recensement ne peut dépasser le cadre des informations détenues par les registres. Ceux-ci ne peuvent apporter des réponses à des questions non traitées par les administrations.

De plus, s’il est relativement aisé de modifier les questions et les définitions d’un recensement classique, le contenu des registres est étroitement lié aux systèmes administratifs qu’ils servent et ne peut être que plus difficilement adapté à de nouvelles exigences statistiques.


2.3. Les cas des Pays-Bas :
une combinaison de sources de données

Peu de temps avant le recensement de 1971, un débat public fut organisé à la télévision néerlandaise à propos du recensement et du respect de la vie privée.

L’opinion publique se divisa, et quelques 300 000 personnes, soit plus de 2 % de la population, refusèrent de remplir le questionnaire du recensement.

En 1979, un recensement expérimentional et facultatif se solda par un échec cuisant. Le taux de non-participation fut de 25 % avec des pointes de 40 à 50 % dans les grandes villes. Devant la désapprobation du public et en l’absence d’un cadre légal suffisant sur la protection des données, le gouvernement néerlandais décida le report du recensement de la population et des logements prévu en 1981.

Afin de répondre aux exigences internationales, le recensement classique fut remplacé par trois autres sources statistiques : les registres de population, les enquêtes sur les forces de travail et les enquêtes sur le logement.
Dans la seconde moitié des années 80, un programme statistique de substitution au recensement classique (incluant la révocation de la loi de 1970 sur les recensements) fut adopté aux Pays-Bas. Ce “programme statistique de substitution” couvre une période d’environ dix ans. Il consiste principalement en un relevé régulier d’informations provenant des registres de population locaux (l’assise du système), de données fournies par les enquêtes sur les forces de travail (menées continuellement) et par les enquêtes sur le logement (menées tous les quatre ans).

Selon les responsables néerlandais, cette méthode remplace avantageusement le recensement décennal. Elle procurerait des informations plus nombreuses et de meilleure qualité que celles d’un recensement. La fréquence des opérations de collecte et la rapidité d’exploitation des résultats seraient également supérieures. Enfin, le coût de cette méthode serait largement inférieur à celui d’un recensement classique.

2.4. Le cas de l’Allemagne : le micro recensement


Dans ces conditions, l’Allemagne a obtenu l’autorisation pour 1991, de remplacer le recensement classique par une enquête par sondage. Cette enquête, également appelée micro-recensement, couvre 1 % de la population totale. Ces enquêtes annuelles existent depuis 1957 et sont parfois complétées par des enquêtes complémentaires de taille plus réduite (taux de sondage de 0,1, 0,25 ou 0,50 %).

Par rapport au recensement classique, les enquêtes fournissent une plus grande variété d’informations. Leur souplesse, leur plus grande fréquence et leur coût plus faible peuvent également être des arguments déterminants. À l’inverse, les enquêtes ne peuvent fournir d’informations au niveau local, voire régional.
L'organisation du micro-recensement en 1991 s'est heurtée à un obstacle imprévu : la réunification de l'Allemagne. L'exécution d'un micro recensement sur le territoire de l'ancienne République démocratique allemande a posé les problèmes suivants :
- l'ancienne République démocratique allemande ne disposait d'aucun recensement démographique récent comme base de sondage ;
- il était impératif de tenir compte des différences qui caractérisent de nombreux secteurs de la vie courante ; les définitions des concepts, les questions posées devaient être adaptées en conséquence ;
- il était nécessaire de créer des offices de statistiques dans chaque Land et de former le personnel requis ;
- la population a dû être sensibilisée à la pratique du micro-recensement.

Ces problèmes sont aujourd'hui résolus, et le micro-recensement a eu lieu au même moment dans les seize Länder, avec un programme d’enquêtes unique. Dans l’ancienne République démocratique allemande, le registre de population a servi de base de sondage. Les expériences réalisées lors d’un "sondage test" en 1990 ont conduit à apporter quelques modifications minimes au questionnaire.

3 - Le recensement de la population comme instrument de décision

Le recensement de la population est l’assise indispensable à toute prise de décision dans de nombreux domaines. C’est également une base solide et nécessaire pour tracer les perspectives d’avenir.

Les données du recensement servent à planifier le nombre et l’emplacement des équipements sociaux (crèches, écoles, hôpitaux, maisons de retraite...), à estimer les besoins en matière de transport en commun, de réseau routier ou encore à déterminer les montants qui seront consacrés aux pensions, au logement, aux programmes d’emploi pour les jeunes...

Dans de nombreux pays, sur la base des données fournies par le recensement, on établira la répartition des sièges au sein des différents organes de décision. Dans le même registre, la répartition des budgets alloués chaque année aux communes par le pouvoir central sera également tributaire des résultats du recensement.

Enfin, avec l’enquête et l’état civil, le recensement de la population constitue une des trois sources classiques d’informations pour la recherche en démographie. Le recensement est également une excellente base de sondage, lors de la préparation d’enquêtes spécifiques. C’est également un élément de référence lorsqu’il s’agit de corriger les estimations annuelles de la population et de mettre à jour les registres de population.
Partie 2

LA COLLECTE
À l'issue du recensement de 1982, la division "Recensements de la population" a souhaité réfléchir à la formulation de la question du bulletin individuel concernant les diplômes possédés. La formulation du RP82 était jugée insatisfaisante pour trois raisons principales :

- le taux élevé de non réponse (40 % des personnes de 15 ans ou plus), lui-même lié à l’absence de modalité "sans diplôme" ;

- la place importante prise sur le BI (1/3 de la première page) ;

- le peu de fiabilité des réponses obtenues (voir les études techniques ci-après).

Pendant deux années, 1987 et 1988, plusieurs recherches ont été menées sur cette question concernant le diplôme. Ces travaux ont permis d’aboutir à la solution retenue pour le bulletin individuel du RP90, reprenant finalement une interrogation sur le diplôme le plus élevé détenu, mais nettement allégée par rapport à l’ancienne formulation.


Le problème de l’inconstance des réponses des individus concernant leurs diplômes est sans doute le problème le plus crucial posé par ce questionnement. Cette inconstance a malheureusement été relevée dans tous les travaux de comparaisons temporelles comme on le verra ci-après.

La collecte
L’essai de recensement de 1985 a testé une question ouverte sur les diplômes divisée en 3 rubriques :

- enseignement général primaire ou secondaire ;
- enseignement technique et formation professionnelle de niveau secondaire ;
- enseignement supérieur.


L’essai de recensement de janvier 1987 a repris la formulation de l’essai 1985 (question ouverte) faute de temps pour mettre au point un nouveau projet de question. J.-P. Baraille a comparé les réponses obtenues à ce test aux réponses données par les mêmes individus au cours de l’enquête emploi, deux mois plus tard (mars 87). La comparaison a concerné 12 000 individus, répartis sur les 10 DR de l’essai. L’étude a révélé l’inconstance des réponses des enquêtés : 52 % de réponses identiques seulement ; cette inconstance est pratiquement la même dans chacune des 3 rubriques de diplômes (général, technique, supérieur). Seules 56 % des non réponses à la question sur les diplômes au test de RP s’expliquent par l’absence de diplôme.

En juin 1987, le B1 du nouvel essai de recensement a comporté une question, non plus sur le diplôme mais sur le niveau d’études atteint, en 7 modalités, regroupées sous 4 rubriques : primaire, secondaire technique, secondaire général et supérieur. Cette nouvelle formulation répondait aux considérations suivantes : ce critère de niveau d’études atteint, plus flou que celui de diplôme possédé, correspond sans doute davantage à la façon dont les individus se perçoivent : de plus, les personnes sans diplôme ne se sentent plus exclues du questionnement et peuvent se retrouver dans une rubrique, au moins celle de niveau primaire. Un questionnaire complémentaire accompagnait ceux du test de recensement, une question sur le diplôme le plus élevé obtenu y figurait de manière à pouvoir comparer les réponses en termes de niveau d’études à celles en termes de diplôme. P.-A. Audirac et C. Galant ont comparé les réponses au test de recensement et au questionnaire complémentaire de 3 600 individus, sur 6 DR. Le bilan de la question sur le niveau d’études en 7 modalités est très satisfaisant : moins de 7 % de non réponses, la cohérence entre les niveaux et les diplômes déclarés est réalisée à 96 %.
Suite à ces résultats très encourageants, la division Recensements a souhaité opter pour une question sur le niveau d'études plutôt que sur le diplôme. Les partenaires concernés, Ministère de l'Éducation et CEREQ, ont fortement critiqué cette option avec les arguments suivants :

- le niveau de l'étude est une donnée plus floue que celle de diplôme possédé ;
- cette solution induirait une rupture des séries temporelles ;
- les comparaisons avec les données d'autres pays seraient rendues délicates alors qu'en 1992 les diplômes étrangers devront avoir des équivalents français ;
- les analyses des liaisons formation-emploi seraient plus difficiles car la corrélation avec le diplôme est plus importante que celle avec le niveau d'étude ;

De plus, les 7 modalités de la question en termes de niveau d'études n'ont pas été jugées pertinentes ni adaptées aux nomenclatures habituellement utilisées.

Le dernier essai de recensement, la répétition générale de juin 1988, a pris en compte ces remarques et deux types de BI ont été testés : sur le premier type figurait une question sur le diplôme le plus élevé (type D), sur le second (type N) une question sur le niveau d'études atteint, en 8 modalités, modalités qu'il n'était plus possible de regrouper en rubriques comme lors de l'essai précédent. Une enquête complémentaire sur la formation était jumelée au ramassage des questionnaires dans certaines zones du test, cette enquête était également typée D ou N et comprenait, entre autres, la question "complémentaire" sur le diplôme si le BI traitait le niveau de formation et réciproquement. La mise au point de cette recherche a été réalisée par J.-P. Baraille, G. Lebouteux et B. Macrakis.

B. Macrakis a assuré l'exploitation du test et la comparaison des réponses des individus au test et à l'enquête complémentaire. Les résultats exploitables ont concerné 1 900 individus, répartis sur 10 DR. Les résultats ont été décevants : 13 % de non réponses à la question sur le niveau de formation, 40 % à celle sur le diplôme, enquête complémentaire assez mal acceptée, incohérence des réponses entre test et enquête complémentaire

L'ensemble des conclusions de la répétition générale de 1988 a amené la division Recensements à réexaminer les conclusions de l'essai précédent (juin 1987). Les agents recenseurs de ce précédent test avaient été recrutés parmi les personnels des DR et de la DG. Ceux-ci, très impliqués, avaient obtenu une bonne qualité de remplissage des questionnaires et une bonne acceptation de l'enquête complémentaire. Pour la répétition générale de mai-juin 1988, les agents recenseurs étaient recrutés par les mairies, ce qui rendait les conditions de cet essai beaucoup plus proches de celles d'un véritable recensement.

Par ailleurs, les 8 modalités de la question sur le niveau de formation de l'essai de 1988 présentaient deux inconvenients par rapport à la formulation de l'essai de l'année précédente : l'absence des 4 rubriques qui avaient sans doute permis aux recensés de se repérer plus facilement et la répétition du terme "CAP" dans deux items différents, items sur lesquels les réponses des recensés ont été les plus mauvaises.

Pour toutes ces raisons, il a paru probable qu'une question sur le niveau de formation posée au RP90, particulièrement si ses modalités respectaient les nomenclatures usuelles (modèle 1988) générerait un taux de non réponses de l'ordre de 10 %. Ce taux ne serait pas considéré comme meilleur qu'un taux supérieur à 35 % pour une question sur le niveau de formation. Le problème le plus crucial de la qualité des réponses des recensés n'est donc pas réglé par une question sur leur niveau de formation. Aussi la division Recensements a-t-elle finalement proposé de conserver une question sur le diplôme le plus élevé détenu, puisque cette formulation était davantage demandée par ses interlocuteurs. La version proposée a été celle de l'essai 1988, plus légère que celle du RP82.

Fallait-il adjoindre une modalité "sans diplôme" à cette question ? Ce point a été discuté, le département de la Démographie a finalement opté pour laisser la version 1988 en l'état du fait principalement que l'adjonction d'une modalité "sans diplôme" n'avait jusqu'à présent fait l'objet d'aucun test, qu'il était donc risqué de lancer cette modification directement en vraie grandeur sur les questionnaires du RP90. Ces options ont finalement été approuvées par le CNIS. Pour les questionnaires des DOM, il a par contre été décidé d'ajouter cette modalité "sans diplôme" à la question sur les diplômes. Le recensement des DOM étant réalisé par des enquêteurs (et non des agents recenseurs se contentant de déposer et de reprendre les questionnaires), la problématique n'était pas tout à fait comparable.

Il restera vraisemblablement à se reposer la question d'une modalité "sans diplôme" pour le prochain recensement, point qui pourra alors être testé préalablement.
2.2 - LA PRÉPARATION ET LE DÉROULEMENT DE LA COLLECTE (MÉTROPOLE)

Pierre-Alain Audirac

La présentation du dispositif de la collecte 90, avant le mois de mars 1990, marquait un effort de communication vers le public, les mairies et le personnel de l’Insee. Dès 1986, il était apparu que la préparation de la collecte serait encadrée par deux contraintes : d’une part, la progression limitée des moyens financiers ; d’autre part, le maintien des grands axes de la méthode de collecte : collecte par la méthode de dépôt-retrait et implication forte des mairies. En définitive, la voie qui s’ouvrait était la recherche d’une meilleure utilisation des moyens disponibles. La plupart des initiatives en matière de communication eurent pour but de favoriser le déroulement de la collecte en informant les différents acteurs des objectifs du recensement et en incitant les uns et les autres à contribuer à la réussite de l’opération. L’exemple le plus simple nous est donné par la campagne de publicité. En l’absence de toute campagne, les agents recenseurs (AR) se seraient heurtés à l’incompréhension générale et chacun d’eux aurait consacré des dizaines d’heures à convaincre les habitants de se faire recenser, sans toujours y parvenir d’ailleurs...

Parmi les critères d’appréciation du succès de la collecte, celui de la rapidité conduit à penser que l’objectif de productivité a été atteint. Les agents recenseurs ont achevé leur travail bien plus rapidement qu’en 1982 et le calendrier de retour des documents en direction régionale (DR), pourtant bien plus serré qu’en 1982, a été beaucoup mieux respecté par les mairies. Cette volonté de réduire les délais était, au demeurant, vigoureusement exprimée par un conseiller technique "grande ville" dans le module vidéo "le rôle du délégué".

1 - La préparation de la collecte

1.1. Une responsabilité clairement définie

Confier à une seule personne la responsabilité de l'ensemble de la collecte fut une mesure efficace et a accru la cohérence de l'opération.

En ce qui concerne les documents, la prise en charge complète des relations avec l'Imprimerie Nationale, du calcul des dotations et de l'organisation des acheminements s'est soldée par un gain très important par rapport à 1982.

Il serait souhaitable cependant que, pour les recensements suivants, le responsable de la collecte soit assisté, au minimum, d'un cadre et d'un contrôleur confirmés, ce qui n'a pas été le cas pour ce RP. La responsabilité très directe d'un budget "impressions" de 30 millions de F et la co-responsabilité d'un budget "collecte" de plus de 500 millions de F sont très lourdes. D'octobre 1987 à janvier 1990, la préparation a été vécue comme une course contre la montre. Nous n'avons connu, sur cette période, ni maladie ni accident, mais il est irréaliste de soumettre une opération de cette importance à de tels aléas. À la différence des opérations situées en aval, aucun retard n'est permis sur la collecte. Trois mois de retard sur la mise au point des programmes de traitement des données (Colibri) passeraient inaperçus à l'extérieur de l'Insee. Trois mois de retard sur la préparation de la collecte se seraient traduits par le report pur et simple.

1.2. Les tests

À l'initiative de A. Lery, chef du projet RP90, quatre essais de recensement ont été réalisés en octobre 1985, novembre-décembre 1986, janvier et juin 1987, ainsi qu'une répétition générale en mai-juin 1988. Une expérimentation aussi poussée présente de très grands avantages :

- elle autorise les innovations ;

- elle permet d'affiner de façon très précise les questionnaires principaux et les instructions ;

(1) Les différentes catégories de personnels intervenant dans la collecte sont listées dans l'article.

(2) Chaîne informatisée de codification assistée par ordinateur.
elle impose la réalité et révèle certaines difficultés. L'expérimentation en 1988 d'un carnet de tournée avec préimpression des situations possibles s'est soldée par un échec. La répétition générale a fait apparaître aussi l'hostilité du corps médical à la distinction des disciplines hospitalières ;

elle permet d'avancer sur tous les terrains. Nos relations avec le studio de conception graphique de l'Imprimerie Nationale ont été amorcées en 1986 ;

elle donne au responsable Collecte l'occasion de traiter le sujet dans son ensemble ; l'essai de 1986 a été déterminant à cet égard, d'autant plus que le responsable de la collecte y a joué le rôle d'agent recenseur ;

elle permet d'initialiser le travail dans les établissements régionaux.

Retenons cependant qu'il faut parfois beaucoup de temps pour dégager tous les enseignements d'un essai. Cette partie de l'expérimentation devra être approfondie lors de la prochaine campagne.

1.3. Accroître la lisibilité des instructions

Le défaut majeur du mode d'exposé des instructions pour le RP82 était son caractère administratif. Il a été décidé de les repenser en se plaçant systématiquement du point de vue de l'utilisateur et dans la perspective des actions qu'il doit réaliser.

Ce point est capital et illustre tout particulièrement le souci d'éviter les gaspillages. En 1982, l'examen en quatrième lecture des instructions aux maires ne nous avait pas permis d'en maîtriser toutes les subtilités et l'enchaînement des opérations nous demeurait obscur. Dans le meilleur des cas, 3 000 délégués et conseillers techniques avaient ainsi perdu plusieurs heures à décrypter ce document, tout comme les responsables en mairie. Dans le pire des cas, ces instructions étaient restées lettre morte faute d'une lecture efficace.

Des progrès ont été faits mais, en dépit des efforts, ces critiques demeurent quant à l'instruction aux maires pour le RP90.

Assurer un exposé clair et compréhensible des règles du recensement, que ce soit en direction des DR, des mairies, des délégués, ou des agents recenseurs et supprimer du même coup les pertes en ligne doit demeurer un impératif.

Il n'a pas été possible d'avoir d'emblée une vision globale des réformes à mettre en œuvre et de pouvoir présenter un an avant le RP un document d'ensemble. Nous avons procédé en effet par une reconstruction progressive : le projet de manuel de l'agent...
recenseur élaboré en 1987 faisait apparaître plusieurs outils de conception nouvelle : notice d’information, carnet de tournée, relevé des logements, notification à la mairie. Le recensement des communautés, en 1988, se plaçait dans la logique du décalage par rapport au recensement des logements, avec la création du bordereau annexe de district. Une fois ce dispositif arrêté, furent étudiés les outils à destination des maires et des délégués : tableau de réception n° 42, fiche communale n° 61, fiche individuelle d’agent recenseur n° 69, tableau de bord de collecte n° 70, etc.

Le manque de temps ne nous a pas permis d’élaborer un manuel du conseiller technique (CT) qui aurait dû d’ailleurs, pour être pleinement utile, être disponible au premier semestre 1989. Manifestement, ce document a fait défaut comme lors des recensements précédents, même si la constitution des équipes en DR s’est beaucoup mieux passée qu’en 1982. Les CT pressentis ont disposé respectivement en mars et en mai 1989 de la version définitive des deux manuels d’agent recenseur. De même, la plupart des DR ont dupliqué les projets d’instructions aux maires et de manuel du délégué. Chaque conseiller technique à constitué ainsi peu à peu son dossier recensement. Ce système n’a pas que des inconvénients : l’information "à petites bouchées" est plus digeste ; le capitainat du responsable RP s’en est trouvé affirmé, par la nécessité de définir l’organisation du travail des CT.

2 - Mieux utiliser les moyens disponibles

On a recherché, dans le dispositif de 1982, les zones de "gaspillage" et les éléments pouvant entraver l’action des personnels de collecte. La liste présentée retient l’essentiel des changements effectués dans cette optique et aborde quelques nouveautés.

2.1. La simplification du recensement des constructions

Lors des recensements précédents, tout immeuble (habité ou non et maison individuelle comme immeuble collectif) était recensé à l’aide d’un bordereau de maison, et l’on faisait valoir deux arguments : d’une part, la nécessité de rechercher des occupants même dans les lieux les moins propices à l’habitation ; d’autre part, le bel ordonnancement de l’ensemble bordereau de maison - feuille de logement - bulletin individuel, simple à exposer.

En 1990 on ne garde ce bordereau que pour l’habitat collectif\(^1\).

---

1 D’où la dénomination DIC : dossier d’immeuble collectif.
Il nous a semblé qu'un effort de pédagogie permettrait de sauvegarder l'exhaustivité et de faire comprendre la différence de traitement entre habitat individuel et habitat collectif. L'essai de novembre-décembre 1986 allait montrer que ce point était facilement assimilé par les agents recenseurs qui n'étaient pas troublés par la coexistence de documents avec ou sans dossier d'immeuble collectif.

La portée de cette réforme est considérable :

- représentons-nous la tournée de distribution des imprimés dans un lotissement pavillonnaire. Dans le nouveau système, l'agent recenseur passe de maison en maison en consacrant quelques instants à chaque ménage. Le seul document à renseigner au moment de la remise est le carnet de tournée. Dans l'ancien système, l'agent recenseur devait manipuler et remplir un imprimé supplémentaire. Le remplissage de 14 millions de bordereaux de maison a ainsi été épargné. Il est difficile d'imaginer que l'agent recenseur ait pu, numérotage compris, consacrer en moyenne moins de 4 minutes à chacun de ces documents. C'est donc un million d'heures de travail qui auraient été économisées, et vraisemblablement bien davantage. Or cela représente 10 heures par agent recenseur, et naturellement, bien plus dans les petites villes. En plus du gain de temps, la réduction de "paperasserie" est appréciée par les agents recenseurs ;

- le stockage des imprimés au retour est considérablement allégé : de 25 à 30 % pour les DR de Lille et Poitiers par exemple, capitales de régions où la part de l'habitat individuel est de l'ordre de 80 %. Le gain en stockage est complété par un gain en saisie. En regard, la perte d'une information largement redondante avec celle figurant sur la feuille de logement est sans grande conséquence ;

- on a économisé l'impression de 19 millions de ces imprimés de 4 pages.

2.2. La spécialisation des agents recenseurs

Au recensement de 1982, 10 % environ des districts comprenaient une ou plusieurs communautés. Les communautés sont caractérisées par la mise en commun de certains services, notamment la cuisine et l'hébergement. Elles se partagent dans le RP90 en 3 types d'"établissements" (militaires, d'enseignement avec internat, pénitentiaires) et 6 catégories de "collectivités" (foyers de travailleurs, cités universitaires, maisons de retraite, hôpitaux (séjours de longue durée), communautés religieuses, centres d'hébergement ou d'accueil, autres). Voir les manuels de collecte et l'instruction aux maires.
Cette situation induisait un gaspillage au niveau de la formation : la plupart des agents recenseurs se voyaient imposer une formation inutile, au demeurant fort complexe. Supprimer cette contrainte allait permettre de simplifier la lecture du manuel de l'agent recenseur, la formation bénéficiant du même allégement.

Cette décision allait avoir d'autres conséquences tout aussi importantes :

- employer d'autres agents, spécialisés dans le recensement des communautés, permettait d'espacer dans le temps les deux opérations "communautés" et "logements". On pouvait désormais escompter, en anticipant davantage le recensement des communautés, anticiper aussi les transferts de bulletins de mairie à mairie(1), et donc le bouclage du recensement dans les petites communes où certains imprimés se faisaient auparavant attendre alors que la collecte était achevée ;

- décalage et spécialisation allaient permettre aux directions régionales de l'INSEE de prendre la responsabilité du recensement des communautés et donc de l'encadrement des agents recenseurs spécialisés. Cela pouvait apparaître difficilement acceptable pour certaines mairies : celles qui avaient coutume de confier les districts concernés à des agents communaux, lesquels y trouvaient une récompense substantielle ; celles aussi qui, par une interprétation très large de la notion de résidence, parvenaient à "gonfler" quelque peu leur chiffre de population.

En règle générale, cette mesure a été assez bien acceptée. Les services municipaux des grandes villes, en particulier, ont salué le transfert sur l'INSEE de la responsabilité d'une opération très "technique". Certaines DR, en revanche, ont considéré que cette charge était assez lourde :

- cette mesure a permis une certaine "vérité des prix". Les agents spécialisés ont été recrutés en général pour recenser des communautés aux effectifs importants. La prise de contact avec un seul responsable d'établissement peut, bien souvent, débloquer le recensement de plusieurs centaines, voire plusieurs milliers de personnes pour lesquelles la rémunération au bulletin doit, pour être juste, être notablement inférieure à celle du bulletin habituel. Certaines DR ont critiqué cette disposition, d'autres l'ont applaudie. Elle représente une économie d'environ 3 millions de francs ;

- les délégués ont été encouragés à procéder eux-mêmes au recensement des communautés lorsque la population concernée dans une commune de leur zone n'excédait pas 300 personnes ;

(1) Les bulletins individuels des élèves internes et militaires sont transférés à leur commune de résidence pour être intégrés dans la FL de leur logement habituel.
- l'opération s'est toutefois trouvée compliquée par la séparation des deux phases de la collecte. Il a fallu créer un bordereau annexe de district, permettant notamment de rémunérer les agents spécialisés. Les informations portées sur ce bordereau ont dû être reportées par les délégués sur les bordereaux normaux. Certaines DR s'étaient inquiétées de la charge imposée aux délégués à cette occasion. Elle nous semblait somme toute raisonnable, les recopies étant fort limitées, et le déroulement de la collecte semble l'avoir confirmé.

2.3. Une utilisation plus rationnelle du temps des délégués

L'un des intérêts principaux de la réforme précédente a été de faire intervenir très tôt les délégués par des actions liées au terrain. Force est de constater que, lors du recensement précédent, la ressource "délégués" avait été employée de façon inégale. Nous avons le souvenir de telle grande ville de province où, à l'occasion d'un essai de recensement, les responsables INSEE se voyaient apostrophés par les responsables en mairie par des propos tels que "en 1982, vos délégués ont passé leur temps à la cafétéria".

Certes, il serait illusoire de penser que tout avait été mauvais en 1982, et que tout a été parfait en 1990. Un délégué médiocre n'a pas fait de miracles. Toutefois, les DR s'accordent à considérer que les délégués ont travaillé beaucoup plus qu'en 1982 et que leur rémunération reste très insuffisante au regard des services rendus.

Dès la mi-janvier, outre le contrôle du découpage communal, les délégués ont pris en main la préparation et l'encadrement du recensement des communautés, par la mise à jour de l'inventaire, la prise de contact avec les responsables de communautés et la formation des AR spécialisés. Beaucoup d'entre eux ont joué le rôle d'agent recenseur spécialisé. Enfin, le contrôle sur toutes les phases d'une opération préliminaire assez complexe les a préparés à la phase essentielle de recensement des ménages.

Au moment d'aborder les étapes suivantes, cette montée en charge rapide de leur activité les avait déjà positionnés, face aux services municipaux, comme techniciens de la collecte et comme responsables.

Dans la phase de collecte, en mars, les délégués ont pu se consacrer entièrement au suivi du travail des agents recenseurs du strict point de vue de la méthode et de la rapidité d'exécution. Dans les recensements précédents, une grande part de leur activité était absorbée, à ce moment, par l'analyse des problèmes de classement de certaines communautés. La courbe d'utilisation du Minitel (messagerie de collecte) l'atteste : c'est en février, et non en mars, que les questions ont été les plus nombreuses.
L’importance du décalage (plus d’un mois) pose cependant certains problèmes pour les établissements à rotation rapide (hôpitaux) ou à mortalité élevée (maisons de retraite). La gestion de ces difficultés devra être anticipée dans les instructions aux délégués.

2.4. L’appui apporté aux agents recenseurs

Refus, personnes difficiles ou impossibles à joindre sont les obstacles principaux à l’avancement du travail des agents recenseurs :

- la campagne de publicité a été un incontestable succès tout comme le suivi des relations avec les médias ;

- de très nombreux maires ont, en milieu urbain, fait éditer une lettre destinée à leurs administrés, le plus souvent sur le modèle présenté dans les instructions aux maires. Cette lettre a été déposée dans les boîtes aux lettres par les agents recenseurs, pendant la tournée de reconnaissance ;

- on a beaucoup insisté sur la nécessité, pour les mairies, de gérer les rattrapages dès réception des notifications de refus ou de difficultés de collecte. Cette recommandation n’a pas toujours été appliquée avec la détermination requise. Il faudra insister davantage encore sur ce point. Par ailleurs, le traitement juridique des refus devra faire l’objet de négociations très précoces avec le Ministère de la Justice. Cela devrait aboutir à la mise à disposition des DR de deux modèles de lettres, l’une pour les réfractaires, l’autre pour les personnes difficiles à joindre, ainsi qu’à la définition des procédures appropriées ;


- ils ont été, en outre, beaucoup mieux suivis dans leur travail et stimulés par la nécessité de remettre assez tôt le relevé des immeubles au délégué. Ils ont été informés, enfin, des contrôles effectués sur leur travail à partir du fichier de la taxe d’habitation. (Voir § 6 ci-après).

- les avis de passage, produits à 5 millions d’exemplaires contre 2 millions en 1982, et les enveloppes pour la remise des imprimés (tirage total : 5 millions) ont été perçus favorablement par les habitants ;

- la présence de plus en plus fréquente des systèmes de digicodes ou d’interphones à l’entrée des immeubles semblait constituer un obstacle redoutable. Un paragraphe du manuel de l’agent recenseur était consacré à cette question. En définitive, les agents recenseurs ont surmonté ces problèmes sans trop de difficultés.

50 INSEE Méthodes n° 52 - 53
2.5. La suppression du numérotage des bulletins individuels

Lors des recensements précédents, l’agent recenseur devait reporter sur chaque bulletin individuel les numéros de district, d’immeuble et de logement. Cela ne compliquait pas trop l’exposé des principes de numérotation. En revanche, à un rythme de 20 secondes par bulletin, ce numérotage représentait trois heures de travail pour chaque agent.

Certains conseillers techniques, accoutumés aux méthodes antérieures, ont montré en 1989 une certaine hostilité à cette suppression, arguant notamment du risque de déclassement. Durant la collecte, cette opposition n’est pas réapparue.

2.6. Le contrôle de l’exhaustivité de la collecte à l’aide des fichiers de la taxe d’habitation (TH)

Une nouveauté du RP 90 a consisté en la réalisation d’un contrôle d’exhaustivité de la collecte par confrontation avec les données du fichier de la taxe d’habitation pour les communes de plus de 20 000 habitants. Ce contrôle portait sur le repérage effectif de tous les immeubles, sur le nombre de logements par immeuble, sur le classement des logements par catégories (résidence principale ou autres)\(^1\).

Les enseignements du recensement lui-même confirment ceux des essais. Même simplifié par le passage du relevé des logements au relevé des immeubles et par le contrôle sur les adresses, ce travail demeure très lourd, fastidieux et apparaît souvent peu productif.

Il faudra poursuivre l’effort entrepris, pour les raisons suivantes :

- les DR signalent des "trouvailles" intéressantes : 450 logements récupérés au Havre, 200 à Voiron, deux immeubles entiers à Bayonne, etc. ;
- la qualité et la rapidité de mise à jour du fichier TH s’améliorent d’année en année ;
- agents recenseurs et mairies se sont sentis contrôlés ;
- la présence des assistants-délégués assoit le dispositif INSEE en mairie et décourage certaines dérives fâcheuses en termes d’utilisation de fichiers administratifs variés.

\(^1\) On a cherché à apprécier l’efficacité de ce contrôle à travers l’enquête post-censitaire, dans ce numéro.
2.7. De nouveaux outils à la disposition des délégués et des mairies

Les documents prévus pour le suivi en 1982 avaient été critiqués. De nouveaux documents ont été élaborés en 1990. Ils ne seront pas détaillés ici, retenons seulement que malgré quelques critiques ponctuelles ils ont été dans la plupart des cas appréciés et ont joué un rôle positif dans la bonne marche de la collecte.

2.8. L’organisation des séances de formation des agents recenseurs par les délégués

En 1982, X. Charoy avait déjà réalisé un investissement très important en affiches de formation et en guide à l’usage des délégués. Les documents élaborés par J-N Pintard pour le RP 90 marquent une nouvelle et sensible progression. Ils ont été très appréciés en DR.

En 1990, les deux séances de formation ont eu lieu avant le début de la collecte. En 1982, la seconde séance de formation avait lieu après plusieurs jours de collecte. Elle était centrée sur le numérotage mais se résumait souvent à un exposé des récriminations des agents recenseurs. En 1990, la seconde séance a été largement consacrée à un approfondissement des acquis de la première. Ainsi, les délégués ont été déchargés ensuite de ce souci de formation (sauf réembauche), et la présentation des difficultés rencontrées par les agents recenseurs s’est faite sereinement, de semaine en semaine, par les visites en mairie (rendez-vous entre le délégué et chaque agent recenseur). Il faudra aller encore plus loin en supprimant, dans la seconde séance, tout exposé des règles de numérotation avec comme seul message : "Interdiction de commencer ce travail avant la fin de la collecte et hors de la présence du délégué".

La présentation de la tournée de reconnaissance comme phase autonome du travail des agents recenseurs entre les deux séances de formation a été extrêmement positive.

2.9. Calcul des dotations et acheminement

Les besoins en imprimés des communes et des DR ont été mieux évalués, grâce à l'actualisation des données du RP 82 à l'aide des évaluations faites grâce au fichier de

---

(1) Voir note 1733/RP90 du 6/8/90 p. 12
la taxe d’habitation. Cependant cette méthode présente certaines imperfections de principe. De plus, la mise à jour des fichiers n’est pas rigoureuse. Enfin, il semble que la reprise de l’activité depuis 1987 se soit traduite par beaucoup d’achèvement de logements et de réaffectations, mal prises en compte sur des fichiers réputés à jour en janvier 1989. Malgré ces défauts, l’utilisation de cette méthode constitue un progrès souligné par la plupart des DR.

La précision accrue des estimations de population a permis d’abaisser la marge de sécurité à +25 % pour les imprimés principaux contre +30 % à +40 % en 1982. Plus qu’une économie sur le budget global d’impression, nous avons recherché une meilleure répartition des imprimés afin d’éviter aussi bien les pénuries que les gaspillages.

Toutefois, il est vraisemblable que la pratique par de nombreux agents recenseurs urbains du dépôt des documents dans les boîtes à lettres se traduit par une sur-utilisation des bulletins individuels. C’est pourquoi, pour cet imprimé, on pourrait suggérer un taux de couverture de +30 % pour les communes de 50 000 à 200 000 habitants, de +35 % pour les communes de plus de 200 000 habitants et de +40 % pour la ville de Paris. Le manque de bulletins individuels dans certaines grandes villes a occasionné des réexpéditions coûteuses pour les DR ou des charges supplémentaires pour les conseillers techniques. A ce niveau, le souci d’économie s’est traduit par un gaspillage de moyens :

- d’autres imprimés ont fait défaut, par suite de la difficulté d’apprécier le lieu de leur utilisation : il s’agit des bulletins de membre de collectivité et des imprimés de l’enquête Famille. Le succès des affichettes 21 x 29,7 et 29,7 x 42 et des avis de passage a parfois conduit au même résultat. Il faudra en tenir compte.

- l’outil informatique de gestion des envois et des stocks a été bien conçu. Cependant, sa mise en place a été lourdement pénalisée par le conflit social de l’automne 1989. Son utilisation s’en est largement ressentie. Dans une situation "idéale", il aurait fallu que le système soit livré aux DR en juin, que l’ensemble des documents soit tiré avant novembre, et que le système accorde à ses utilisateurs un droit à l’erreur. Là encore, on peut penser que l’acquis 1990 rend cet objectif réaliste pour le prochain RP.

- les DR ont salué la décision de diviser par trois le poids des colis de documents et de réduire le seuil de population pour l’envoi direct des imprimés principaux aux communes. Un gaspillage de moyens très important a ainsi été évité aux DR, sans

(1) L’enquête Famille est menée conjointement au recensement sur un échantillon de districts ; remplie par les femmes de 19 à 64 ans, elle porte des questions complémentaires au RP sur la construction de la descendance, l’origine sociale de la femme, sa vie professionnelle et familiale.
répercussion sur la sécurité des approvisionnements, les erreurs du transporteur ayant été extrêmement rares.

- le choix de cartons renforcés et la mise à disposition des DR d'emballages complémentaires ont grandement facilité la préparation des envois et contribué à l'image de marque INSEE auprès des services municipaux. On peut penser aussi que ces conditionnements variés ont souvent permis d'éviter la détérioration des documents.

3 - Les actions de communication

3.1. La rémunération des agents recenseurs

C'est dans cette rubrique communication que nous classerons la redistribution du budget collecte entre les agents recenseurs. Notre expérience du RP 82 nous avait révélé le mécontentement des agents recenseurs des grandes villes.

Dans les centre-ville, en particulier, la présence de très nombreux logements inoccupés, le nombre réduit des personnes par ménage, la difficulté à joindre les habitants et parfois à pénétrer dans les immeubles, rendaient ce mécontentement légitime, les agents recenseurs étant en effet payés au bulletin rempli.

Il nous est apparu préférable de ne pas nous attaquer à cette difficulté par des taux de rémunération "à la carte", c'est-à-dire modulés selon les zones concernées ;

Nous avons opté pour une modification de la rémunération relative des différents imprimés, visant à favoriser les agents recenseurs "grandes villes" :

- le rapport entre rémunération du BI et rémunération de la FL était de 2,45 en 1982. Il a été ramené à 2. Il ne faut pas aller trop loin dans cette voie, sinon on s'exposerait à voir certains agents recenseurs classer comme logements beaucoup de locaux commerciaux ou de maisons en ruine ;

- le bordereau de district a été rémunéré à un taux correspondant à 10 FL. C'est dans les communes découpées en îlots que les agents sont susceptibles de recenser plusieurs districts. Cette mesure permet aussi de justifier devant les agents recenseurs le remplissage des bordereaux de district, qui ne faisait pas l'objet d'une rémunération en 1982 ;
- le remplissage du relevé des immeubles dans les communes de plus de 10 000 habitants a valu un forfait de 90 F aux agents concernés ;

Par ailleurs, une augmentation générale des différents taux a été permise par différentes mesures d’économie :

- la suppression du bordereau de maison pour l’habitat individuel a dégagé une somme à répartir de l’ordre de 30 millions de francs ;
- la spécialisation des agents recenseurs des communautés a dégagé une somme à répartir sur l’ensemble des agents de l’ordre de 3 millions de francs ;
- enfin, tous les agents recenseurs ont perçu un forfait de 27 F pour leurs dépenses de téléphone.

L’ensemble de ces mesures a permis d’accroître de quelques centaines de francs la rémunération des agents recenseurs des grandes villes.

Il est clair que c’est surtout de la réduction du travail en regard de cette rémunération (suppression de la numérotation des bulletins individuels) que provient l’amélioration de la situation des agents recenseurs.

3.2. La notice départementale d’information

Ce document est issu d’un constat. La notice 1982 n’avait qu’un seul mérite : rappeler le thème de la campagne de publicité à des gens qui avaient déjà ouvert leur porte à l’agent recenseur. Le recensement allemand de 1987 offrait une image plus caricaturale encore. La notice du recensement britannique de 1981 était aussi dépourvue d’intérêt, mais elle était illustrée de quelques dessins évoquant la vie des habitants : un cultivateur sur son tracteur, une usine, des personnes à un arrêt d’autobus, etc.

Nous avons pensé que le meilleur moyen de justifier le recensement était de présenter quelques résultats intelligibles pour le plus grand nombre et proches des préoccupations des habitants, c’est-à-dire ciblés géographiquement. Il y a là en outre un souci du service public. Aux personnes qui se prêtent à une enquête, l’INSEE restitue le fruit de leur participation passée :

73 % des personnes interrogées par BVA se souviennent d’avoir reçu la notice. Il est certes regrettable que beaucoup d’agents recenseurs aient omis de distribuer ce document et l’on peut sans doute insister davantage là-dessus. Néanmoins, un impact des 3/4 n’est pas un mauvais score. La non-distribution ne semble pas liée au caractère urbain/rural des communes.
Parmi les personnes ayant reçu la notice, la moitié déclarent l’avoir lue. Parmi celles-ci, les 3/4 déclarent l’avoir trouvée plutôt intéressante ou très intéressante. En définitive, 24 % des personnes de l’échantillon ont déclaré avoir reçu et lu avec intérêt ce document. Par ailleurs, les notices ont été présentées à beaucoup d’élèves des écoles et collèges. Il est raisonnable de penser que 10 millions de personnes ont lu la notice avec intérêt. Ce document a valorisé l’INSEE et le RP et a été repris par beaucoup de journaux peu avant le démarrage de la collecte, ainsi que par des bulletins municipaux.

Le passage de 2 à 4 pages de cette notice a été financé par la réduction du tirage des dossiers d’immeuble collectif de plus de 20 millions à seulement 4 millions d’exemplaires. Par ailleurs, les dotations communales ont été calculées sur la base du nombre estimé de résidences principales et non du nombre estimé de logements comme en 1982. On pourrait d’ailleurs faire valoir que le passage à 4 pages a en fait réduit le gaspillage en conférant quelque intérêt à un document institutionnel et donc obligatoire qui n’en avait pratiquement aucun.

Les DR ont massivement approuvé ce projet, en regrettant l’austérité du papier utilisé. Nous pensons que l’économie de 1 500 000 F réalisée sur ce poste était nécessaire au vu de la contrainte budgétaire.

La préparation des 96 modèles départementaux a été très lourde. En amont, les itérations pour figer le produit sur sa forme et son contenu ont été fort laborieuses. La présence de personnel supplémentaire devrait faciliter les travaux pour le prochain RP.

3.3. La présentation des documents de collecte

L’aspect et la clarté des principaux imprimés ont été largement améliorés en 1990.

Autant le sacrifice sur la qualité du papier de la notice nous semblait acceptable, autant il nous paraissait indispensable de s’en tenir à du papier de la meilleure qualité pour les deux questionnaires remis aux ménages. Conformément à nos prévisions, la feuille de logement et le bulletin individuel ont été présentés dans chacun des journaux télévisés du 5 mars.

De façon générale, il est essentiel de maintenir une politique de qualité. Agents recenseurs, délégués, habitants, ont reçu des documents très soignés. Notre sentiment est qu’ils en ont retiré l’impression d’être traités comme des partenaires et non comme des entités comptables ou des exécutants corvéables. C’est, en définitive, ce que le secteur privé baptise "respect du client".
Nous manifestons un regret, celui de n’avoir pas consacré plus de moyens à la page de couverture des instructions aux maires. Une illustration par des photos aurait été appréciée et le taux de lecture aurait été meilleur.

3.4. La vidéo

Il nous semblait difficile de faire l’impasse sur cette technique, même si nous ne pouvions lui consacrer beaucoup de temps. En fait, c’est le scénario "Collecte" pour l’essai de 1988 qui a été très coûteux en heures de préparation. Au total, l’écriture du scénario des trois modules du RP 90 n’a représenté qu’un mois de travail.

Nous avions rejeté d’emblée la conception "télévision scolaire" qui nous semblait débilitante. A cet égard, les expériences allemandes et australiennes avaient complètement échoué. Il nous a paru plus efficace de "raconter l’histoire" d’un agent recenseur ou d’un délégué. Moins didactique, cette démarche est plus conviviale : la personne qui assiste à la formation se sent davantage "accueillie" puisqu’on prend soin de lui présenter un spectacle et non une accumulation de règles. Une présentation réaliste du travail permet de rassurer les personnels sur le contenu des tâches à effectuer. C’est aussi un intermède.

Ainsi défini, l’outil vidéo doit être relativisé. Son utilité est très grande et il faudra persévérer dans cette voie, mais ce n’est pas l’essentiel des actions de formation. La "bataille d’Hernani" autour du module "Numérotation", certainement inadéquat, était bien excessive de part et d’autre.

3.5. Les relations avec les mairies

La lettre aux maires envoyée en juillet 1989 à toutes les mairies est un progrès important. Les instructions aux maires devront être expédiées séparément des autres imprimés du RP et plus tôt. Nous avions gagné un mois sur la date de disponibilité de ce document (novembre 1989) mais il faudrait faire encore mieux.

Le renforcement de l’encadrement en délégués et conseillers techniques devra être poursuivi. Peut-être faudra-t-il envisager, pour l’Île-de-France, des délégations supplémentaires de travaux à d’autres directions régionales après le succès du recensement de la Seine-et-Marne par la DR de Reims.
L’un des axes d’effort pour le prochain recensement devra être la formation des personnels municipaux par les conseillers techniques. Il y a là une marge de progression considérable sur le suivi de la collecte et la préparation des opérations.

**3.6. Les relations DG-DR**

Les multiples essais entrepris ont été l’occasion de créer un climat de collaboration.

Toutes les instructions principales mais aussi beaucoup de documents de collecte ont été soumis aux responsables régionaux. Leurs remarques ont été largement prise en compte. C’est là une innovation importante.

Le minitel a été pour les DR un apport essentiel. Partout, on se demande "comment on faisait auparavant" pour assurer le lien responsable régional RP -conseiller technique-délégué.

L’intérêt du minitel a été presque aussi grand pour les relations DR-DG. À ce stade, pourtant, notre dispositif a été imparfait. Une personne ne suffit pas à préparer les réponses aux questions des DR. Il faudrait une petite équipe permanente. Malheureusement, la collecte a eu lieu dans une période où les autres responsables de la division étaient surchargés.

Les fiches de collecte transmises par la messagerie interne à l’INSEE de la DG aux DR ont été un outil efficace mais certaines dispositions ont été contestées par les DR.

De façon générale, les questions posées par les DR appellent une décision très rapide. Il faut "légitérer" de façon quasi-instantanée. Pour améliorer la qualité de notre réponse globale aux besoins des DR, il faudra, outre le travail d’équipe évoqué plus haut :

- avoir gagné deux mois sur la mise de tous les imprimés à disposition des DR ;
- avoir intégré en annexe au manuel du délégué certains cas particuliers rencontrés en 1990 ;
- avoir verrouillé le système de recensement des communautés en envisageant tous les cas particuliers.

Pour conclure ce point sur une note plus optimiste, le dialogue par minitel a permis de traiter de façon ultra-rapide les manques d’imprimés en DR et les demandes d’information sur les personnels.
4 - Quelques thèmes à approfondir

4.1. Les relations avec le Ministère de l’Intérieur

Il conviendrait, avant et pendant la collecte, de disposer d’une structure légère de concertation permanente.

Il nous paraît souhaitable de faire le point sur les conséquences financières, pour les communes, des changements démographiques et des changements de définition apparus en 1990, en mettant l’accent sur certains aspects comme les effets de seuil. Cela pourrait être l’occasion de publier dans Economie et Statistique un article comparable à celui de M. Vernet, "Le nombre et la loi". Il faudrait aussi explorer certaines hypothèses de travail sur les problèmes de localisation des habitants. (Voir ci-après).

4.2. Analyse de la qualité du remplissage des questionnaires

Une étude précise à partir d’un échantillon représentatif devrait permettre d’évaluer la qualité des réponses aux questions. À titre d’exemples, sur le bulletin individuel, il sera intéressant d’analyser :

- la compréhension de l’aiguillage séparant actifs et inactifs ;
- la qualité des réponses lorsqu’une précision est demandée (nationalité détaillée, ancienne profession, etc.) ;
- la précision du libellé de la profession à la question 12 ;
- le refus de déclarer l’employeur ;
- la qualité des réponses aux 3 dernières questions.

---

(1) Les résultats du RP reçoivent de nombreuses utilisations réglementaires, généralement fonction de l’appartenance de la population de la commune à une tranche donnée ; il y a donc des effets de seuil liés à l’intérêt présenté par le passage à une tranche supérieure.

(2) Économie et Statistiques N°36, Juillet-Août 1972.
Cette étude peut être menée de façon autonome. Elle peut être menée aussi par
comparaison avec un autre échantillon de bulletins du RP 82 (si tout n’a pas été détruit)
et par appariement avec l’enquête Emploi. Ce dernier travail avait été réalisé par
J.-P. Baraille en 1987 mais sur la base d’un bulletin de 4 pages qui devait ensuite être
abandonné.

4.3. La localisation des habitants

S’il n’y avait pas eu de problème sur ce sujet, le déroulement de la collecte aurait été
tout à fait tranquille. Cette question a été beaucoup plus ressentie qu’en 1982 pour
plusieurs raisons : les difficultés ont été atténuées sur la plupart des autres dossiers du
recensement ; certaines des populations concernées se sont fortement accrues.

Pourtant, les acquis du RP 90 ne sont pas négligeables. La création d’un nouvelle
catégorie de logement, les logements utilisés occasionnellement, a été positive. De
même, les nomenclatures d’établissements et de collectivités ont été largement simplifiées
et la terminologie est moins désuète : c’est une satisfaction pour nous que d’avoir
"tordu le cou" aux établissements "de population comptée à part" (création de la
monarchie de juillet) et aux ménages "ordinaires" qui ont encombré si longtemps nos
publications. La correspondance d’un type de feuille d’établissement à chaque type
d’établissement a sans doute été une bonne chose.

En revanche, il faudra "remettre à plat" les questions suivantes :

- imputation de population aux communes de résidence secondaire et de logement
  utilisé occasionnellement ;

- recensement des étudiants ayant une chambre en ville universitaire ;

- circulation des bulletins des membres de collectivités ;

- recensement des militaires en stage ;

- liste figurant sur la feuille de logement.

Tous ces points devront être abordés avec le souci du réalisme et celui de la rigueur.
D’une part, il faut éviter des constructions intellectuelles trop complexes et inapplica-
bles sur le terrain, tout en cherchant à se concilier l’accord des municipalités. D’autre
part, tous les outils d’analyse doivent être mis en œuvre pour mesurer les enjeux au
regard des contraintes méthodologiques envisageables. Il convient en effet de garder à
l’esprit que, en dépit de ces problèmes, le recensement a été mené à bien.
2.3 - L’ENQUÊTE DE MESURE DU DEGRÉ D’EXHAUSTIVITÉ DU RECENSEMENT DE 1990

Nicole Coeffic

Le propre d’un recensement est de viser à l’exhaustivité du dénombrement des personnes. Malheureusement, malgré tout le soin que l’on peut apporter au recueil des données, des erreurs sont inévitables. Or, le recensement est une opération coûteuse, qui tient une place fondamentale dans notre système de connaissance statistique. Il est donc important de tenter d’apprécier la marge d’erreur commise, d’analyser quelles sont les causes de ces erreurs, et de déterminer si certaines populations sont moins bien couvertes que d’autres. À la fin du recensement de 1990, l’INSEE a donc mené une enquête spécifique, ayant pour objectif principal une mesure et une étude des erreurs de dénombrement.

Cette enquête post-censitaire n’est qu’un élément d’un plus large dispositif mis en place par l’INSEE et consacré au contrôle de la qualité, non seulement de la collecte, mais aussi du chiffrement et de la saisie des documents du recensement de 1990 (voir article de G. Badeyan et de J.C. Deville). Mais, à la différence des autres opérations menées dans le cadre de ce dispositif, l’enquête n’a pas pour but d’agir sur cette qualité et d’améliorer les résultats ; elle est seulement un instrument d’étude méthodologique.

Pour mettre au point cette enquête, nous nous sommes inspirés d’une première expérience que l’INSEE avait menée après le recensement de 19621. Que l’INSEE n’ait pas réalisé d’enquête post-censitaire entre 1962 et 1990 a sans doute pour explication qu’une telle étude représente une tâche très délicate, comme on le verra.

D’après l’enquête post-censitaire de 1990, le taux d’omission de la population au recensement est proche de 2% ; ces omissions sont, d’une certaine façon, compensées partiellement par des doubles compte, qui représentent un peu moins de 1% de la population.

Après avoir décrit la méthode de l’enquête, nous en présenterons les résultats. Les points sur lesquels la méthode de 1990 diffère de celle de 1962 sont mentionnés en annexe.

1 - Les grands principes du recensement et les origines possibles des erreurs de recensement

Afin de mieux faire comprendre certaines des options retenues pour la méthode de l’enquête, nous rappellerons en préalable selon quels grands principes la population a été recensée, et les sources possibles des erreurs de dénombrement.

Au recensement, les personnes, qui ne vivent pas dans une "communauté", doivent être recensées dans leur résidence principale, c’est-à-dire le logement "où le ménage demeure la plus grande partie de l’année". Les élèves internes et les militaires sont recensés respectivement dans leur internat ou leur caserne, mais le bulletin individuel qu’ils ont rempli dans l’une de ces institutions doit être "réintégré" dans le logement correspondant à leur résidence familiale (généralement celle de leurs parents), s’ils ont donné au recensement l’adresse de cette résidence familiale. Les personnes vivant dans d’autres communautés (cité universitaire, foyer de travailleurs...) sont recensées dans ces communautés.

Les principales sources des erreurs de recensement sont les suivantes :

- des logements ou des communautés peuvent être omis. Si le logement omis est une résidence principale ou une communauté, leurs occupants ont toute chance d’être omis au recensement, sauf si, par exception, ils sont recensés ailleurs ;

- des logements peuvent être classés dans une catégorie erronée. Ainsi des résidences principales peuvent être considérées à tort comme "résidences secondaires", "logements utilisés occasionnellement" pour des raisons professionnelles ou "logements vacants", donc recensées sans occupants. Ce type d’erreur se produit principalement pour les ménages absents pendant la période de recensement, ceux dont les membres effectuent des déplacements fréquents et ont des horaires particuliers, ceux qui partagent leur temps entre deux résidences... Les personnes qui résident dans ces résidences principales mal classées risquent alors d’être omises au recensement, sauf si elles sont recensées ailleurs (par exemple, par erreur, dans leur résidence secondaire).
À l'inverse, des résidences secondaires, des logements utilisés occasionnellement ou des logements vacants peuvent être classés à tort résidences principales. Les personnes qui y sont recensées risquent alors d'être comptées deux fois au recensement, si elles sont aussi recensées dans leur résidence principale. Ces classements erronés peuvent provenir d'erreurs de collecte, mais aussi résulter de fraudes. On sait en effet qu'il arrive que des mairies cherchent par ce biais à "gonfler" le nombre des habitants de leur commune.

Une mairie qui voudrait frauder peut aussi créer des logements avec une adresse fictive et des occupants fictifs.

Certaines personnes refusent de se faire recenser. Certes, ces refus n'entraînent pas systématiquement des omissions au recensement. En effet, si tout le ménage refuse, normalement, l'agent recenseur doit tout de même le recenser. Il a la consigne de se renseigner auprès des voisins, et d'insérer dans une feuille de logement le nombre de bulletins individuels qui convient et portant, pour le moins, la mention "refus". Bien évidemment, cette situation risque tout de même de provoquer des erreurs : l'agent recenseur peut oublier d'appliquer cette consigne ; ou obtenir des voisins des renseignements erronés sur le nombre de bulletins individuels à créer.

Par contre, si, à l'intérieur d'un ménage, une personne veut échapper au recensement, alors que les autres membres du ménage remplissent un bulletin individuel, l'erreur de recensement est quasiment certaine, car il y a peu de chances que l'agent recenseur puisse repérer ce type d'omission et la corriger.

Des personnes peuvent aussi être omises isolément dans leur résidence principale, si la personne qui remplit les documents du recensement oublie de les déclarer. Ces omissions involontaires concernent assez souvent les personnes ne faisant pas partie du noyau familial (salariés logés chez leur employeur, étudiants logés chez un particulier, personnes âgées hébergées chez leurs enfants...).

À l'inverse, dans une résidence principale recensée, des personnes peuvent être comptées en trop, si elles résident la plus grande partie de l'année ailleurs. Elles risquent alors d'être comptées deux fois. Par exemple, un étudiant, qui vit habituellement dans une chambre en ville, peut, par erreur, être déclaré par ses parents comme habitant encore avec eux.

De même, les personnes vivant dans une communauté et ayant conservé des liens avec un logement familial peuvent être recensés à la fois dans cette communauté et dans ce logement. La méthode de recensement est conçue pour limiter le risque de double compte pour les élèves internes et les militaires du contingent, du fait que l'on réintègre le bulletin collecté dans la caserne ou l'internat au logement de la résidence familiale (on peut alors supprimer les doubles, au moment de la réintégration), mais on n'a pas prévu les mêmes précautions pour les membres des autres communautés.
Enfin, le déménagement pendant la période de recensement est source d’erreurs. La période de recensement des logements s’est étalée sur plusieurs semaines. Quand un ménage, ou une personne dans un ménage, change pendant ce laps de temps de résidence principale, il risque d’être omis, ou au contraire d’être compté deux fois.

Ce même type de problème se pose pour les personnes ayant quitté un logement pour vivre en communauté ou vice-versa. Le recensement des établissements militaires a commencé le 12 février 1990, et celui des autres communautés le 29 janvier, alors que celui des logements n’a débuté que le 5 mars.

2 - La méthode de l’enquête post-censitaire de 1990

2.1. Place de l’enquête post-censitaire

Nous situerons ici la présente enquête post-censitaire parmi d’autres méthodes possibles de la mesure de la qualité du dénombrement des personnes par le recensement.

Une première méthode aurait pu consister à tirer des échantillons dans des fichiers administratifs (fichier de sécurité sociale, fichier de retraités...) et à rechercher les individus ainsi désignés dans les documents du recensement. Cette méthode n’aurait pas permis de détecter des doubles comptes, mais aurait pu permettre de se faire une idée des omissions. Mais leur mesure aurait été assez imparfaite, car on aurait compté comme omise une personne n’ayant pas été recensée à son adresse administrative, alors qu’elle aurait pu être recensée ailleurs. Autre inconvénient de la méthode : les fichiers dont on aurait pu disposer ne couvrent que des champs partiels de la population vivant en France, et on n’aurait pas été capable de donner une estimation globale de la qualité du recensement.

Une étude des erreurs de recensement, du moins des doubles comptes, pourrait être faite par l’INSEE à partir de son "échantillon démographique permanent", appelé également "panel démographique". Il s’agit d’un fichier regroupant toutes les personnes nées certains jours de l’année, enrichi en permanence par les données de l’état civil et celles des recensements successifs. Pour introduire dans le panel les renseignements collectés lors du recensement de 1990, on a donc isolé les bulletins individuels des personnes nées pendant les jours de l’année désignés, et cherché à les apprêcher avec les personnes nées à la même date et de même nom figurant dans l’échantillon démographique. En conséquence, on doit pouvoir compter les personnes du panel pour lesquelles on n’a pas pu retrouver de bulletin individuel rempli au recensement de 1990, et celles pour lesquelles on en a trouvé deux. Toutefois, cette méthode ne permet sans doute pas de mesurer correctement les omissions. En effet, si, pour une personne, on n’a pas trouvé de document de recensement, ce n’est pas nécessairement parce qu’elle a été omise.
Sa date de naissance peut être, au recensement, fausse ou non renseignée (ce type d’erreur n’est pas rare) ; son bulletin individuel est alors introuvable. Mais, surtout, elle peut ne pas habiter en France, et il serait alors abusif de la compter comme omise. Par contre, une mesure des doubles comptes à partir du panel donnera très certainement des résultats plus fiables, même si leur nombre risque d’être quelque peu sous-estimé, en raison des erreurs sur les dates de naissance au recensement. Malheureusement, pour le moment, ces résultats ne sont pas disponibles.

La méthode de l’enquête post-censitaire consiste à dénombrer tous les logements situés dans un échantillon de zones géographiques. Pour chaque logement, un questionnaire approfondi permet d’en connaître les occupants, même occasionnels, de noter comment ces occupants partagent leur temps entre ce logement et d’éventuels autres lieux d’habitation, d’obtenir l’adresse précise de ces autres lieux de résidence. Puis, dans un premier temps, on recherche tous les documents du recensement remplis pour les logements situés sur les zones-échantillon, et on regarde si les personnes enquêtées ont été recensées à l’adresse du logement enquêté. En cas d’incohérence inexplicable entre l’enquête et le recensement, les enquêteurs doivent retourner sur le terrain pour obtenir des explications et corriger l’enquête si elle comporte des erreurs. Dans un deuxième temps, on regarde dans les documents du recensement si les enquêtés ont été ou non recensés à l’une de leurs éventuelles autres adresses (autres que celle du logement enquêté). Par cette méthode, on peut donc déterminer, pour chaque enquêté, s’il a été recensé 0 fois, 1 fois (à l’adresse du logement enquêté ou à une autre adresse) ou 2 fois.

2.2. La collecte

2.2.1 Calendrier

Le calendrier de la collecte était soumis à deux contraintes. L’enquête devait se dérouler le plus près possible de la date de fin du recensement. En effet, des mouvements pouvaient se produire entre la date du recensement et celle de l’enquête : déménagements, décès et naissances. Ceux-ci rendaient le travail de comparaison entre le recensement et l’enquête plus délicat. On avait donc intérêt à ce qu’ils soient les moins nombreux possibles. À l’inverse, il fallait attendre que les documents du recensement remplis soient revenus à l’INSEE, ou du moins en mairie, pour envoyer les enquêteurs sur le terrain. Il fallait en effet éviter que le recensement ne puisse être modifié du fait du passage de l’enquêteur, puisque le but de l’enquête n’était pas d’améliorer le recensement, mais d’en mesurer les limites.

Finalement, la majorité des interviews (98 %) ont eu lieu entre la mi-avril et la fin juin (72 % pour le seul mois de mai), une petite minorité ont été réalisées plus tardivement.
2.2.2 Échantillon et champ de la collecte

L'enquête a été réalisée auprès d'un **échantillon aréolaire**. Cet échantillon n'a pas été tiré pour la circonstance ; il avait été tiré en 1983 sur la base du recensement de 1982 et aurait dû être utilisé pour une des enquêtes sur l'Emploi d'octobre. La série des enquêtes Emploi d'octobre ayant été supprimée (seule celle de mars a été conservée), on a pu utiliser un échantillon qui n'avait jamais servi.

Les aires avaient été tirées en 1983 à l'intérieur de strates définies par le croisement des 21 régions et de 10 catégories de communes (2 catégories de communes rurales, et 8 d'unités urbaines classées en fonction du nombre d'habitants\(^1\)). Chaque aire avait des limites repérables sur le terrain ; dans les zones constituées de gros immeubles collectifs, les aires étaient constituées de parties d'immeuble bien identifiées : par exemple, les logements de tel à tel étage ou les logements donnant sur telle cage d'escalier... Des plans d'aire précis avaient été mis au point lors du tirage.

L'échantillon de l'enquête post-censitaire, représentatif de l'ensemble du territoire métropolitain (Corse exceptée), était constitué d'environ 650 aires d'une quarantaine de logements, représentant au total à peu près 30 000 logements. Le taux de sondage correspondant était d'environ 1/900.

Lors de l'enquête post-censitaire, l'enquêteur devait repérer de manière exhaustive tous les logements se trouvant à l'intérieur des limites des aires-échantillon, y compris les logements neufs construits depuis la date du tirage de l'échantillon. A la différence de ce qui est fait pour l'enquête Emploi, on n'a pas tiré **d'échantillon spécial** représentant la construction neuve des immeubles construits sur permis de plus de 10 logements depuis 1982, faute de moyens. En conséquence, on peut supposer que les gros immeubles collectifs, construits depuis 1982, sont assez mal représentés dans l'enquête.

L'INSEE a choisi d'utiliser un échantillon aréolaire, parce que c'était la seule solution permettant de repérer à l'enquête les logements oubliés au recensement, grâce à un *ratissage* systématique des logements situés sur les aires-échantillon, effectué par des enquêteurs de qualité.

Par contre, ce type d'échantillon a un inconvénient. Certaines erreurs au recensement peuvent dépendre de circonstances locales (capacités de l'agent recenseur, structure de l'habitat) et ne pas être réparties uniformément sur tout le territoire. Or, l'échantillon aréolaire enregistre ces *effets de grappe*, ce qui diminue un peu la précision des estimations. Toutefois, les aires tirées sont de petite taille, ce qui réduit les effets de grappe.

---

\(^1\) Enquête sur l’emploi de 1989 - résultats détaillés, INSEE RÉSULTATS, Série Emploi-Revenus, N° 6-7

INSEE Méthodes n° 52 - 53
Du fait de son échantillon aréolaire, l’enquête mesure assez mal le volume des fraudes pouvant être commises par certaines mairies et consistant en ajouts de logements qui n’existent pas ou de personnes dans des logements déjà recensés. Ces fraudes sont en effet étroitement localisées. De plus, les logements qui, n’existent pas, sont créés frauduleusement avec une adresse fictive, ont parfois une probabilité quasiment nulle de figurer dans une aire d’un échantillon aréolaire. Heureusement, en 1990, les risques de fraude ont été limités au maximum par un accroissement des opérations de contrôles et de suivi pendant la collecte du recensement.

L'idéal aurait été de pouvoir tirer un échantillon spécifique pour l’enquête, permettant de sur-représenter les zones où les risques d’erreur de recensement étaient supposés importants. Mais cela n’a pas été possible.

Seuls les logements étaient dans le champ de l’enquête : autrement dit, les communautés (au sens du recensement), telles que casernes, internats, foyers de travailleurs... n’étaient pas enquêtées. Cependant, on verra plus loin que les personnes logeant dans ces communautés sont prises en compte dans l’enquête, dès lors qu’elles conservaient un lien avec un ménage ordinaire et étaient enquêtées dans ce ménage. Les personnes sans domicile fixe et les personnes vivant habituellement dans une habitation mobile sont aussi hors champ de l’enquête.

2.2.3 Exécution de l’enquête sur le terrain

Tout d’abord l’enquêteur devait bien repérer sur le terrain les limites de l’aire-échantillon qu’il avait à enquêter, et, à l’intérieur de ces limites, faire un inventaire complet des immeubles et des logements. Il devait faire la liste de tous les immeubles, en précisant leur adresse et quelques-unes de leurs caractéristiques, sur un état descriptif d’aire. Pour chaque immeuble comportant deux logements d’habitation ou plus, l’enquêteur remplissait un état descriptif d’immeuble, donnant la liste des logements avec le nom de leur occupant et leur situation sur le palier. Pour éviter qu’il n’oublie un logement, on demandait à l’enquêteur de signaler aussi les locaux non habités, mais pouvant servir à l’habitation, et de ne pas oublier les pièces indépendantes, telles les chambres de bonne rattachées à un logement principal.

Pour chaque logement, l’enquêteur devait remplir un questionnaire A. Ensuite, dès qu’il réussissait à contacter les occupants du logement, l’enquêteur devait ouvrir un questionnaire B.

L’enquêteur ne devait pas se contenter d’interviewer les ménages déclarant que le logement enquêté était leur résidence principale, mais il devait interviewer tous les ménages qu’il lui était possible de joindre dans l’aire, et n’arrêter toute tentative de prise de contact que s’il avait l’assurance (par les voisins généralement) que le logement était occupé moins de 3 mois par an. On voulait, par cette méthode, couvrir un champ...
un peu plus large que les résidences déclarées principales par leurs occupants, et ne pas risquer de manquer les résidences que les enquêtés (ou leurs voisins) déclarent secondaires, alors qu'en réalité elles sont occupées par les ménages plus longtemps sur l'année que leur(s) autre(s) résidence(s). En effet, on a vu qu'un certain nombre d'omissions de personnes au recensement peuvent provenir de ce que certaines résidences principales sont classées à tort secondaires. Il fallait éviter que les enquêteurs, à leur tour, ne reproduisent le même type d'erreur.

Quand le logement était vacant, ou occupé moins de 3 mois par an, ou que le ménage ne pouvait être interrogé suite à un refus ou une absence pendant toute la période d'enquête, un questionnaire A était rempli, mais très succinctement : l'enquêteur n'avait à y inscrire que la catégorie du logement : vacant, résidence secondaire, son adresse précise, et, pour les ménages non répondants, quelques renseignements (fournis par les voisins) pouvant servir au redressement des non réponses, tels que le nombre, le sexe et l'âge approximatif des occupants du logement.

Quand l'enquête était acceptée par un ménage, le questionnaire A était complété par quelques renseignements portant sur le logement, et un questionnaire B était ouvert. Sur ce questionnaire B étaient portées les données principales suivantes :

- liste des personnes, résidant même très occasionnellement, dans le logement, ou y ayant résidé depuis le 1er janvier ;

- puis, pour chaque personne :

  . date de son arrivée éventuelle dans le logement ou de son départ (en cas d'emménagement, déménagement, décès, naissance après le 1er janvier) ; durée approximative d'occupation du logement depuis un an ;

  . liste des autres résidences, dont elle dispose en permanence (par exemple, résidence secondaire...), ou dans lesquelles elle a logé, même occasionnellement, depuis le 1er janvier (par exemple, logement de ses parents, établissement collectif : caserne, internat, foyer...) ; adresse précise de ces autres résidences et durée approximative d'occupation depuis un an ;

  . enfin, toujours pour chaque personne, renseignements sur son état civil (sexe, âge, nationalité...), et description de sa situation professionnelle, dans un Bulletin individuel, proche de celui du recensement, quoique un peu allégé.

(1) En réalité, ils n'ont pas été utilisables, car trop peu souvent fournis.
Sur ce questionnaire B, on peut faire plusieurs remarques :

- il permettait de connaître les mouvements récents de personnes dans le logement (emménagements, déménagements...) susceptibles d’expliquer les divergences entre la liste des personnes enquêtées et celle des personnes recensées. Au lieu de prendre comme période de référence, pour l’observation de ces mouvements, l’intervalle séparant la date du recensement et celle de l’enquête, on a préféré faire débuter cette période au 1er janvier, pour deux raisons : on ne voulait pas focaliser l’interview sur la date du recensement, pour éviter que les enquêtés ne cherchent à être cohérents avec ce qu’ils avaient déclaré lors du recensement. Par ailleurs, couvrir une période plus large que celle commençant à la date du recensement nous paraissait permettre d’éviter une sous-estimation de ces mouvements ;

- le questionnaire était conçu de manière à enquêter la population la plus large possible. Au stade de l’enquête sur le terrain on voulait éviter avant tout d’omettre des personnes. Ainsi, dans le questionnaire, étaient rattaché au logement enquêté un nombre très large de personnes, notamment toutes celles qui y résidaient occasionnellement. En effet, bien que résidant très occasionnellement dans un logement enquêté, certaines catégories de personnes pouvaient n’avoir aucune chance d’être prises en compte par l’enquête ailleurs : celles qui effectuent des déplacements longs et fréquents ; celles qui vivent habituellement dans une communauté ; les personnes, sans résidence personnelle, qui sont hébergées tantôt chez les uns, tantôt chez les autres... Aussi, après que l’enquêté eut donné spontanément la liste des occupants du logement, l’enquêteur lui demandait de citer toutes les personnes ayant habité le logement depuis le 1er janvier, puis toutes les personnes absentes et logeant temporairement ailleurs, "par exemple, jeune vivant en caserne... ; personne à l’hôpital ou logée ailleurs pour son travail". Enfin, il posait une question sur la présence éventuelle de locataires ou de sous-locataires, de salariés logés, d’enfants en nourrice. Ces catégories risquaient en effet d’être souvent spontanément omises à l’enquête (comme au recensement) ;

- on demandait aux enquêtés la liste de leurs autres résidences, pour pouvoir vérifier ultérieurement s’ils y avaient été ou non recensés. Là encore, on a essayé de "ratisser large". L’objectif était de ne pas omettre d’"autre résidence", et on a considéré qu’il n’était pas gênant de prendre en compte, à l’inverse, des "autres résidences" en trop, c’est-à-dire des résidences où les personnes avaient très peu de chances d’avoir été recensées.

2.3. Premier appariement manuel et retours sur le terrain

Une fois la collecte terminée, dans chaque direction régionale de l’Insee, le personnel chargé de l’enquête devait procéder à un appariement manuel des questionnaires de
l’enquête et des documents remplis au recensement, pour tous les logements situés sur les aires de l’enquête. Il s’agissait de regarder si on avait bien repéré les mêmes logements lors de l’enquête et lors du recensement sur le territoire correspondant à l’aire-échantillon. En même temps, on devait apparié les personnes.

L’appariement des logements n’était pas toujours simple. Par exemple, dans les zones rurales, les adresses sont souvent très peu formalisées et étaient très imprécises sur les documents du recensement, en particulier dans les hameaux. Pour essayer de rapprocher les logements enquêtés de ceux qui avaient été recensés, il a fallu parfois retourner sur le terrain.

Mais l’appariement des personnes était encore plus délicat. En effet, le champ de l’enquête était plus large que celui du recensement, puisqu’on enquêtait des personnes dont le logement enquêté n’était pas la résidence principale ; par ailleurs, des mouvements de personnes pouvaient s’étre produits entre la date du recensement et celle de l’enquête. Certaines divergences entre le recensement et l’enquête étaient donc normales. Ainsi, il était normal qu’une personne enquêtée n’ait pas été recensée à l’adresse du logement enquêté, si elle y avait emménagé après la date du recensement ; si le logement enquêté n’était pas sa résidence principale ; si, à la date du recensement, elle n’était pas encore née, etc. Par contre, chaque fois qu’une divergence entre le recensement et l’enquête était, au vu des renseignements collectés lors de l’enquête, inexplicable, la direction régionale devait demander aux enquêteurs de retourner sur le terrain pour obtenir des compléments d’information. Ces retours auprès des enquêtés permettaient non seulement de compléter les questionnaires, mais aussi de corriger d’éventuelles erreurs commises lors de l’enquête. Il fallait en effet absolument éviter que des erreurs de l’enquête puissent être imputées à tort au recensement.

La difficulté de ce 1er appariement était donc de bien isoler les cas de non-appariements nécessitant un retour terrain.

En outre, lors de ce 1er appariement, les DR étaient amenées à vérifier que les personnes enquêtées n’avaient pas omis de déclarer une résidence : par exemple, elles devaient s’assurer que les emménagés récents avaient bien déclaré l’adresse de leur résidence antérieure (car ils pouvaient y avoir été recensés). Si une adresse était manquante, un “retour terrain” devait être demandé.

2.4. Deuxième appariement manuel

Rappelons que, sur le questionnaire de l’enquête, figurait, pour chaque personne enquêtée, la liste de toutes leurs autres résidences, autres que celles correspondant au logement enquêté. Un 2e appariement (complètement séparé du 1er) a consisté à
rechercher les documents du recensement correspondant à toutes ces autres résidences, et à noter si les enquêtés y avaient été ou non recensés.

Les autres résidences des personnes, bien souvent, n’étaient pas situées dans la même région que le logement où elles avaient été enquêtées. En conséquence, il a fallu prévoir que des fiches-navette puissent circuler entre directions régionales. Chaque direction régionale devait recopier sur des fiches-navette, à partir des questionnaires, l’adresse précise des autres résidences à appairer, le nom et la date de naissance des enquêtés y résidant, puis transmettre ces fiches-navette pour appariement à la direction régionale compétente.

Quand l’autre résidence n’était pas retrouvée dans les documents du recensement, les agents devaient vérifier sur le Minitel que l’adresse donnée par l’enquêté n’était pas erronée. Si, par le Minitel, ils trouvaient une autre adresse, une nouvelle recherche était tentée.

Ce deuxième appariement n’a pas permis dans tous les cas de conclure au recensement ou au non recensement de l’enquêté dans son autre résidence. Les adresses données à l’enquête étaient parfois manquantes, ou trop incomplètes, ou visiblement fausses.

Si l’adresse donnée par l’enquêté était précise mais ne figurait pas sur le Minitel, et qu’on ne retrouvait aucune trace de l’autre résidence dans les documents du recensement, on en a conclu que, probablement, les enquêtés n’avaient pas été recensés dans cette autre résidence. En effet, on n’en était pas tout à fait certain puisqu’il y avait un petit risque que l’adresse soit fausse et que l’enquêté ait été recensé à une autre adresse. Toutefois, il était fort probable que, dans la très grande majorité des cas, l’adresse soit bonne, bien que ne figurant pas sur le Minitel.

Bien évidemment, si l’adresse de l’autre résidence figurait sur le Minitel et qu’on ne retrouvait aucun document de recensement correspondant, on devait conclure que cette autre résidence avait été omise au recensement et que, à coup sûr, les enquêtés n’avaient pas été recensés dans cette résidence.

### 2.5. Exploitation

Lors de la collecte, on devait enquêter toutes les personnes logeant ou ayant logé depuis le 1er janvier 1990, même très occasionnellement, dans les logements de l’échantillon. Si, au stade de l’exploitation, on avait gardé dans le champ des tableaux tous les enquêtés affectés de leur coefficient de pondération initial, on aurait sur-représenté les personnes ayant logé dans plusieurs logements entre le 1er janvier et la date de l’enquête, puisque celles-ci avaient une probabilité plus forte d’apparaître dans l’enquête.
Il s’agissait donc d’éviter ce biais. Pour ce faire, on a choisi la méthode suivante.

On a déterminé quelle était, pour chaque enquêté, sa résidence principale à la date de l’enquête. Si cette résidence principale correspondait au logement-échantillon, celui où la personne avait été enquêtée, on gardait la personne dans le champ de l’exploitation. Quand cette résidence principale correspondait à une autre résidence que le logement enquêté, la personne était exclue du champ de l’exploitation.

Pour déterminer quelle était la résidence principale des personnes à la date de l’enquête, on ne se fondait pas sur la déclaration spontanée des enquêtés. Si on l’avait fait, on risquait de reproduire certaines des erreurs du recensement. Au contraire, la détermination de la résidence principale à la date de l’enquête a été effectuée a posteriori par les directions régionales de l’INSEE selon des règles précises. Lorsque la personne enquêtée avait plusieurs résidences, la règle générale était de classer en résidence principale celle où elle logeait la plus grande partie de l’année (les durées d’occupation de ses différents logements figuraient sur le questionnaire). Par contre, pour un certain nombre de cas particuliers1 - par exemple, celui d’un salarié logé par son employeur, ayant par ailleurs un domicile familial - la règle établie ne tenait pas compte de la durée d’occupation, mais elle désignait clairement le logement qui devait être classé conventionnellement en résidence principale (dans l’exemple cité, c’était le logement chez l’employeur). Autre cas particulier, celui des personnes résidant habituellement dans une communauté, mais rattachées lors de l’enquête à leur logement familial : on considérait que ce logement familial était leur résidence principale.

On aurait pu choisir de garder dans le champ de l’exploitation les personnes dont le logement enquêté était la résidence principale à la date du recensement, et non à celle de l’enquête. Cette méthode aurait été plus difficile à mettre en œuvre, c’est pourquoi elle n’a pas été choisie. En effet, il aurait fallu connaître, pour les personnes ayant quitté le logement-échantillon entre la date du recensement et celle de l’enquête, l’adresse de leur nouveau logement. Pour obtenir cette adresse, puisque les personnes n’étaient plus là pour répondre, il aurait fallu interroger les nouveaux occupants du logement enquêté, ou les voisins. Par ailleurs, du fait du départ de ses occupants, le logement-échantillon pouvait, à la date de l’enquête, être devenu vacant. Dans cette situation, l’enquêteur, souvent, n’aurait même pas su que le logement était encore récemment habité.

(1) Les cas particuliers désignés (salariés logés chez un employeur, étudiants ayant une chambre en ville pour leurs études, enfants en nourrice chez des particuliers) correspondent à des personnes susceptibles d’avoir échappé fréquemment au recensement, du fait de l’incertitude sur le lieu où elles auraient dû être recensées.
2.6. Traitement des non-réponses

Bien évidemment, dans cette enquête, il fallait essayer de limiter au maximum les non-réponses, car on pouvait supposer a priori une corrélation forte entre la non-réponse du ménage à l’enquête et un recensement incorrect de ses membres. Aussi avait-on prévu diverses formes de rattrapages pour essayer d’obtenir des renseignements auprès des ménages non répondants. À un ménage qui refusait l’enquête, les directions régionales commençaient par adresser, comme d’habitude, une lettre de rappel ; si le ménage persistait dans son refus, la direction régionale lui adressait un questionnaire postal succinct. Pour un ménage absent, la procédure de rattrapage était la suivante. L’enquêteur essayait d’obtenir auprès des voisins l’adresse d’une éventuelle autre résidence, et des tentatives d’interview par téléphone devaient être effectuées à cette 2e résidence. En réalité, ce mode de rattrapage a été très peu utilisé par les DR. Par contre, plus systématiquement, lors de la phase des retours terrain, les enquêteurs ont tenté de contacter à nouveau les ménages précédemment absents. Puis, en dernier recours, la direction régionale leur a envoyé un questionnaire postal.

Finalement, les rattrapages ont permis de récupérer l’interview de 2,7 % des résidences principales. Il subsiste néanmoins un taux d’échec (rapporté aux seules résidences principales) de 2,5 %.

Si le taux de non-réponses est faible par rapport à une autre enquête, il est élevé si on compare ce pourcentage au taux d’erreur de recensement mesuré par l’enquête. Aussi, a priori, il nous avait semblé important de tenter un redressement des non-réponses. Généralement, on redresse les non-répondants en les supposant semblables aux répondants de mêmes caractéristiques. Dans cette enquête, il nous a paru préférable de les assimiler plutôt aux ménages qui, n’ayant pas répondu lors de la phase principale de l’enquête, avaient finalement fait l’objet d’un rattrapage (par questionnaire postal, par téléphone ou lors des retours terrain). En effet, nous pensions que ces ménages non-répondants finalement répondants, difficiles à joindre lors de la phase principale de l’enquête, ou parfois récalcitrants pour répondre, avaient été sans doute recensés dans des conditions plus proches de celles des non-répondants que de celles des ménages ayant répondu à l’enquête sans difficulté. Pour effectuer ce redressement, on a donc réparti ménages non-répondants et ménages non-répondants finalement répondants en 4 classes. La 1ère était constituée des ménages dont le logement enquêté (qui était leur résidence principale à la date de l’enquête) avait été, au recensement, soit oublié, soit classé dans une autre catégorie (résidence secondaire, logement vacant ou occasionnel). Les autres ménages, dont le logement enquêté avait été recensé comme résidence principale, étaient répartis en 3 classes, selon la taille d’agglomération (variable corrélée avec le risque d’erreur au recensement). Dans chaque classe ainsi définie, on a corrigé le poids des non-répondants finalement répondants de manière à compenser les ménages défaillants.
Finalement, on s’aperçoit qu’un tel redressement ne modifie qu’à la marge le taux global d’omission au recensement et le taux de double compte : il les augmente, respectivement, de 0,1 point. Or, ce redressement est fragile, du fait que le nombre de ménages non répondants finalement répondants est faible (à peine supérieur au nombre de ménages non répondants). Par ailleurs, d’autres erreurs, en dehors de celles résultant de la non-réponse, peuvent entacher les données de l’enquête (erreurs de collecte, d’appariement et de chiffrement). En outre, une partie des omissions et des doubles comptes détectés par l’enquête ne sont que probables ou possibles, et non pas certains, comme on le verra plus loin. Pour toutes ces raisons, on a choisi de présenter ci-après les résultats détaillés bruts, c’est-à-dire non redressés.

Pour terminer cette présentation de la méthode de l’enquête, on notera que les résultats de l’enquête doivent être considérés avec précaution. L’imperfection de l’échantillon, la non-exhaustivité du champ couvert, les non-réponses entachent les résultats d’une certaine incertitude. Quand les effectifs sur lesquels portent les résultats sont faibles, les erreurs aléatoires peuvent être relativement importantes. Surtout, malgré toutes les précautions prises, on ne peut pas être sûr que certaines erreurs de recensement ne se sont pas reproduites à l’enquête (omissions de personnes, omissions de logements). Enfin, les erreurs de déclaration lors de la collecte et les erreurs de chiffrement peuvent influer sur la qualité de certains résultats.

3 - Les résultats

3.1. Les erreurs de recensement de résidences principales

On verra ci-après qu’une part importante des omissions de personnes provient de ce que leur résidence principale, là où elles auraient dû être recensées, a été oubliée au recensement, ou a été recensée comme logement vacant, résidence secondaire, ou logement utilisé occasionnellement pour des raisons professionnelles (donc sans occupants). Aussi, avant d’aborder l’étude des erreurs de dénombrement de personnes, commencerons-nous par donner quelques éléments sur les omissions ou les mauvais classements de résidences principales au recensement.

D’après l’enquête post-censitaire, en 1990, environ 1,6 % des résidences principales ont été recensées de manière incorrecte : 0,7 ont été omises ; 0,9 % ont été classées à tort dans une autre catégorie, soit 0,3 % en logements vacants et 0,6 % en résidences secondaires ou en logements occasionnels. Toutefois, ce dernier chiffre doit être considéré avec prudence ; pour certains ménages, il n’est pas facile de désigner celle de leur résidence qui est leur résidence principale, et il n’est pas toujours certain que le choix fait dans l’enquête soit meilleur que dans le recensement.
Ces pourcentages, et ceux qui suivent, sont calculés sur le champ des résidences principales, après exclusion de celles dont le ménage a déménagé en mars 1990 ou après. Si on assimile l’omission d’un ménage ayant déménagé à une erreur de recensement de résidence principale, le pourcentage d’erreurs rapporté cette fois l’ensemble des résidences principales atteint 1,9 %.

Ce n’est pas parce qu’un ménage n’est pas recensé dans sa résidence principale qu’il est obligatoirement omis, il peut avoir été recensé ailleurs. Le pourcentage de ménages omis en entier est proche de 1,5 %.

3.1.1 Les risques d’erreur diffèrent selon le type de logement, la taille de l’immeuble

Les erreurs de recensement des résidences principales sont particulièrement fréquentes pour certains types de logements marginaux (Cf Tableau 1) : les annexes indépendantes, les chambres d’hôtel, les habitations de fortune, mais le nombre de résidences principales situées dans ces logements est faible. Les logements dans des immeubles principalement à usage industriel, commercial ou administratif... ont été assez souvent oubliés (2,5 %), alors que les logements situés dans l’enceinte d’une communauté ont été mieux repérés (0,9 % d’omis). Pour les fermes, le taux d’omission est faible (0,4 %) : sauf celles qui sont très isolées, ce sont des constructions suffisamment visibles pour que l’agent recenseur ne les manque pas.

Les autres logements, les plus nombreux, sont les maisons individuelles, et les appartements (hors annexes indépendantes) situés dans des immeubles collectifs principalement ou uniquement à usage d’habitation. On remarque que les appartements sont beaucoup plus souvent classés à tort logement vacant, occasionnel ou résidence secondaire (1,5 %) que les maisons individuelles (0,5 %), ce qui n’est guère étonnant. Les ménages qui vivent en appartement sont plus difficiles à joindre qu’en pavillon, surtout quand la porte de l’immeuble est protégée par un interphone ou un code d’accès. De plus, dans les immeubles collectifs, souvent, on ne connaît pas ses voisins, et on ne peut pas donner à l’agent recenseur de renseignements fiables sur la durée d’occupation des autres appartements de l’immeuble. Par contre, les appartements ne sont qu’un peu plus souvent omis que les pavillons (0,8 % contre 0,5 %). Ce résultat doit cependant être nuancé en fonction de la taille de l’immeuble (Cf Tableau 2). Le taux d’omission des appartements est relativement élevé dans les petits immeubles (2,2 % dans les immeubles de 2 à 5 logements) ; dans les moyens et gros immeubles, il est nettement plus faible (il tourne autour de 0,4 %). Cette différence s’explique par deux facteurs. Tout d’abord, les petits immeubles sont parfois oubliés en entier. Mais ces omissions

(1) Logements rattachés à un logement principal mais séparés de lui, tels que, par exemple, "chambres de bonne".
n’expliquent qu’une partie de l’écart entre les taux. Un deuxième phénomène joue sans doute : dans les immeubles d’une certaine taille, la disposition des appartements est la même à chaque étage, ceux-ci sont donc assez faciles à repérer, alors que les petits immeubles ont souvent une structure plus complexe, moins apparente.

On peut répartir les omissions de résidences principales en fonction de la nature de l’erreur : oubli de recensement d’un immeuble entier, omission d’un appartement dans un immeuble recensé. Ainsi, sur 100 logements omis, 36 sont des maisons individuelles ; 44 sont des appartements oubliés dans un immeuble d’habitation recensé (dont 5 annexes indépendantes) ; 8 correspondent à un immeuble oublié en entier (en général un petit immeuble de 2 à 5 logements). Les cas restants (12) sont des logements d’un type particulier : chambre d’hôtel, habitation de fortune, logement dans un immeuble industriel, etc.

L’enquête donne quelques éléments partiels sur les raisons, appréciées par les enquêteurs, de l’omission de ces résidences principales : 15 % seulement des omissions viendraient de ce que l’accès du logement est peu visible ; 27 % de ce que le travail de l’agent recenseur est de mauvaise qualité sur le district qui lui a été confié ; mais 34 % des omissions n’ont pas de raison apparente.

3.1.2 Deux fois plus d’erreurs pour les logements occupés par une seule personne

Le taux d’omission des résidences principales est relativement élevé pour les personnes vivant seules : 1,4 % (Cf Tableau 3) ; il diminue ensuite nettement, quand augmente le nombre de personnes du ménage, puisqu’il n’est plus que de 0,6 % pour les ménages de 2 personnes, puis de 0,3 % pour les ménages de 3 ou 4 personnes ; puis il remonte un peu pour les ménages d’une taille supérieure. Les proportions de résidences principales qui, au recensement, ont été classées dans une autre catégorie (logement vacant, occasionnel, résidence secondaire) suivent une courbe analogue : pour les ménages d’une personne, le pourcentage atteint 2,0 % ; il diminue ensuite nettement avec l’augmentation de la taille du ménage, pour remonter ensuite très légèrement quand cette taille dépasse 5.

Il n’est pas étonnant que les résidences principales des ménages de très petite taille soient assez souvent omises ou mal classées. Ce sont plus souvent de petits logements, ou même des logements marginaux, tels que annexes indépendantes, qui sont donc plus difficiles à repérer. Par ailleurs, les couples sans enfants, et surtout les personnes vivant seules, sont plus souvent absents de leur domicile que les couples ayant des charges familiales. Enfin, plus un ménage comporte de membres, plus l’agent recenseur a de chances d’en contacter au moins un. Par contre, on s’explique assez mal que le taux d’erreur remonte quand la taille du
ménage dépasse 4 personnes. Il se peut que cette remontée ne soit due qu'à des erreurs aléatoires.

3.1.3 Le degré d’urbanisation influe sur la qualité du recensement...

Les omissions et les erreurs de classement des résidences principales ont été nettement plus fréquentes dans les grosses agglomérations (de plus de 200 000 habitants) que dans les unités urbaines plus petites ou dans les communes rurales (Cf Tableau 4). Que les résidences principales soient plus difficiles à recenser en ville s’explique par la plus grande mobilité de la population, l’absence d’intégration sociale, la densité et la structure de l’habitat. Pour cette même raison, et, parce que dans les campagnes, souvent, l’agent recenseur connaissait bien les logements et les ménages de sa commune (ce pouvait être le secrétaire de mairie), on aurait dû observer une meilleure qualité du recensement dans les communes rurales que dans les agglomérations moyennes. Si cette différence n’apparaît pas, c’est que le contrôle dit TH (pour Taxe d’Habitation), que nous allons décrire maintenant, a brouillé ce phénomène de différenciation en fonction de la taille de l’agglomération.

3.1.4 ... mais les omissions ont été réduites dans les zones urbaines où le contrôle dit TH a été mis en œuvre

En effet, pour améliorer le recensement des logements, réduire les omissions et les classements erronés, un contrôle TH a été mis en place pendant le recensement de 1990. Il s’agissait de vérifier que, sur chaque district, le nombre de logements recensés, classés par catégorie, était à peu près conforme au nombre figurant sur les listings issus du fichier de la taxe d’habitation. En cas d’incohérences trop fortes, les agents recenseurs devaient retourner sur le terrain pour corriger le recensement. Ces contrôles devaient être réalisés principalement sur les communes de plus de 20 000 habitants. Dans l’enquête post-censitaire, pour chaque aire enquêtée, on sait, par un chiffrement, si le contrôle TH a été effectué pour le(s) district(s) de recensement sur le(s)quel(s) était située l’aire. L’enquête doit donc pouvoir indiquer si ce contrôle a eu une certaine efficacité. Le Tableau 5 fait apparaître qu’en effet, à taille d’agglomération égale, le taux d’omission des résidences principales est plus faible sur les zones soumises au contrôle TH que sur les autres. Certes, ce phénomène n’est pas perceptible sur l’ensemble de l’agglomération parisienne, mais c’est en raison d’un effet de structure : les zones

(1) On pourrait supposer que cette évolution provient de ce que les ménages nombreux sont plus souvent que les autres des ménages d’étrangers. En effet, on vera plus loin que les étrangers sont un peu plus souvent oubliés au recensement que les français. Mais cette hypothèse ne peut être retenue, car, si on elimine les étrangers du tableau, les taux d’erreur augmentent encore légèrement quand la taille du ménage atteint 5 personnes.
de l'agglomération non couvertes par le contrôle TH sont principalement constituées de petites communes (de moins de 20 000 habitants), où le recensement des logements est plus facile, et le taux d'omission faible. Si on élimine de l'agglomération parisienne les communes de moins de 20 000 habitants, le taux d'omission sur les zones contrôlées devient manifestement plus faible que sur celles non couvertes par le contrôle

Si le contrôle TH a visiblement permis d'éviter des omissions de logements, par contre, l'enquête ne met pas en évidence qu'il a empêché des classements erronés de résidences principales dans une autre catégorie de logement (résidence secondaire...). Mais, sur ce point, on l'a déjà noté, les résultats de l'enquête sont un peu fragiles.

À la lumière de ces résultats, on peut revenir à l'analyse du Tableau 4. Il est clair qu'en l'absence des contrôles TH, on aurait eu une progression plus continue des taux d'omission en fonction de la taille de l'agglomération. Ces taux auraient été plus élevés sur les grosses agglomérations et les unités urbaines de taille intermédiaire, et ils auraient dépassé ceux des petites unités urbaines. Au contraire, on note que, quand la taille de l'agglomération atteint 20 000 habitants, le taux d'omission baisse brutalement de 0,6 % à 0,2 %. Si on se reporte à nouveau Tableau 5, cette rupture a, pour origine évidente que les contrôles TH, assez rares pour les unités urbaines de moins de 20 000 habitants, ont été par contre très fréquents pour celles d'une taille supérieure. De plus, il est remarquable que, pour les unités urbaines de 20 000 à 100 000 habitants, les contrôles TH ont permis d'obtenir, sur les zones couvertes, un taux d'omission très bas (0,1 %).

On peut aussi observer des différences régionales dans la qualité du recensement (Cf Tableau 6). Certes, ces différences ne sont pas complètement indépendantes des écarts liés au degré d'urbanisation, mais il semble qu'existent aussi des facteurs de disparité purement régionaux.

3.2. Les erreurs de recensement des personnes

3.2.1 Résultats globaux

Tout d'abord, l'enquête donne des résultats globaux sur l'erreur de couverture de la population par le recensement : omissions, doubles comptes. Par contre les ajouts de

---

(1) Cet effet de structure fait qu'en observant les écarts des taux d'omission par taille d'agglomération, toutes catégories de commune confondues, on sous-estime un peu la mesure de l'efficacité du contrôle TH. Pour éliminer cet effet de structure, on aurait pu, pour chaque catégorie d'agglomération, restreindre le champ du tableau 5 aux seules communes de plus de 20 000 habitants. Mais cette restriction n'était pas possible pour les unités urbaines de 20 000 habitants à 200 000 habitants, car, pour ces unités urbaines, très peu de communes de plus de 20 000 habitants n'ont pas été contrôlées, et les résultats auraient été peu significatifs.
personnes comptées à tort (personnes fictives ajoutées par fraude, enfants nés après la date de recensement) sont mal mesurés

Dans l’enquête, on peut classer les *omissions*, en trois catégories, selon leur plus ou moins grand degré de certitude :

- les *omissions caractérisées* correspondent à des oubliis de recensement repérés sans ambiguïté par l’enquête ;

- les *omissions probables ou possibles* correspondent à des enquêtés non recensés dans le logement enquêté (qui était leur résidence principale à la date de l’enquête) et ayant une deuxième résidence qui n’a pas pu être appariée avec le recensement.

Dans le cas des *omissions probables*, l’enquêté a donné l’adresse précise de son autre résidence, mais cette adresse, n’ayant pas été trouvée sur le Minitel, est de ce fait un peu incertaine. Aussi, quand, pour cette deuxième résidence, on n’a pas pu trouver de document de recensement correspondant, on peut en conclure qu’il est fort probable que l’enquêté n’y a pas été recensé ; mais ce n’est pas certain, car, si l’adresse de la deuxième résidence donnée lors de l’enquête est erronée, on n’a pas pu vérifier que l’enquêté n’avait pas été recensé à la bonne adresse.

Plus de la moitié des personnes *probablement omises* sont des enquêtés ayant déménagé en mars 1990 ou après : leur deuxième résidence, non retrouvée dans les documents du recensement, est celle qu’ils occupaient avant leur déménagement ; ils y résidaient donc encore fin février ou début mars. Cette forte proportion de personnes ayant déménagé parmi les omissions probables n’est pas étonnante. En effet, en cas de déménagement d’un ménage entier, quand on ne trouvait pas pour lui de documents de recensement à l’adresse de sa résidence antérieure, généralement on n’a pas pu conclure à son omission certaine dans cette résidence, mais seulement à son omission probable, puisque son ancienne adresse ne figurait plus sur le Minitel à la date, relativement tardive, du 2ᵉ appariement. Pour s’assurer que l’adresse fournie par le ménage enquêté était bonne, il aurait fallu consulter les annuaires téléphoniques, qui, eux, ne sont pas mis à jour instantanément.

Dans le cas des *omissions possibles*, la deuxième résidence n’a pas pu être appariée parce qu’on n’en connaissait pas l’adresse. On sait que la personne a été omise dans le logement enquêté (qui était sa résidence principale à la date de l’enquête), mais on ne sait pas si elle n’a pas été recensée ailleurs.

Là encore, plus de la moitié des *omissions possibles* correspondent à des enquêtés ayant changé de domicile en mars 1990 ou après. Il est possible que ces enquêtés aient été recensés dans leur ancienne résidence, celle qu’ils occupaient début mars, mais on n’a pas pu trancher.
Parmi les personnes *recensées une fois*, on isole :

- les personnes dont on sait *sans ambiguïté* qu’elles ont été *recensées une fois* ;

- les personnes *probablement ou peut-être recensées une fois* : cette catégorie est constituée de personnes recensées au moins une fois, mais ayant déclaré une autre résidence qui n’a pas pu être appariée avec le recensement (ou bien l’adresse de cette deuxième résidence était manquante, incomplète ou fausse ; ou l’adresse était précise, mais elle n’a pas été trouvée au Minitel, et on ne sait pas si le non-appariement avec le recensement n’est pas due à une erreur d’adresse). Une petite partie de ces personnes "probablement ou peut-être recensées une fois" ont sans doute été en réalité comptées deux fois au recensement.

Pour les *doubles comptes*, on isole deux catégories :

- les *doubles comptes caractérisés* : on a retrouvé pour les personnes deux bulletins individuels remplis, chacun dans deux logements différents ;

- les *doubles comptes probables*. La majorité (70 %) d’entre eux correspondent à des élèves internes ou des militaires pour lesquels il semble, sans qu’on en soit toujours certain, que leur bulletin individuel n° 5 rempli dans la caserne ou l’internat n’a pas pu être réintégré dans leur logement familial et a été pris en compte, parmi les *réintégrables non réintégrés*, en sus du bulletin individuel n° 2 rempli à leur nom dans leur logement familial.

D’après l’enquête, les erreurs au recensement se répartissent comme suit (*Cf Tableau 7*) :

<table>
<thead>
<tr>
<th>Catégorie</th>
<th>% de la population enquêtée</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omissions caractérisées</td>
<td>1,5</td>
</tr>
<tr>
<td>Omissions probables</td>
<td>0,3</td>
</tr>
<tr>
<td>Omissions possibles</td>
<td>0,2</td>
</tr>
</tbody>
</table>

Le pourcentage des doubles comptes caractérisés est de 0,6 %. Si on y ajoute celui des doubles comptes probables (0,1 %), le pourcentage des doubles comptes atteint 0,7 %. Mais, si une partie des *recensés probablement ou peut-être une fois* l’a été en réalité deux fois, il est possible que le taux de double compte soit supérieur à 0,7 % ;

Finalement, le solde, la part des personnes recensées (à coup sûr, probablement ou peut-être) une seule fois est de 97,3 % ; mais ce taux inclut une frange d’incertitude, constituée des personnes "probablement ou peut-être recensées une fois", qui représentent 3,4 % de la population enquêtée. Comme on vient de le rappeler, une petite partie de cette marge a sans doute été recensée deux fois. En conséquence, le pourcentage total des personnes "recensées une fois" est sans doute un peu inférieur.

Ces résultats n’ont pas été redressés des non-réponses. Si l’on effectuait un tel redressement, les taux d’erreur pourraient être légèrement majorés, de 0,1 point pour le taux d’omission et de 0,1 point, également, pour le taux de double compte. Le taux
d’omission (regroupant omissions caractérisées et probables) serait alors de 1,9 %, le taux de double compte (regroupant doubles comptes caractérisés et probables) serait de 0,8 %.

Rappelons que sont exclues du champ de l’enquête les personnes vivant en communauté et non rattachées à un ménage "ordinaire", les personnes vivant dans des habitations mobiles, les personnes sans domicile.

3.2.2 Sources des erreurs de recensement des personnes

Remarque préalable concernant la présentation des résultats

Dans la suite de l’étude, et sauf précision contraire, les taux d’omission mentionnés sont ceux qui regroupent omissions caractérisées et probables. Les taux de double compte concernent l’ensemble des doubles comptes caractérisés et probables.

Dans les tableaux qui accompagnent le texte, on a généralement distingué omissions caractérisées et omissions probables ; par contre, on a regroupé tous les doubles comptes. Dans quelques cas particuliers, quand la marge des omissions possibles était élevée, on l’a fait figurer ; de même, quand les doubles comptes probables étaient relativement nombreux, on les a distingués des doubles comptes caractérisés. Dans les tableaux, les effectifs sont donnés, principalement pour donner une idée de la validité des taux d’omission et de double compte. Quand les effectifs sont petits, ces taux sont évidemment fragiles. Généralement, ces effectifs sont sous-estimés, puisqu’ils ne sont pas redressés des non-réponses et que, très vraisemblablement, une partie de la population résidant dans des immeubles construits depuis la date de tirage de l’échantillon (1982) n’est pas représentée dans l’enquête.

Des situations à risque : le déménagement pendant la période de recensement, le rattachement à plusieurs résidences...

L’enquête de 1990 donne des éléments sur la nature et les causes des erreurs de dénombrement des personnes.

L’enquête permet d’abord de repérer que certaines situations occasionnent un risque relativement élevé d’erreurs de recensement : le déménagement pendant la période de recensement ; le rattachement des personnes à plusieurs résidences (communautés incluses).

Les personnes qui ont changé de résidence principale pendant la période de recensement risquent d’être omises dans leurs deux logements successifs, ou d’être comptées deux...
fois, une fois dans chacun d’eux. De fait, le taux d’omission des personnes ayant déménagé en mars 1990 ou après (entre la date du recensement et celle de l’enquête) est élevé : 14,4 %, peut-être supérieur si on y adjoint une partie des omissions possibles qui représentent un taux de 5,6 % (Cf Tableau 8). Le taux de double compte est, lui aussi, relativement important (3,9 %).

Même si elles n’ont pas changé de logement, les personnes qui ont un point d’attache dans plusieurs logements ont un risque assez grand d’être oubliées au recensement (2,5 % d’entre elles le sont) ou d’être comptées en double (3,6 %). Il s’agit de personnes disposant d’une résidence secondaire, d’un logement occasionnel, mais aussi d’étudiants ayant un logement en ville en même temps qu’une résidence familiale, d’enfants de parents divorcés ayant deux résidences familiales...

Les personnes rattachées à un ménage ordinaire et vivant par ailleurs dans une communauté à la date de l’enquête, ou y ayant vécu entre le 1er janvier 1990 et la date de l’enquête forment un autre cas particulier de "doubles résidences". Leurs taux d’omission et de double compte sont particulièrement élevés, et diffèrent assez nettement selon la catégorie de communauté (Cf tableau 9). Ainsi, d’après l’enquête, les résidents de cité universitaire sont très souvent mal recensés (près d’un cinquième sont omis, un cinquième comptés deux fois). Les élèves internes paraissent être mieux recensés. Certes, parmi eux, la part des doubles comptes reste assez élevée, surtout si on y inclut les "doubles comptes probables", mais celle des omis (2,4 %) est relativement faible, si on la compare à celle des autres catégories de population résidant en communauté. Les militaires se caractérisent par un taux d’omission élevé, et un taux de double compte plutôt incertain. Les résultats relatifs aux autres catégories de personnes rattachées à une communauté ont un moindre intérêt, puisqu’ils ne portent que sur une faible partie de leurs membres (ceux qui ont conservé des liens avec un ménage ordinaire).

Il n’est pas étonnant de constater que les taux de doubles comptes des élèves internes et des militaires du contingent sont nettement inférieurs à ceux des étudiants en cité universitaire, puisque le recensement des premiers était organisé de manière à éviter les doubles comptes (par la réintégration du bulletin individuel rempli dans la caserne ou l’internat dans le logement familial).

Les résultats concernant ces différentes catégories de personnes résidant en communauté sont certes un peu fragiles. Les effectifs sont assez faibles ; les doubles comptes sont peut-être sous-estimés du fait que retrouver les communautés dans les documents du recensement s’est avéré n’être pas toujours facile. Par ailleurs, la liste des membres de certaines communautés était parfois fort longue, et il se peut que certains enquêtés, recensés, n’y aient pas été repérés. Par contrec, il est satisfaisant de constater que les effectifs de militaires, d’élèves internes et d’étudiants en cité universitaire représentés dans l’enquête sont assez proches de ceux du recensement (Cf Tableau 10).
Précisons par ailleurs que les omissions de personnes rattachées à une communauté ne résultent sans doute pas toutes d’un échec de recensement à l’adresse de la communauté. En effet, les résultats présentés portent sur l’ensemble des personnes qui résidaient en communauté à la date de l’enquête, ou qui y ont vécu entre le 1er janvier 1990 et la date de l’enquête. Par conséquent, une petite fraction d’entre elles n’y vivaient pas encore à la date du recensement de la communauté (mi-février pour les militaires, fin janvier pour les autres). Ce décalage de temps n’a sans doute pas beaucoup d’importance pour les élèves internes, qui généralement passent toute leur année scolaire dans le même internat. Il ne devrait pas être non plus très gênant pour les étudiants de cité universitaire. Par contre, pour les militaires du contingent, les incorporations s’échelonnent sur l’année tous les deux mois. Après celle de début février, une vague de recrutement a eu lieu début avril, une autre début juin. Les militaires représentés dans le tableau 9 incluent donc une grande partie des nouveaux entrants d’avril, et peut-être quelques entrants de juin.

Toutefois, le fort taux d’omission des militaires résulte peut-être justement en partie de l’échelonnement des recrutements du contingent. En effet, le recensement des logements s’est souvent prolongé au-delà de mars, sur les mois d’avril et de mai. De ce fait, certains militaires incorporés en avril peuvent avoir échappé au recensement, si, à la date du recensement de leur résidence familiale, ils l’avaient déjà quittée pour la caserne.

On vient de voir que le déménagement pendant la période de recensement, et plus généralement le rattachement des personnes à plusieurs résidences, étaient à l’origine d’erreurs de recensement fréquentes. Les personnes, qui ne disposent que d’un seul logement, sont mieux recensées. Leur taux de double compte est nul. Mais leur taux d’omission atteint tout de même 1,2 %.

... Toutefois près de 60 % des personnes omises n’avaient qu’un seul logement

On peut regrouper ces différents résultats dans une autre perspective, c’est-à-dire essayer de faire une typologie des omissions et des doubles comptes, montrant comment se répartissent les erreurs en fonction de leur origine probable.

Le Tableau 11 montre que, sur 100 personnes omises :

- 9 sont rattachées à une communauté ;
- 17 ont déménagé en mars 1990 ou après (ils n’ont été recensés, ni dans leur nouveau logement, ni dans l’ancien) ;
- les autres (74) auraient dû être recensées dans le logement enquêté, qui était leur résidence principale à la date de l’enquête et à la date du recensement. Or, pour 25
personnes, le logement enquêté a été omis au recensement ; pour 24 autres, celui-ci a été classé vacant, occasionnel ou résidence secondaire. Le solde est constitué de 25 personnes, omises isolément dans le logement enquêté (celui-ci a été recensé comme résidence principale).

Autre résultat intéressant : parmi les 74 personnes omises qui n’ont pas déménagé entre la date du recensement et la date de l’enquête, et qui auraient du être recensées dans leur résidence principale, 57 ne disposaient que d’un seul logement ; 17 ont été oubliées dans leur résidence principale et dans leur(s) autre(s) logement(s).

On observe que les erreurs de recensement des résidences principales sont à l’origine d’une part élevée des omissions de personnes. Pour calculer leur importance relative par rapport à l’ensemble des omissions, il faudrait ajouter aux cas mentionnés précédemment, et qui représentent déjà 49 % de cet ensemble, les omissions de ménages ayant déménagé en mars 1990 ou après, et qui ont échappé en totalité au recensement. On peut en effet considérer que ces ménages ont été omis en totalité du fait d’une erreur de recensement de leur résidence principale. D’après l’enquête, 108 000 personnes omises (11 % des omissions) appartiennent à un ménage non recensé dans sa totalité. Au total, le pourcentage des omissions qui résulte d’une erreur de recensement de la résidence principale est d’environ 60 %. Cela signifie que 60 % des omissions de personnes correspondent à des oubliés de recensement de ménages entiers, et que 40 % des personnes omises l’ont été isolément dans des logements recensés.

On ne s’explique pas toujours très bien pourquoi des personnes sont omises isolément, quand leur résidence principale a été recensée avec d’autres membres de leur ménage. Il se peut que certaines d’entre elles aient souhaité délibérément échapper au recensement. Mais l’enquête ne repère pas ce type de raison, et elle permet sans doute de ne récupérer qu’une petite partie de cette catégorie d’omissions. On peut en tout cas repérer dans l’enquête que les personnes qui sont éloignées du “noyau familial” ont une probabilité plus forte d’être omises que celles qui appartiennent à ce noyau. Ainsi (Cf Tableau 12), les personnes de référence, les conjoints et les enfants sont mieux recensés que les salariés logés par leur employeur, les amis, les pensionnaires, sous-locataires... Il est clair que la personne du ménage qui répond au recensement oublie assez souvent de mentionner les personnes qui ne font pas du noyau stable de sa famille. Ces mêmes catégories particulières ont d’ailleurs assez fréquemment une autre résidence. En conséquence, leur taux de double compte est élevé.

Puisque une forte proportion des omissions de personnes ont pour origine une erreur de recensement de leur résidence principale, il n’est pas étonnant de constater que la taille du ménage (de même que, on le verra après, la taille de l’agglomération) influe manifestement sur la probabilité qu’ont les individus d’être oubliés au recensement, puisque la qualité de recensement des logements est corrélée avec ces deux critères. Ainsi, 3,2 % des personnes vivant seules sont omises (Cf Tableau 13) ; ce pourcentage baisse ensuite avec la taille du ménage pour remonter à partir de 5 personnes. On peut
aussi faire remarquer que le taux de double compte des personnes seules (1,4 %) est nettement supérieur aux autres, et deux fois plus élevé que le taux global.

Enfin, on observe que le taux d’omission des personnes s’accroît régulièrement avec la taille de l’agglomération (Cf Tableau 14). Il n’est que de 0,9 % dans les communes rurales appartenant à un canton entièrement rural et progresse jusqu’à 2,9 % pour l’agglomération parisienne. Par contre, le risque de double compte ne semble pas différer selon que les personnes ont leur résidence principale en ville ou dans une commune rurale.

L’effet du contrôle TH est moins visible dans ce tableau que dans celui portant sur les omissions de résidences principales. Mais, bien évidemment, ces contrôles, en réduisant les omissions de logements ont diminué les oubli de recensement de personnes. En l’absence de contrôles TH, la hiérarchie des taux d’omission par taille d’agglomération aurait été plus accentuée.

3.3. Résultats complémentaires

L’enquête post-censitaire de 1990 permet aussi d’indiquer si les erreurs de dénombrement touchent plus particulièrement certaines catégories de population.

Le taux d’omission est un peu plus élevé pour la tranche d’âge des 20-29 ans (Cf Tableau 15). Le taux d’omission diminue ensuite légèrement en même temps qu’augmente l’âge, pour remonter un peu au-delà de 70 ans. Les nouveaux-nés, ceux nés en 1990 avant la date du recensement sont assez souvent omis, mais ils sont peu nombreux, et leur taux d’omission pèse peu sur l’ensemble. Le taux global d’omission est un peu plus important pour les hommes que pour les femmes. Cet écart entre hommes et femmes se vérifie à presque tous les âges. Quanti aux doubles comptes, ils sont fréquents aux âges jeunes, surtout entre 20 et 24 ans, plus rares pour les très jeunes (moins de 10 ans) ou les personnes d’âge mûr (après 35 ans). Cet effet d’âge n’est en réalité que la traduction de différences dans les modes de vie. Quand on est jeune adulte, que les liens avec le domicile familial commencent à se distendre et qu’on n’est pas encore installé dans un logement et une situation professionnelle stable, la mobilité est plus forte, le rattachement à plusieurs résidences plus fréquent, et donc, les risques d’omissions et de doubles comptes au recensement accru.

D’après l’enquête, les étrangers ont été environ deux fois plus souvent oubliés au recensement que les français (Cf Tableau 16) ; parmi les étrangers, ceux qui proviennent du Maghreb sont un peu plus fréquemment omis que ceux qui proviennent d’un pays de la CEE. Mais ces résultats sont partiels, puisque l’enquête ne couvre que la population des ménages ordinaires. Les étrangers qui vivent dans des communautés,
dans des foyers de travailleurs par exemple, et ne sont pas rattachés à un logement, sont exclus du champ de l’enquête.

Les taux d’erreur varient en fonction de la situation professionnelle (Cf Tableau 17) : pour les actifs ayant un emploi, les taux d’omission et de double compte sont identiques aux taux moyens. À l’opposé, les inactifs de 60 ans et plus, et les retraités sont un peu moins souvent oubliés qu’en moyenne (le tableau par âge montre que c’est surtout vrai des jeunes retraités). Par contre, les chômeurs échappent fréquemment au recensement (4,2 %). Parmi les étudiants âgés de plus de 18 ans, on compte beaucoup d’omis (5 %), et de doubles comptes (3,6 %). En réalité, toutes ces disparités en fonction de la situation professionnelle s’analysent plutôt, comme les différences par âge, par référence à la plus ou moins grande stabilisation des personnes dans un logement fixe. Les chômeurs, qui n’ont pas un travail assuré, sont sans doute plus mobiles que les actifs occupés ; et ils sont en moyenne plus jeunes. Les étudiants se caractérisent par un rattachement fréquent à plusieurs logements, la résidence de leurs parents, et un lieu d’habitation qu’ils occupent pour leurs études : chambre en ville, internat, cité universitaire (Cf Tableau 18). De plus, ceux qui vivent dans un logement indépendant en changent sans doute assez souvent. La proportion des étudiants qui ont déclaré à l’enquête plus d’une résidence atteint 41 %. Il n’est donc pas étonnant qu’une assez forte proportion d’entre eux aient été omis ou comptés 2 fois au recensement.

Le Tableau 17 donne aussi une idée des différences des taux d’erreurs en fonction de la catégorie socio professionnelle des actifs occupés. Elles sont peu marquées. On peut tout de même noter que les agriculteurs sont très peu souvent oubliés (0,7 %), à l’encontre de la catégorie des "artisans, commerçants, chefs d’entreprises" qui, elle, l’est un peu plus souvent qu’en moyenne (2,3 %). Une analyse plus fine montre d’ailleurs que ce sont les commerçants et les chefs d’entreprises qui sont assez souvent omis, et non les artisans. L’explication en est peut-être que les premiers ont plus de déplacements professionnels. À l’opposé, les agriculteurs travaillent généralement sur leur lieu d’habitation, la ferme, domicile qui, on l’a vu, est assez peu souvent oublié au recensement.

(1) On remarquera que, dans l’enquête, l’effectif des étrangers est relativement élevé, puisqu’il est un peu supérieur à l’ensemble des étrangers recensés (3 580 000). Il est fort probable que certains enquêteurs aient affecté d’office, sans même poser la question, la même nationalité aux enfants d’étrangers qu’à leurs parents, alors que ces enfants peuvent être français.
4 - La qualité du recensement de 1990 : un essai de bilan global

Pendant le recensement de 1990, un soin particulier a été porté à la collecte : les délégués de l’INSEE, recrutés en plus grand nombre que pour les recensements précédents, ont dispensé une formation approfondie aux agents recenseurs et réalisé un suivi attentif de leur travail. De même, ils ont bien encadré le travail du personnel des mairies chargé du recensement. Des contrôles de l’exhaustivité de la collecte des logements ont été effectués, soit par rapprochement avec les résultats du recensement de 1982, soit - on l’a déjà noté - par rapprochement avec le fichier de la Taxe d’Habitation.

Des indicateurs partiels suggèrent d’ailleurs que le recensement de 1990 est meilleur que celui de 1982. Mais, on n’a pas de moyen de donner une mesure globale du gain de qualité obtenu entre ces deux derniers recensements.

Le seul point de comparaison dont on dispose est le recensement de 1962. Le Tableau 7 fait apparaître des résultats voisins aux deux dates. Compte tenu des efforts consentis en 1990, il peut paraître décevant de n’avoir pu obtenir un niveau des taux d’erreur plus bas qu’en 1962. Mais, il faut rappeler que le recensement de 1962, comme celui de 1990, était considéré, par ceux qui y avaient participé, comme étant de bonne qualité. De plus, l’apparente égalité des résultats est peut-être trompeuse. En effet, on peut considérer que, depuis 1962, les difficultés de recensement se sont sans doute aggravées : en raison de l’urbanisation et de l’évolution des conditions de vie des français : ils ont plus souvent une résidence secondaire ; les femmes travaillent plus souvent à l’extérieur ; les étudiants sont plus nombreux... Certes, à l’inverse, dans les villes, il y a plus d’immeubles neufs qui sont plus faciles à recenser que les anciens, mais, en raison de la croissance du sentiment d’insécurité, les ménages citadins sont plus difficiles à joindre. Aussi, pour un travail des agents recenseurs de qualité égale, on aurait dû avoir une augmentation sensible du taux d’erreur. Il n’en est rien, ce qui est un résultat plutôt satisfaisant.
Méthodologie de l’enquête de 1962.

**Principales différences avec celle de 1990**

Si, dans ses grandes lignes, l’enquête de 1990 est assez semblable à celle de 1962, on peut souligner deux différences principales :

1) En 1962, le *premier appariement* (à l’adresse des logements enquêtés) a été fait par les enquêteurs (et non par le personnel des directions régionales). Ceux-ci partaient en effet sur le terrain avec les documents du recensement. Ils devaient faire l’enquête sans les consulter, et, dès la fin de l’interview, faire sur place l’appariement, et obtenir des explications sur les divergences. Cette méthode a l’avantage de faciliter le travail de 1er appariement et augmente sans doute sa qualité. Par contre, la garantie d’indépendance entre l’enquête et le recensement est moins forte.

2) En 1962, tous les logements sans exception, résidences secondaires incluses, devaient faire l’objet d’un questionnaire complet dans lequel les enquêteurs avaient à inscrire la liste des occupants et l’adresse de leurs autres résidences (on suppose que, lorsque le ménage était absent, l’enquêteur interrogeait les voisins ou la mairie ?). Les ménages disposant d’une résidence secondaire étaient donc représentés deux fois dans l’enquête. Pour corriger ce biais, au stade de l’exploitation, leur coefficient de pondération était divisé par deux. Ce même traitement s’appliquait aussi aux personnes qui, n’ayant pas de résidence personnelle, effectuaient des séjours temporaires chez des membres de leur famille, des amis...

Par contre, les cas particuliers de double résidence, tels que personne logeant à la fois dans une communauté et dans un logement, étudiant vivant dans sa famille et dans une chambre en ville, salarié logé par son employeur ont été traités comme dans l’enquête de 1990.

Nous n’avons pas voulu retenir cette méthode pour l’enquête de 1990. En effet, pour pouvoir l’appliquer correctement, il aurait fallu qu’on puisse connaître avec une bonne fiabilité les noms des occupants des résidences secondaires de l’échantillon, l’adresse de leur résidence principale également, et de toutes leurs autres résidences, puisque, pour ces personnes, au même titre que pour celles enquêtées dans leur résidence principale, il aurait fallu vérifier si elles avaient été comptées une, deux ou zéro fois au recensement. Or, les personnes disposant de résidences secondaires ne les occupent généralement qu’épisodiquement, et on avait toute chance de ne pas pouvoir en joindre...
un grand nombre pendant la durée de l’enquête. De plus, souvent, il n’aurait pas été possible d’obtenir des voisins les renseignements indispensables, notamment dans les zones à forte densité de résidences secondaires. En 1962, à une date où les résidences secondaires étaient moins nombreuses, et sans doute plus dispersées, le problème ne se posait peut-être pas dans les mêmes termes.

3) En 1962, le champ de l’enquête était moins large qu’en 1990. En 1962, l’enquêteur devait recueillir la liste la plus exhaustive possible des occupants du logement à la date de l’enquête, y compris les absents temporairement, y compris les occupants temporaires. Puis, il pouvait consulter immédiatement les documents du recensement pour vérifier si d’autres personnes n’avaient pas été recensées dans le logement. Si tel était le cas, elles étaient ajoutées sur le questionnaire de l’enquête, et l’enquêteur devait mentionner la raison pour laquelle elles n’habitaient plus le logement. Dans un dernier temps, une question de rattrapage était posée : "En outre, considérez-vous que certaines personnes (membres de la famille, salariés logés ou autres) auraient pu être recensées ici, bien qu’elles ne l’aient pas été ?".

Par contre, en 1990, l’enquêteur n’avait pas avec lui les documents du recensement. Aussi, il lui était demandé de faire la liste de tous les occupants du logement depuis le 1er janvier. Les personnes ayant quitté le logement ou y ayant emménagé depuis le recensement pouvaient donc être repérées dès la collecte, ce qui facilitait l’appariement que les directions régionales avaient à effectuer, une fois la phase principale de collecte terminée.

On peut penser que la méthode de 1990 n’est pas sans inconvénient. On pensait au départ qu’il n’était pas gênant d’enquêter un champ trop large de personnes si, au stade de l’exploitation, on veillait à éliminer celles qui auraient été représentées plusieurs fois dans l’enquête, c’est-à-dire celles dont le logement enquêté n’était pas la résidence principale (voir chapitre 2.4 : Exploitation). Or, cette délimitation du champ de l’exploitation reposait sur un chiffrement qui, malheureusement, n’a pas toujours été de bonne qualité. Les erreurs de chiffrement ont pour beaucoup pu être corrigées par informatique, mais sans doute pas toutes. Or, elles pourraient conduire à une petite sur-estimation des omissions, dans le cas où, suite à un chiffrement erroné, on aurait retenu dans le champ de l’exploitation des personnes hébergées occasionnellement dans le logement enquêté, qui, de ce fait, n’y ont pas été recensées, et dont on n’aurait pu obtenir, lors de la collecte, l’adresse de la résidence principale. En conséquence, des contrôles minutieux ont été effectués pour minimiser cette catégorie d’erreurs.

Par ailleurs, du fait qu’on enquêtait des personnes dont le logement enquêté n’était pas la résidence principale à la date de l’enquête, ni à la date du recensement, il était normal qu’on ne trouve pas, pour elles, de bulletin individuel à l’adresse du logement enquêté lors du 1er appariement. Or, que ces cas normaux de non-appariement puissent être relativement fréquents compliquait un peu le travail des agents et risquait de les rendre un peu moins attentives au traitement des cas anormaux, nécessitant un retour terrain.
4) Le questionnaire de 1990 a été conçu de manière à effectuer un repérage plus exhaustif des autres résidences qu’en 1962. En théorie, le concept d’autre résidence était aussi large dans les deux enquêtes, puisque la question posée en 1962 était celle-ci : "Cette personne loge-t-elle parfois AILLEURS (Pied à terre, famille, amis, maison louée, etc.) à une époque quelconque de l’année". Mais, cette question était unique, et une ligne seulement du questionnaire était prévue pour noter la catégorie de cette autre résidence et son adresse. En 1990, au contraire, plusieurs questions étaient posées pour qu’aucune catégorie d’autre résidence ne soit omise, et il était explicitement prévu que chaque enquêté puisse avoir plusieurs autres résidences, puisqu’on pouvait remplir pour lui autant de sous-questionnaires autre résidence que nécessaire.

Le pourcentage de la population ayant déclaré à l’enquête plus d’une résidence a fortement augmenté entre 1962 et 1990, passant de 10 % à 17 %. Toutefois, dans le même temps, le nombre de doubles résidences s’est sans doute nettement accru. Aussi est-on bien incapable de dire si le questionnaire de 1990 a été plus efficace que celui de 1962.

(1) Par exemple, d’après les recensements de 1962 et de 1990, le nombre de résidences secondaires rapporté à celui des résidences principales a doublé entre les deux dates, passant de 6,5 % à 13,1 %.
La collecte
<table>
<thead>
<tr>
<th>Type de logement</th>
<th>Erreurs de recensement</th>
<th>Nombre de résidences principales (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z en ligne</td>
<td>Z de résidences omises</td>
</tr>
<tr>
<td>Maison individuelle</td>
<td>0,5</td>
<td>0,5</td>
</tr>
<tr>
<td>Appartement dans un immeuble principalement ou uniquement à usage d'habitation</td>
<td>0,8</td>
<td>1,5</td>
</tr>
<tr>
<td>Annexe indépendante</td>
<td>16,7</td>
<td>6,9</td>
</tr>
<tr>
<td>Logement dans l'enceinte d'une communauté</td>
<td>0,9</td>
<td>0</td>
</tr>
<tr>
<td>Logement dans un immeuble principalement à usage industriel, commercial, administratif</td>
<td>2,5</td>
<td>0,8</td>
</tr>
<tr>
<td>Ferme</td>
<td>0,4</td>
<td>0</td>
</tr>
<tr>
<td>Logement-foyer pour personnes âgées</td>
<td>0,6</td>
<td>0</td>
</tr>
<tr>
<td>Chambre d'hôtel, construction provisoire, habitation de fortune</td>
<td>24,7</td>
<td>0</td>
</tr>
<tr>
<td>ENSEMBLE**</td>
<td>0,7</td>
<td>0,9</td>
</tr>
</tbody>
</table>

*Résidences principales à la date de l'enquête, après exclusion de celles dont le ménage a déménagé entre la date du recensement et celle de l'enquête*  

**Y compris type de logement non renseigné (57 000 logements)
### TABLEAU 2

**Erreurs de recensement des résidences principales* par taille de l'immeuble d'habitation**

$I$ en ligne

<table>
<thead>
<tr>
<th>I de résidences principales classées en logements vacants, en résidences secondaires, en logements occasionnels</th>
<th>Nombre de résidences principales (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>comises</td>
<td></td>
</tr>
<tr>
<td>Maison individuelle</td>
<td>0,5</td>
</tr>
<tr>
<td>Immeuble de 2 à 5 logements</td>
<td>2,2</td>
</tr>
<tr>
<td>Immeuble de 6 à 10 logements</td>
<td>0,5</td>
</tr>
<tr>
<td>Immeuble de 11 à 20 logements</td>
<td>0,4</td>
</tr>
<tr>
<td>Immeuble de 21 logements et plus</td>
<td>0,3</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td><strong>0,6</strong></td>
</tr>
</tbody>
</table>

*Résidences principales à la date de l'enquête, après exclusion de celles dont le ménage a déménagé entre la date du recensement et celle de l'enquête

Le champ est constitué des maisons individuelles, et des appartements (hors annexes indépendantes) situés dans un immeuble principalement ou uniquement à usage d'habitation.*
TABLEAU 3

Erreurs de recensement des résidences principales* selon la taille du ménage

<table>
<thead>
<tr>
<th>Nombre de personnes du ménage</th>
<th>Nombre de résidences principales classées en logements vacants, en résidences secondaires, en logements occasionnels</th>
<th>Nombre de résidences principales (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,4</td>
<td>2,0</td>
</tr>
<tr>
<td>2</td>
<td>0,6</td>
<td>0,8</td>
</tr>
<tr>
<td>3</td>
<td>0,3</td>
<td>0,5</td>
</tr>
<tr>
<td>4</td>
<td>0,3</td>
<td>0,3</td>
</tr>
<tr>
<td>5</td>
<td>0,4</td>
<td>0,3</td>
</tr>
<tr>
<td>6 et plus</td>
<td>0,5</td>
<td>0,4</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>0,7</td>
<td>0,9</td>
</tr>
</tbody>
</table>

* Résidences principales à la date de l’enquête, après exclusion de celles dont le ménage a déménagé entre la date du recensement et celle de l’enquête
<table>
<thead>
<tr>
<th></th>
<th>I de résidences principales</th>
<th>classées en logements vacants, en résidences secondaires, en logements occasionnels</th>
<th>Nombre de résidences principales (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>omises</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commune rurale appartenant à un canton entièrement rural</td>
<td>0,5</td>
<td>0,7</td>
<td>2 509</td>
</tr>
<tr>
<td>Commune rurale appartenant à un canton partiellement urbain</td>
<td>0,6</td>
<td>0,5</td>
<td>2 892</td>
</tr>
<tr>
<td>Unité urbaine de moins de 20 000 habitants</td>
<td>0,6</td>
<td>0,8</td>
<td>3 346</td>
</tr>
<tr>
<td>Unité urbaine de 20 000 à moins de 100 000 habitants</td>
<td>0,2</td>
<td>0,7</td>
<td>2 802</td>
</tr>
<tr>
<td>Unité urbaine de 100 000 à moins de 200 000 habitants</td>
<td>0,5</td>
<td>0,7</td>
<td>1 465</td>
</tr>
<tr>
<td>Unité urbaine de 200 000 habitants ou plus, sauf agglomération parisienne</td>
<td>0,9</td>
<td>1,3</td>
<td>4 095</td>
</tr>
<tr>
<td>Agglomération parisienne</td>
<td>1,3</td>
<td>1,5</td>
<td>3 254</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td><strong>0,7</strong></td>
<td><strong>0,9</strong></td>
<td><strong>20 363</strong></td>
</tr>
</tbody>
</table>

*Résidences principales à la date de l'enquête, après exclusion de celles dont le ménage a déménagé entre la date du recensement et celle de l'enquête*
**TABLEAU 5**

Taux d'omission des résidences principales selon que le contrôle TH (contrôle par rapprochement avec le fichier de la taxe d'habitation) a été ou non effectué*  

<table>
<thead>
<tr>
<th>Zones soumises au contrôle TH</th>
<th>Zones non soumises au contrôle TH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taux d'omission des résidences principales</td>
<td>Nombre de résidences principales (en milliers)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Unité urbaine de moins de 20 000 habitants (commune rurale exclue)</td>
<td>0,5</td>
</tr>
<tr>
<td>Unité urbaine de 20 000 à moins de 100 000 habitants</td>
<td>0,1</td>
</tr>
<tr>
<td>Unité urbaine de 100 000 à moins de 200 000 habitants</td>
<td>0,3</td>
</tr>
<tr>
<td>Unité urbaine de 200 000 habitants ou plus (agglomération parisienne exclue)</td>
<td></td>
</tr>
<tr>
<td>. dont communes de 20 000 habitants et plus</td>
<td>0,8</td>
</tr>
<tr>
<td></td>
<td>0,9</td>
</tr>
<tr>
<td>Agglomération parisienne</td>
<td>1,3</td>
</tr>
<tr>
<td>. dont communes de 20 000 habitants et plus</td>
<td>1,4</td>
</tr>
</tbody>
</table>

* Les communes rurales sont exclues du champ du tableau, car elles n'ont généralement pas été soumises au contrôle TH. On a exclu également le petit nombre de logements (52 000) pour lesquels on ne sait pas s'ils sont situés sur une zone soumise ou non au contrôle TH.
TABLEAU 6
Erreurs de recensement des résidences principales* par ZEAT

<table>
<thead>
<tr>
<th>Z.E.A.T</th>
<th>% de résidences principales classées en logements vacants, en résidences secondaires, en logements occasionnels</th>
<th>Nombre de résidences principales (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile de France</td>
<td>1,1                                                               1,4                                                  3 724</td>
<td></td>
</tr>
<tr>
<td>Bassin parisien</td>
<td>0,3                                                               0,6                                                  3 714</td>
<td></td>
</tr>
<tr>
<td>Nord</td>
<td>0,9                                                               0,2                                                  1 352</td>
<td></td>
</tr>
<tr>
<td>Est</td>
<td>0,3                                                               0,7                                                  1 602</td>
<td></td>
</tr>
<tr>
<td>Ouest</td>
<td>0,5                                                               0,9                                                  2 579</td>
<td></td>
</tr>
<tr>
<td>Sud-Ouest</td>
<td>0,5                                                               1,1                                                  2 275</td>
<td></td>
</tr>
<tr>
<td>Centre-Est</td>
<td>0,3                                                               0,6                                                  2 445</td>
<td></td>
</tr>
<tr>
<td>Méditerranée</td>
<td>1,5                                                               1,5                                                  2 672</td>
<td></td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>0,7                                                               0,9                                                  20 363</td>
<td></td>
</tr>
</tbody>
</table>

* Résidences principales à la date de l'enquête, après exclusion de celles dont le ménage a déménagé entre la date du recensement et celle de l'enquête
TABLEAU 7

Erreurs de recensement des personnes : comparaison des résultats globaux de 1990 à ceux de 1962

I en colonne

<table>
<thead>
<tr>
<th>1990</th>
<th>1962</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omissions caractérisées</td>
<td>1,5</td>
</tr>
<tr>
<td>Omissions probables</td>
<td>0,3</td>
</tr>
<tr>
<td>Omissions possibles</td>
<td>0,2</td>
</tr>
<tr>
<td>Recensés une seule fois</td>
<td>93,9</td>
</tr>
<tr>
<td>Recensés probablement ou peut-être une fois (doubles comptes possibles)</td>
<td>3,4</td>
</tr>
<tr>
<td>Doubles comptes possibles</td>
<td>0,1</td>
</tr>
<tr>
<td>Doubles comptes caractérisés</td>
<td>0,6</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>100</td>
</tr>
</tbody>
</table>

* Il s'agit des militaires stationnés en Algérie : un grand nombre des bulletins du recensement remplis aux Armées ont été égarés, du fait des événements politiques, lors de leur transfert en métropole. Il semble que la plupart de ces omissions aient pu être corrigées.
### Tableau 8

**Erreurs de recensement des personnes en fonction de diverses situations**

(déménagement, rattachement à plusieurs résidences)

<table>
<thead>
<tr>
<th>Situation</th>
<th>% d'omissions caractérisées</th>
<th>% de doubles comptes caractérisés</th>
<th>% de doubles comptes possibles</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A changé de résidence principale (logement ordinaire) en mars 1990 ou après</td>
<td>7,8</td>
<td>3,9</td>
<td>0</td>
<td>1 135</td>
</tr>
<tr>
<td>Rattaché à plusieurs logements ordinaires</td>
<td>1,9</td>
<td>0,6</td>
<td>0,3</td>
<td>6 541</td>
</tr>
<tr>
<td>Rattaché à une communauté (et au logement enquêté)</td>
<td>4,7</td>
<td>1,0</td>
<td>1,3</td>
<td>1 492</td>
</tr>
<tr>
<td>A une seule résidence (= logement enquêté)</td>
<td>1,2</td>
<td>0</td>
<td>0</td>
<td>44 850</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td>1,5</td>
<td>0,3</td>
<td>0,2</td>
<td>54 018</td>
</tr>
</tbody>
</table>

La collecte
<table>
<thead>
<tr>
<th>Communauté</th>
<th>Z d'omissions caractérisées</th>
<th>Z d'omissions probables</th>
<th>Z de doubles comptes caractérisés</th>
<th>Z de doubles comptes probables</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSEMBLE</td>
<td>4,7</td>
<td>1,0</td>
<td>1,3</td>
<td>5,1</td>
<td>2,6</td>
</tr>
<tr>
<td>dont</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caserne</td>
<td>4,9</td>
<td>0,3</td>
<td>1,8</td>
<td>0,8</td>
<td>2,2</td>
</tr>
<tr>
<td>Internat</td>
<td>1,7</td>
<td>0,7</td>
<td>0,3</td>
<td>2,2</td>
<td>4,9</td>
</tr>
<tr>
<td>Cité Universitaire</td>
<td>17,5</td>
<td>0,7</td>
<td>3,4</td>
<td>20,0</td>
<td>0</td>
</tr>
</tbody>
</table>
### TABLEAU 10

**Effectifs vivant en communauté dans le recensement et dans l'enquête**

<table>
<thead>
<tr>
<th></th>
<th>Recensement</th>
<th>Enquête*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internat ou caserne**</td>
<td>871 000</td>
<td>832 000</td>
</tr>
<tr>
<td>Cité universitaire</td>
<td>138 000</td>
<td>115 000</td>
</tr>
<tr>
<td>Foyer de travailleurs</td>
<td>176 000</td>
<td>39 000</td>
</tr>
<tr>
<td>Maison de retraite</td>
<td>357 000</td>
<td>14 000</td>
</tr>
<tr>
<td>Autre communauté (y compris, dans l'enquête, communauté non définie)</td>
<td>334 000</td>
<td>82 000</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td>1 876 000</td>
<td>1 082 000</td>
</tr>
</tbody>
</table>

* Dans ce tableau, les effectifs de l'enquête correspondent aux personnes vivant en communauté à la date de l'enquête (à l'exclusion de celles qui y ont vécu entre le 1er janvier et une date antérieure à l'enquête) et rattachées à un ménage ordinaire.

** Au recensement, l'effectif se répartit comme suit : 525 000 internes réintégrés, 247 000 militaires réintégrés, et 99 000 internes ou militaires non réintégrés (à la date de publication de ce document, la distinction de ces deux dernières catégories dans le recensement n'est pas disponible). Dans l'enquête, on compte 567 000 élèves internes et 265 000 militaires.
TABLEAU 11

Typologie des omissions (caractérisées et probables) de personnes

<table>
<thead>
<tr>
<th>ENSEMBLE, dont :</th>
<th>Nombre de personnes omises</th>
<th>I (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnes rattachées à une communauté</td>
<td>961 000</td>
<td>100</td>
</tr>
<tr>
<td>Personnes ayant déménagé en mars 1990 ou après</td>
<td>163 000</td>
<td>17</td>
</tr>
<tr>
<td>AUTRES CAS</td>
<td>713 000</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nombre de personnes omises</th>
<th>I (1)</th>
<th>dont rattachées à un seul logement</th>
<th>I (1)</th>
<th>dont rattachées à plusieurs logements</th>
<th>I (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTRES CAS, dont :</td>
<td>713 000</td>
<td>74</td>
<td>551 000</td>
<td>57</td>
<td>162 000</td>
</tr>
<tr>
<td>. leur résidence principale a été omise au recensement</td>
<td>235 000</td>
<td>25</td>
<td>194 000</td>
<td>21</td>
<td>41 000</td>
</tr>
<tr>
<td>. leur résidence principale a été classée dans une autre catégorie au recensement</td>
<td>232 000</td>
<td>24</td>
<td>180 000</td>
<td>18</td>
<td>52 000</td>
</tr>
<tr>
<td>. personnes omises isolément (leur résidence principale a bien été recensée)</td>
<td>246 000</td>
<td>25</td>
<td>177 000</td>
<td>18</td>
<td>69 000</td>
</tr>
</tbody>
</table>

(1) I par rapport à l'ensemble des omissions
<table>
<thead>
<tr>
<th>Personne vivant seule</th>
<th>2,8</th>
<th>0,4</th>
<th>1,4</th>
<th>5 465</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personne de référence (d'un ménage d'au moins deux personnes)</td>
<td>1,1</td>
<td>0,2</td>
<td>0,6</td>
<td>15 502</td>
</tr>
<tr>
<td>Conjoint</td>
<td>1,0</td>
<td>0,2</td>
<td>0,5</td>
<td>13 237</td>
</tr>
<tr>
<td>Enfant</td>
<td>1,4</td>
<td>0,3</td>
<td>0,8</td>
<td>18 024</td>
</tr>
<tr>
<td>Petit-enfant</td>
<td>7,1</td>
<td>0,3</td>
<td>1,0</td>
<td>269</td>
</tr>
<tr>
<td>Ascendant</td>
<td>2,3</td>
<td>0,5</td>
<td>1,4</td>
<td>504</td>
</tr>
<tr>
<td>Autre parent</td>
<td>5,9</td>
<td>0</td>
<td>1,1</td>
<td>425</td>
</tr>
<tr>
<td>Ami</td>
<td>11,6</td>
<td>2,3</td>
<td>3,6</td>
<td>320</td>
</tr>
<tr>
<td>Salarié logé, étudiant logé, pensionnaire, sous-locataire</td>
<td>8,5</td>
<td>2,8</td>
<td>2,8</td>
<td>166</td>
</tr>
<tr>
<td>ENSEMBLE*</td>
<td>1,5</td>
<td>0,3</td>
<td>0,7</td>
<td>54 018</td>
</tr>
</tbody>
</table>

* Y compris lien avec la personne de référence non renseigné (106 000 personnes)
**TABLEAU 13**

Erreurs de recensement des personnes selon la taille du ménage

<table>
<thead>
<tr>
<th>Nombre de personnes du ménage</th>
<th>% d'omissions caractérisées</th>
<th>% d'omissions probables</th>
<th>% de doubles comptes</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,8</td>
<td>0,4</td>
<td>1,4</td>
<td>5 465</td>
</tr>
<tr>
<td>2</td>
<td>1,7</td>
<td>0,3</td>
<td>0,8</td>
<td>12 540</td>
</tr>
<tr>
<td>3</td>
<td>1,6</td>
<td>0,2</td>
<td>0,8</td>
<td>11 398</td>
</tr>
<tr>
<td>4</td>
<td>0,8</td>
<td>0,3</td>
<td>0,5</td>
<td>12 833</td>
</tr>
<tr>
<td>5</td>
<td>1,3</td>
<td>0,2</td>
<td>0,6</td>
<td>6 935</td>
</tr>
<tr>
<td>6 et plus</td>
<td>1,6</td>
<td>0,2</td>
<td>0,5</td>
<td>4 847</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td>1,5</td>
<td>0,3</td>
<td>0,7</td>
<td>54 018</td>
</tr>
<tr>
<td></td>
<td>I d'omissions</td>
<td>I de doubles comptes</td>
<td>Effectifs (en milliers)</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>caractérisées</td>
<td>probables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commune rurale appartenant à un canton entièrement rural</td>
<td>0,8</td>
<td>0,1</td>
<td>6 867</td>
<td></td>
</tr>
<tr>
<td>Commune rurale appartenant à un canton partiellement urbain</td>
<td>1,0</td>
<td>0,1</td>
<td>8 375</td>
<td></td>
</tr>
<tr>
<td>Unité urbaine de moins de 20 000 habitants</td>
<td>1,3</td>
<td>0,2</td>
<td>8 737</td>
<td></td>
</tr>
<tr>
<td>Unité urbaine de 20 000 à moins de 100 000 habitants</td>
<td>1,3</td>
<td>0,3</td>
<td>7 180</td>
<td></td>
</tr>
<tr>
<td>Unité urbaine de 100 000 à moins de 200 000 habitants</td>
<td>1,6</td>
<td>0,3</td>
<td>4 023</td>
<td></td>
</tr>
<tr>
<td>Unité urbaine de 200 000 habitants ou plus, sauf agglomération parisienne</td>
<td>2,1</td>
<td>0,3</td>
<td>10 468</td>
<td></td>
</tr>
<tr>
<td>Agglomération parisienne</td>
<td>2,5</td>
<td>0,4</td>
<td>8 368</td>
<td></td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td><strong>1,5</strong></td>
<td><strong>0,3</strong></td>
<td><strong>54 018</strong></td>
<td></td>
</tr>
</tbody>
</table>

La collecte 105
TABLEAU 15

Erreurs de recensement des personnes par âge et sexe

<table>
<thead>
<tr>
<th>Age atteint au 31/12/1990 (en ans)</th>
<th>% d'omissions caractérisées</th>
<th>% d'omissions probables</th>
<th>% de doubles comptes</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (nés en 1990)</td>
<td>3,3</td>
<td>1,6</td>
<td>0,0</td>
<td>108</td>
</tr>
<tr>
<td>1 à 9</td>
<td>1,3</td>
<td>0,2</td>
<td>0,4</td>
<td>6 823</td>
</tr>
<tr>
<td>10 à 19</td>
<td>1,2</td>
<td>0,3</td>
<td>0,9</td>
<td>7 746</td>
</tr>
<tr>
<td>20 à 24</td>
<td>3,8</td>
<td>0,8</td>
<td>2,5</td>
<td>4 161</td>
</tr>
<tr>
<td>25 à 29</td>
<td>2,6</td>
<td>0,5</td>
<td>0,9</td>
<td>4 285</td>
</tr>
<tr>
<td>30 à 34</td>
<td>1,6</td>
<td>0,2</td>
<td>0,9</td>
<td>4 079</td>
</tr>
<tr>
<td>35 à 39</td>
<td>1,3</td>
<td>0,2</td>
<td>0,4</td>
<td>3 946</td>
</tr>
<tr>
<td>40 à 49</td>
<td>1,2</td>
<td>0,1</td>
<td>0,3</td>
<td>6 918</td>
</tr>
<tr>
<td>50 à 59</td>
<td>1,2</td>
<td>0,1</td>
<td>0,5</td>
<td>5 525</td>
</tr>
<tr>
<td>60 à 69</td>
<td>0,8</td>
<td>0,1</td>
<td>0,6</td>
<td>5 341</td>
</tr>
<tr>
<td>70 à 79</td>
<td>1,0</td>
<td>0,0</td>
<td>0,3</td>
<td>3 051</td>
</tr>
<tr>
<td>80 et plus</td>
<td>1,5</td>
<td>0,3</td>
<td>0,8</td>
<td>1 984</td>
</tr>
<tr>
<td>ENSEMBLE*</td>
<td>1,5</td>
<td>0,3</td>
<td>0,7</td>
<td>54 018</td>
</tr>
</tbody>
</table>

| HOMMES                          | 1,7                         | 0,2                     | 0,8                 | 26 226                 |
| FEMMES                          | 1,4                         | 0,2                     | 0,7                 | 27 783                 |

* Y compris âge non renseigné (49 000 personnes), ou sexe non renseigné (9 000 personnes)
TABLEAU 16

Erreurs de recensement des personnes par nationalité

<table>
<thead>
<tr>
<th>Nationalité</th>
<th>I d'omissions caractérisées</th>
<th>I de doubles comptes</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Français de naissance</td>
<td>1,4</td>
<td>0,2</td>
<td>48 472</td>
</tr>
<tr>
<td>Français par acquisition</td>
<td>1,3</td>
<td>0,4</td>
<td>1 418</td>
</tr>
<tr>
<td>Etrangers*</td>
<td>3,3</td>
<td>0,3</td>
<td>3 694</td>
</tr>
<tr>
<td>- pays de la CEE</td>
<td>2,3</td>
<td>0,2</td>
<td>1 266</td>
</tr>
<tr>
<td>- Maghreb</td>
<td>4,2</td>
<td>0,4</td>
<td>1 560</td>
</tr>
<tr>
<td>- autre pays</td>
<td>3,4</td>
<td>0,3</td>
<td>805</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong>*</td>
<td><strong>1,5</strong></td>
<td><strong>0,3</strong></td>
<td><strong>54 018</strong></td>
</tr>
</tbody>
</table>

* Y compris nationalité non renseignée (434 000 personnes, dont 63 000 étrangers)

NB. Les étrangers vivant en communauté (foyers de travailleurs,...) et non rattachés à un ménage ordinaire sont exclus du champ de l’enquête.
### Erreurs de recensement des personnes âgées de 14 ans et plus selon leur situation professionnelle

I en ligne

<table>
<thead>
<tr>
<th></th>
<th>I d'omissions caractérisées</th>
<th>I de doubles comptes</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actifs ayant un emploi</td>
<td>1,6</td>
<td>0,7</td>
<td>21 126</td>
</tr>
<tr>
<td>- Agriculteurs</td>
<td>0,6</td>
<td>0,4</td>
<td>975</td>
</tr>
<tr>
<td>- Artisans, commerçants et chefs d'entreprise</td>
<td>2,2</td>
<td>0,2</td>
<td>1 617</td>
</tr>
<tr>
<td>- Cadres, professions intellectuelles supérieures</td>
<td>1,8</td>
<td>0,8</td>
<td>2 388</td>
</tr>
<tr>
<td>- Professions intermédiaires</td>
<td>1,5</td>
<td>0,7</td>
<td>3 989</td>
</tr>
<tr>
<td>- Employés</td>
<td>1,6</td>
<td>0,8</td>
<td>5 836</td>
</tr>
<tr>
<td>- Ouvriers</td>
<td>1,7</td>
<td>0,6</td>
<td>6 321</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>3,5</td>
<td>0,6</td>
<td>2 370</td>
</tr>
<tr>
<td>Elèves, étudiants, stagiaires non rémunérés</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Âgés de 14 à 18 ans</td>
<td>0,8</td>
<td>0,8</td>
<td>3 557</td>
</tr>
<tr>
<td>- Âgés de plus de 18 ans</td>
<td>4,1</td>
<td>3,6</td>
<td>2 064</td>
</tr>
<tr>
<td>Retraités</td>
<td>1,0</td>
<td>0,5</td>
<td>8 390</td>
</tr>
<tr>
<td>Inactifs*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- de moins de 60 ans</td>
<td>1,7</td>
<td>0,9</td>
<td>4 833</td>
</tr>
<tr>
<td>- de 60 ans et plus</td>
<td>1,0</td>
<td>0,4</td>
<td>1 867</td>
</tr>
<tr>
<td><strong>ENSEMBLE</strong></td>
<td>1,6</td>
<td>0,8</td>
<td>44 207</td>
</tr>
</tbody>
</table>

* Y compris militaires du contingent
TABLEAU 18
Erreurs de recensement des étudiants, élèves, stagiaires non rémunérés âgés de plus de 18 ans selon le nombre et la nature de leurs résidences

<table>
<thead>
<tr>
<th></th>
<th>Z en ligne</th>
<th>I d'omissions</th>
<th>I de doubles comptes</th>
<th>Effectifs (en milliers)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>caractérisées</td>
<td>probables</td>
<td>possibles</td>
</tr>
<tr>
<td>Rattachés à deux résidences, dont :</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Cité universitaire</td>
<td>19,6</td>
<td>0,8</td>
<td>3,0</td>
<td>20,7</td>
</tr>
<tr>
<td>- Internat</td>
<td>2,6</td>
<td>1,4</td>
<td>0,6</td>
<td>1,5</td>
</tr>
<tr>
<td>- Logement en ville d'étudiant</td>
<td>6,5</td>
<td>1,6</td>
<td>0</td>
<td>11,9</td>
</tr>
<tr>
<td>- Autre</td>
<td>4,0</td>
<td>3,1</td>
<td>1,4</td>
<td>4,0</td>
</tr>
<tr>
<td>Rattachés à un seul logement</td>
<td>2,5</td>
<td>0,1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>4,1</td>
<td>0,9</td>
<td>0,5</td>
<td>3,0</td>
</tr>
</tbody>
</table>

La collecte
DOCUMENTS DE COLLECTE

- ETAT DESCRIPTIF D'ARÉA

Seules les pages 1 à 3 du document ont été reproduites ci-après. En effet, les pages 4-5 et 6-7, non reproduites, étaient identiques aux pages 2-3.

- ETAT DESCRIPTIF D'IMMEUBLE

Seules les pages 1 à 3 du document ont été reproduites ci-après. En effet, les pages 4-5 et 6-7, non reproduites, étaient identiques aux pages 2-3.

- QUESTIONNAIRE A (questionnaire rempli pour chaque logement)

- QUESTIONNAIRE B (questionnaire rempli pour chaque ménage interviewé)

Le questionnaire B comprenait 5 "bulletins individuels" et 3 sous-questionnaires "autre résidence". Si nécessaire, l'enquêteur pouvait même insérer dans le questionnaire des bulletins individuels ou des sous-questionnaires "autre résidence" supplémentaires. Le questionnaire B, reproduit ci-après, ne comprend que le sous-questionnaire "autre résidence" n° 1 (pages 6-7 du questionnaire B) et le "bulletin individuel" n° 1 (pages 13-14 du questionnaire B). Les pages manquantes correspondent aux sous-questionnaires "autre résidence" n° 2 et 3, et aux "bulletins individuels" n° 2 à 5.

- CARTE "CODES"

Cette carte a permis aux enquêteurs de chiffrer le code "Lien avec la personne de référence" (Cf questionnaire B, p 2), et les codes "catégorie de logement" et "catégorie de collectivité" des "autres résidences" (Cf questionnaire B, pages 6-7).
# État descriptif d'aire

<table>
<thead>
<tr>
<th>Identifiant de l'aire:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RG</td>
<td>TECH</td>
<td>STTSECH</td>
<td>A</td>
<td>SA</td>
<td>CLE</td>
<td></td>
</tr>
</tbody>
</table>

Département:  
Commune*:  
Nom de l'enquêteur:  
*Si l'aire est à cheval sur 2 communes, ouvrir 2 états descriptifs d'aire.

<table>
<thead>
<tr>
<th>Appariement avec le RP (réservé à la DR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N° d'identification de la commune</td>
</tr>
<tr>
<td>Numéro des districts sur lesquels l'aire est située</td>
</tr>
</tbody>
</table>
ÉNUMÉRATION DE TOUS LES IMMEUBLES DE L'aire
(Y compris ceux qui ne sont pas à usage d'habitation).
Utiliser une ligne pour chaque immeuble. Si un immeuble est décomposé en corps de bâtiment ou en escaliers, utiliser une ligne pour chaque corps de bâtiment, et pour chaque cage d'escalier.

Classer de préférence les immeubles par N° croissant dans la rue, côté des N° impairs d'une part, côté des N° pairs d'autre part.

<table>
<thead>
<tr>
<th>Nom de la voie (rue, place, impasse, etc.) ou nom du lieu-dit</th>
<th>N° dans la voie (ou toute autre indication permettant de localiser l'immeuble). Pour un immeuble collectif, préciser, le cas échéant, le corps de bâtiment et l'escalier</th>
<th>Type d'usage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1. Commune de l'existence d'un logement collectif.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Immeuble entièrement ou principalement à usage industriel, commercial ou administratif. Préciser l'usage.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Immeuble entièrement ou principalement à usage d'habitation.</td>
<td></td>
</tr>
<tr>
<td>TYP</td>
<td>N° LOG</td>
<td>N° immeuble</td>
<td>L</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>-------------</td>
<td>---</td>
</tr>
<tr>
<td>Fermie</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hôtel, pension de famille, garni</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction provisoire, habitation de fortune</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maison individuelle, immeuble collectif</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caravane, habitation mobile</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Correspondance avec les immeubles recensés au RP**

(Réservé à la DR)

- Si 3 en colonne précédente
  1. Fermie
  2. Hôtel, pension de famille, garni
  3. Construction provisoire, habitation de fortune
  4. Maison individuelle, immeuble collectif
  5. Caravane, habitation mobile

- Si 1 ou 2 en colonne précédente, si y a eu moins un logement d'habitation dans l'immeuble et dans l'aire,
  - N° de l'immeuble dans l'enquête

- N° du district au recensement
- Rang de l'immeuble dans le district
- Éventuellement, observations
**ÉTAT DESCRIPTIF D’IMMEUBLE**

(à remplir si l’immeuble comprend au moins 2 logements d’habitation situés dans l’aire de l’enquête)

<table>
<thead>
<tr>
<th>Identifiant de l’aire:</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG 2 2 TECH STTSECH A 2 SA CLE</td>
</tr>
<tr>
<td>N° de l’immeuble dans l’enquête IMM</td>
</tr>
</tbody>
</table>

**Adresse:**

- N°  Voie (ou lieu-dit) ____________________________
- Commune ____________________________ Département ____________________________
- Corps de bâtiment, escalier, ou toute autre indication permettant de repérer l’immeuble: ____________________________

Nombre de niveaux au-dessus du rez-de-chaussé (y compris ceux qui sont situés hors de l’aire de l’enquête): ______

<table>
<thead>
<tr>
<th>Appariement avec le RP (réservé à la DR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N° de district au RP IL FIL CIL</td>
</tr>
<tr>
<td>Rang du ou des immeubles dans le district RGIMM</td>
</tr>
</tbody>
</table>
ÉNUMÉRATION DES LOCAUX DE L'IMMEUBLE SITUÉS DANS L'AIRE
(On remplira une ligne par local, quel que soit l'usage de ce local).

Décrire les locaux étage par étage en partant du bas de l'immeuble.

<table>
<thead>
<tr>
<th>Étage</th>
<th>Localisation du local dans l'immeuble</th>
<th>Si le local n'est pas à usage d'habitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Situation sur le palier (par rapport à la montée de l'escalier)</td>
<td>Usage du local</td>
</tr>
<tr>
<td></td>
<td>Aborder les locaux toujours dans le même sens</td>
<td></td>
</tr>
</tbody>
</table>

118  
INSEE Méthodes n° 52 - 53
Si le local est à usage d'habitation

<table>
<thead>
<tr>
<th>Nom(s) des occupants</th>
<th>N° du logement à l'enquête</th>
<th>ANNEXE pour les logements avec annexes indépendantes</th>
<th>Rang de l'immeuble dans le district</th>
<th>Rang du logement dans l'immeuble</th>
<th>Eventuellement, observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

La collecte

119
ENQUÊTE MÉTHODOLOGIQUE
POST-CENSITAIRE 1990
questionnaire A

Identifiant du logement:

RG TECH STTSECH A SA CLE

IMM LOC

*Dans le cas d'un logement comportant une ou plusieurs annexes indépendantes, inscrire 0 pour le logement principal, inscrire 1 à 9 pour la ou les annexes indépendantes (quel que soit leur usage).*

Nom de l'enquêteur: _______________________

N° de l'enquêteur: _______________________

Réservé à la DR

Contrôle TH effectué:

1 - oui

2 - non

TH

1. ADRESSE DU LOGEMENT

N° ____ Rue (ou lieu-dit) ___________________________________________________________

Commune ____________________________ Département ____________________________

(Pour Paris, Lyon, Marseille, préciser l'arrondissement)

2. LOCALISATION DU LOGEMENT DANS L'IMMEUBLE:

a. Escalier __________

b. Étage __________

c. Situation sur le palier ________________________________

Si la porte d'entrée du logement a un numéro, indiquer ce numéro _______________________

3. NOM(S) DES OCCUPANTS à la date de l'enquête:

_________________________________________________________

(Éventuellement) n° de téléphone: __ __ __ __ __ __

Visa n° 90 A 023 Ec du ministre d'État, ministre de l'Économie, des Finances et du Budget, valable pour l'année 1990.

Enquête statistique obligatoire (au n° 51-711 du 7 juin 1951 modifiée sur l'obligation, la coordination et le secret en matière de statistique). Tout défaut de réponse ou une réponse sciemment inexacte peut entraîner l'application d'une peine d'amende.

Questionnaire confidentiel destiné à l'INSEE.

Ce loi n° 78.17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés garantit aux individus un droit d'accès et de rectification pour les informations les concernant. Ce droit d'accès peut être exercé pendant le délai où ces informations sont gardées sous forme nominative auprès des Directions et Services régionaux de l'INSEE.
4. **TYPE DE LOGEMENT:**

- Logement-foyer pour personnes âgées
- Ferme
- Chambre d'hôtel
- Construction provisoire, habitation de fortune
- Annexe indépendante (en général, pièce indépendante ayant sa propre entrée)
- Maison individuelle
- Logement ordinaire situé dans l'enceinte d'une communauté (caserne, internat...)
- Logement dans un immeuble collectif entièrement ou principalement à usage d'habitation
- Logement dans un immeuble principalement à usage industriel, commercial, administratif
- Caravane, habitation mobile faisant office de résidence principale

> Remplir Q.7 (résultat de l'interview).
> Si l'interview est accepté, inscrire sur un questionnaire B la liste des personnes habitant le logement. Sinon, essayer au moins d'indiquer le nombre de personnes en Q.10.

5. **Si le logement est une annexe indépendante rattachée à un logement principal (05 coché à la question précédente)**

a. Cette annexe indépendante est située:

- Sur l'aire de l'enquête
- En dehors de l'aire de l'enquête (le logement principal étant, lui, sur l'aire de l'enquête)

b. La classer dans l'une de ces 2 situations:

- Annexe louée, sous-louée ou prêtée à des particuliers n'appartenant pas au ménage du logement principal
- Autre cas (annexe n'ayant pas d'occupants ou dont tous les occupants sont rattachés au ménage du logement principal)
6. DÉFINITION DU CHAMP DES LOGEMENTS À ENQUÊTER:

- Logement meublé, loué ou à louer pour des séjours touristiques de courte durée dans les stations balnéaires ou de sports d'hiver
- Logement vacant
- Autre

<table>
<thead>
<tr>
<th>CHAMP</th>
<th>Fin de l'enquête</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

7. RÉSULTAT DE L'INTERVIEW (lors de la 1ère vague d'enquête)

- Interview accepté
- Ménage contacté ; enquête impossible à faire
- Ménage contacté ; refus
- Ménage présent ; impossible à joindre
- Absent de longue durée

<table>
<thead>
<tr>
<th>R1</th>
<th>Partie II Q17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

OBSERVATIONS ÉVENTUELLES DE L'ENQUÊTEUR SUR LES PROBLÈMES RENCONTRÉS POUR CONTACTER LE MÉNAGE.
### PARTIE I

A remplir pour tous les ménages non-répondants lors de la 1ère vague d’enquête.

Se faire aider des voisins, du concierge...

#### 8. LE LOGEMENT EST-IL HABITÉ...

- Le plus grande partie de l’année
- Occasionnellement, moins de 3 mois par an ou d’une semaine par mois
- Autre durée d’occupation
- Durée d’occupation inconnue

<table>
<thead>
<tr>
<th>Durée d'occupation</th>
<th>Q. 8, puis fin de l'enquête</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DUR_HAB</td>
</tr>
</tbody>
</table>

#### 9. CATÉGORIE DU LOGEMENT:

- Résidence principale (où le ménage demeure la plus grande partie de l’année)
- Logement utilisé occasionnellement (pour des raisons professionnelles)
- Résidence secondaire
- Catégorie inconnue

<table>
<thead>
<tr>
<th>Catégorie du Logement</th>
<th>CLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

#### 10. NOMBRE DE PERSONNES HABITANT LE LOGEMENT

- Si ce nombre n'est pas connu, chiffrer 99

#### 11. SEXE ET ÂGE APPROXIMATIF DES PERSONNES HABITANT LE LOGEMENT

<table>
<thead>
<tr>
<th>N° individuel</th>
<th>Sexe 1. Masculin</th>
<th>2. Féminin</th>
<th>Âge approximatif</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sexe A</td>
<td>AGE A</td>
<td></td>
</tr>
<tr>
<td>01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
12. COORDONNÉES DE LA PERSONNE FOURNISSANT LES RENSEIGNEMENTS
sur le ménage non-répondant (nom, adresse):

_________________________________________________________________________

• Si le ménage est absent de longue durée, poursuivre en 0,13
• Pour tous les autres cas (résultat de l’interview = 2 à 4), FIN DE L’ENQUÊTE (1er vague)

13. A) LE OU LES OCCUPANTS DE CE LOGEMENT DISPOSENT-ILS D’UNE DEUXIÈME
RESIDENCE, D’UNE RESIDENCE SECONDAIRE, PAR EXEMPLE; OU LOGENT-ILS UNE
PARTIE DE L’ANNÉE AILLEURS, PAR EXEMPLE, CHEZ DES PARENTS?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oui</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ne sait pas</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B) SI OUI, ADRESSE DE CET AUTRE LOGEMENT:

N° ___________________________ Voie (ou lieu-dit) ________________________________
Commune __________________________ Département _____________________________

N° de téléphone: __________________________

->FIN DE L’ENQUÊTE (1er vague)
### RATTRAPAGE DES NON-RÉPONDANTS
(sauf si le logement est habité moins de 3 mois par an ou moins d’une semaine par mois - 2 coché en Q.8)

### RATTRAPAGE PAR TÉLÉPHONE
(pour tous les non-répondants, sauf les refus)

<table>
<thead>
<tr>
<th>RÉSULTAT</th>
<th>Description</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Réalisé avec succès</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>au logement décrit en Q.13 b</td>
<td>Parle II, Q.17</td>
</tr>
<tr>
<td>3</td>
<td>au logement situé sur l’aire</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Refus</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Pas de réponse au téléphone</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Rattrapage impossible à faire</td>
<td>RATEL</td>
</tr>
</tbody>
</table>

**Remplir un questionnaire postal:**

- Ne pas remplir la partie II du questionnaire A, passer directement au questionnaire B. Attention, le logement où s’effectue l’enquête n’est pas celui de l’échantillon (voir consignes dans l’introduction).

### RATTRAPAGE PAR ENQUÊTEUR LORS DE LA 2e VAGUE D’ENQUÊTE
(pour tous les non-répondants – sauf les refus – n’ayant pu être rattrapés par téléphone).

<table>
<thead>
<tr>
<th>RÉSULTAT</th>
<th>Description</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interview acceptée</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Ménage contacté; enquête impossible à faire</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Ménage contacté; refus</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Ménage présent; impossible à joindre</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Absent de longue durée</td>
<td></td>
</tr>
</tbody>
</table>

### RATTRAPAGE PAR QUESTIONNAIRE POSTAL
(réservé à la DR)
(pour tous les refus lors de la 1e vague, et tous les autres non-répondants n’ayant pu être rattrapés lors de la 2e vague)

<table>
<thead>
<tr>
<th>RÉSULTAT</th>
<th>Description</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Questionnaire retourné par l’enquêté</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>et correctement rempli</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>mais inexploitable (non rempli ou incomplet)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Questionnaire non retourné par l’enquêté</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Questionnaire retourné par la poste</td>
<td></td>
</tr>
</tbody>
</table>

**Remplir un questionnaire et à partir du questionnaire postal:**

* Ne pas remplir la partie II du questionnaire A, passer directement au questionnaire B. Attention, le logement où s’effectue l’enquête n’est pas celui de l’échantillon (voir consignes dans l’introduction).
PARTIE II
A remplir pour tous les ménages ayant accepté l'interview lors de la 1ère ou de la 2ème vague d'enquête.
(1 coché en Q.7 ou en Q.16).

17. ÉLÉMENTS DE CONFORT DU LOGEMENT
a) Nombre de pièces d'habitation ___________________________ [ ] PIECE
- Compter les pièces telles que chambres à coucher, salle à manger, salle de séjour, quelle que soit leur surface.
- Ne compter la cuisine que si sa surface est supérieure à 12 m².
- Ne pas compter les pièces telles que couloir, salle de bains, W-C, buanderie, etc.
- Ne pas compter les pièces à usage exclusivement professionnel (exemples: cabinet de médecin, atelier d'artisan, etc.).

b) Avez-vous une baignoire ou une douche?
- Baignoire ___________________________ 1
- Douche seulement ______________________ 2
- Ni baignoire, ni douche __________________ 3

BAIN

c) Les W-C sont-ils situés à l'intérieur du logement?
- Oui ___________________________ 1
- Non ___________________________ 2

WC

18. STATUT D'OCCUPATION DU LOGEMENT
Étes-vous?
- Propriétaire du logement (y compris les différentes formes d'accèsion à la propriété) ___________________________ 1
- Logé gratuitement, par exemple par des parents, des amis, ou votre employeur (y compris la cas des personnes occupent un logement qu'elles ont vendu en viager ou dont elles ont la jouissance par usufruit) ___________________________ 2
- Locataire ou sous-locataire d'un logement loué vide ___________________________ 3
- Locataire ou sous-locataire d'un logement loué meublé ou d'une chambre d'hôtel ___________________________ 4
- Autre cas (squatter) ___________________________ 5

SO
19. POUR TOUS LES LOGEMENTS, SAUF LES ANNEXES INDEPENDANTES

a) Votre logement a-t-il une annexe indépendante, par exemple une chambre de bonne?

- Oui
- Non

* Ouvrir un questionnaire A pour cette annexe, quel que soit son usage.

b) Si oui en question a,
Depuis le 1er janvier, cette annexe indépendante vous a-t-elle servi à loger, même temporairement, un membre de votre famille ou un salarié à votre service?

- Oui
- Non

* Inscrire la ou les personnes sur le questionnaire B du logement principal (c'est le logement en cours d'enquête).

Ne pas oublier de poser la question suivante (c).

c) Quelle que soit la réponse en b,
Actuellement, cette annexe indépendante est-elle louée, sous-louée ou prêtée à des particuliers?

- Oui
- Non

* Inscrire la ou les personnes sur le questionnaire B se rapportant à l'annexe indépendante.

POUR TOUS LES LOGEMENTS, PASSER AU QUESTIONNAIRE B.
RÉSERVÉ À LA DR
RÉSULTAT DE L’APPARIEMENT
(reservé à la DR)

20. CATÉGORIE DU LOGEMENT AU RECENSEMENT

a) Logement recensé comme:

• Résidence principale
• Logement utilisé occasionnellement
• Résidence secondaire ou logement loué (ou à louer) pour les loisirs ou les
  vacances
• Logement vacant
• Caravane, habitation mobile
• Pièce indépendante ne formant pas un logement à part entière
• Logement oublié au recensement

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>recensé</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) (À remplir le cas échéant). Logement recensé, mais qui n’existait pas à la date du RP (fictif)

RÉCIP

Ne pas oublier d’ouvrir un questionnaire B avec la liste des personnes recensées, et de recopier les BI du
recensement.

21. CATÉGORIE DU LOGEMENT À L’ENQUÊTE (à chiffrer par la DR)

• Résidence principale
  – Logement en ville d’un étudiant
  – Autre résidence principale (où le ménage demeure la plus grande partie
    de l’année)
• Logement utilisé occasionnellement
• Résidence secondaire ou logement loué (ou à louer) pour les loisirs ou les
  vacances
• Logement vacant
• Annexe indépendante ne formant pas un logement à part entière
• Catégorie inconnue ou incertaine (pour certains ménages non-répondants)

<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>12</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>recensé</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CLENQ

INSEE Méthodes n° 52 - 53
22. SI LE LOGEMENT A ÉTÉ OUBLIÉ AU RP, INDIQUER LES RAISONS APPARENTES DE CETTE OMISSION :

- Logement dans un immeuble principalement à usage industriel, commercial ou administratif
- Logement ordinaire dans une communauté
- Logement dont l'accès est peu visible et situé :
  - au dernier étage d'un immeuble
  - au sous-sol d'un immeuble
  - autre situation
- Logement dans un local qui normalement n'est pas utilisable pour l'habitation (cave, hangar)
- Le travail de l'agent recenseur est de mauvaise qualité sur le district qui lui a été confié
- Autre raison (précisez)
- Pas de raison apparente

RAIOUB

23. SI LA CATÉGORIE DE LOGEMENT AU RECENSEMENT N'EST PAS MÈME QU'À L'ENQUÊTE, INDIQUER, SI POSSIBLE, L'ORIGINE DE LA DIVERGENCE :

- Changement de catégorie entre la date du recensement et celle de l’enquête
- Erreur manifeste au recensement
- Catégorie du logement difficile à déterminer (cas des ménages ayant deux résidences occupées chacune à peu près à mi-temps, par exemple)
- Ne sait pas

ORDIV

La collecte
À REMPLIR PAR LA DR
ÉVENTUELLEMENT, CONSIGNES À L’ENQUÊTEUR EN VUE DE LA 2e VAGUE D’ENQUÊTE

À REMPLIR PAR L’ENQUÊTEUR
RÉSULTATS DE LA 2e VAGUE D’ENQUÊTE

NB : L’enquêteur devra notamment préciser les modifications qu’il a apportées au questionnaire, suite à cette deuxième vague d’enquête.
Identifiant du logement (à reporter)

RG TECH STIECH A SA DUE

IAM LOC

ANNEXE

DATE DE COLLECTE JOUR MOIS

Visa n° 90A033 Ec du Ministre d'Etat, Ministre de l'Economie, des Finances et du Budget, valable pour l'année 1990.

Enquête statistique obligatoire (loi n° 51-711 du 7 juin 1951 modifiée sur l'obligation, la coordination et le secret en matière de statistiques). Tout défaut de réponse ou une réponse sciemment inexacte peut entraîner l'application d'une peine d'amende.

Questionnaire confidentiel destiné à l'INSEE.

La loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés garantit aux individus un droit d'accès et de rectification pour les informations les concernant. Ce droit d'accès peut être exercé pendant le délai où ces informations sont gardées sous forme nominative auprès des Directions et Services régionaux de l'INSEE.
1. Liste des personnes ayant habité le logement depuis le 1er janvier 1990

1. Quelles sont les personnes qui habitent dans ce logement, même si elles sont temporairement absentes, même si elles n’y vivent pas régulièrement ?
   - Les inscrire ci-dessous. Pour chaque personne, remplir une ligne jusqu’au double trait.

   Entourer le prénom du ou des répondants.

<table>
<thead>
<tr>
<th>NOM</th>
<th>Sexe</th>
<th>Date de naissance</th>
<th>Liens avec la personne de référence du ménage</th>
<th>M. loge-t-il ici (ou logeait-il ici) la plus grande partie de l’année ?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>JOUR MOIS ANNEES 5 derniers chiffres ct. carte</td>
<td>1. oui</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>o trouver colonne g suivante</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b)</th>
<th>(c)</th>
<th>(d1)</th>
<th>(d2)</th>
<th>(d3)</th>
<th>(e)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEXE</td>
<td>JNAIS</td>
<td>MNAIS</td>
<td>ANAIS</td>
<td>LRMM1/LRMM2</td>
<td>LGPA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Depuis le 1er janvier dernier, y-a-t-il d'autres personnes qui ont logé ici, avec vous, même temporairement ?

<table>
<thead>
<tr>
<th>OUI</th>
<th>NON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

*Les inscrire ci-dessous. Pour chaque personne, remplir une ligne jusqu'en dernière colonne (1 ou 2).*

3. N'avez-vous oublié personne ? Par exemple, depuis le 1er janvier dernier, avez-vous logé :

- un locataire ou un sous-locataire ?
- un pensionnaire ou un enfant placé en nourrice nuit et jour chez vous ?
- un salarié à votre service, par exemple, un apprenti, une jeune fille au pair, un domestique, un salarié agricole ?

<table>
<thead>
<tr>
<th>SI</th>
<th>NON</th>
<th>ABSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

*(si 3 ou 4 en colonne I précédente)*

<table>
<thead>
<tr>
<th>(g)</th>
<th>(h1)</th>
<th>(h2)</th>
<th>(i1)</th>
<th>(i2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>MEM</td>
<td>DE</td>
<td>MOE</td>
<td>JOE</td>
</tr>
<tr>
<td></td>
<td>mois</td>
<td>mois</td>
<td>si décès,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>durée</td>
<td></td>
</tr>
</tbody>
</table>

4. N'avez-vous pas oublié une personne habitant normalement ici, mais absente actuellement, parce qu'elle loge temporairement ailleurs, par exemple :

- un jeune vivant en caserne, en internat, en pension, dans un foyer ?
- une personne à l'hôpital, ou logée ailleurs pour son travail ?

*Si 1 ou 3 en colonne I, les inscrire ci-dessous (si 1, 3 ou 4 en colonne I)*
II. CATÉGORIE DU LOGEMENT ENQUÊTÉ

5. A remplir directement par l'enquêteur, d'après les réponses aux questions précédentes, ou en demandant un complément d'information. Éventuellement, commencer par poser les questions ci-dessous (la 0.6 notamment).
- Logement en ville d'un étudiant
- Résidence principale (ou le menage demeure la plus grande partie de l'année)
- Logement utilisé occasionnellement (pour des raisons professionnelles)
- Résidence secondaire

Observations éventuelles de l'enquêteur (en cas de classement délicat):

III. LISTE DES AUTRES RÉSIDENCES

6. Disposez-vous d'une autre résidence où vous pouvez aller à tout moment, par exemple, d'une résidence secondaire, d'un logement utilisé occasionnellement pour des raisons professionnelles ?

Si oui, inscrire (en clair ou en abrégé) la catégorie de résidence sur les lignes correspondant aux personnes qui logent dans cette autre résidence, même occasionnellement et même si elles n'y sont pas allées depuis le 1er janvier.

7. Depuis le 1er janvier, M. (citer chaque personne de la liste ci-contre) a-t-il vécu dans un autre logement : dans un logement indépendant ou chez quelqu'un qui l'hébergeait, par exemple, chez des amis, de la famille, chez un employeur ?

Si oui pour une personne, inscrire en clair la catégorie de résidence sur la ligne correspondante.
8. Depuis le 1er janvier, l'une des personnes citées a-t-elle logé dans :
- une caserne ?
- un internat, une cité universitaire ?
- une pension ?
- un foyer de travailleurs ?
- ou une autre collectivité ?
Si oui pour une personne, inscrire en clair la catégorie de collectivité sur la ligne correspondante.

9. Depuis le 1er janvier, l'une des personnes citées a-t-elle effectué un séjour à l'hôpital ou dans un autre établissement de soins pour une durée d'au moins un mois ?
- oui
- non

a) faire une croix sur la case correspondante.

b) à quelle date, cette(s) personne(s) a-t-elle (ont-elles) été hospitalisé(s) ?

Inscrire ci-dessous les dates, même approximatives.

C) pour les personnes hospitalisées fin janvier

Inscrire en clair la catégorie de collectivité sur les lignes correspondantes.
1ère AUTRE RÉSIDENCE

1. Logement ordinaire
   Collectivité

2. ADRESSE :
   Nom de l'occupant ou de la collectivité ________________________________
   _________________________________________________________________
   N° __________ Rue (ou lieu dit) _________________________________
   _________________________________________________________________
   Commune __________________________ Département ________________
   (Pour Paris, Lyon ou Marseille, essayer de préciser l'arrondissement)

3. CATÉGORIE DE RÉSIDENCE ET DURÉE DE SÉJOUR

<table>
<thead>
<tr>
<th>N° individuel</th>
<th>Cocher le NOI de la ou des personnes logeant ou ayant logé dans cette résidence</th>
<th>S'il s'agit d'un logement ordinaire</th>
<th>Si 3 ou 4 en colonne b</th>
<th>Si 1, 3, 4 en colonne c</th>
<th>M. habite-t-il encore dans ce logement actuellement ?</th>
<th>Catégorie de logement ordinaire cf. carte Code A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M. loge-t-il (ou logeait-il) dans ce logement la plus grande partie de l'année ?</td>
<td>Précisions sur la durée de séjour dans le logement depuis un an :</td>
<td>M. habitait-t-il encore dans ce logement avant l'année indiquée ?</td>
<td>Catégorie de logement ordinaire cf. carte Code A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. oui __________ ou ce logement est _ voyageur personnel _</td>
<td>par exemple, nombre de mois depuis un an, ou nombre approximatif de jours par mois</td>
<td>1. oui __________</td>
<td>Catégorie de logement ordinaire cf. carte Code A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. non : y loge (logeait) de temps à autre, occasionnellement</td>
<td>ou nombre approximatif de jours par mois</td>
<td>2. non : y logeait (logeait) pendant une partie de l'année</td>
<td>Catégorie de logement ordinaire cf. carte Code A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. non : y est (était) hôtelier provisoirement</td>
<td>M. habitait-il encore dans ce logement avant l'année indiquée ?</td>
<td>3. non __________</td>
<td>Catégorie de logement ordinaire cf. carte Code A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. non : (autre cas)</td>
<td>ou nombre approximatif de jours par mois</td>
<td>4. non __________</td>
<td>Catégorie de logement ordinaire cf. carte Code A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) (b) (c) (d) (e)

<table>
<thead>
<tr>
<th>NOI</th>
<th>LGPAB</th>
<th>DEB</th>
<th>CALO</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Réservé à la DR (aux colibristes)

APPARIEMENT A L'ADRESSE INDIQUÉE CI-CONTRE

1. Résultat de l'appariement des logements
   - Logement recensé
   - Logement non retrouvé au RP
     - car vraisemblablement oublié au recensement
     - car l'adresse ci-contre est imprécise ou fausse

2. Si l'adresse ci-dessus, et si le logement recensé est un logement ordinaire, catégorie du logement au recensement :
   - Résidence principale
   - Logement utilisé occasionnellement
   - Résidence secondaire
   - Logement vacant
   - Caravane, habitation mobile

<table>
<thead>
<tr>
<th>S'il s'agit d'une collectivité</th>
<th>RESERVE À LA DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. loge-t-il encore dans cet établissement (caserne, internat, etc.) ?</td>
<td>Category of collectivity</td>
</tr>
<tr>
<td>1. oui</td>
<td>cl. carte</td>
</tr>
<tr>
<td>2. non : il a quitté définitivement (après le 1er janvier)</td>
<td>Code E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEC</th>
<th>MDEC</th>
<th>CACOL</th>
<th>RAB</th>
<th>CABI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

La collecte
BULLETIN INDIVIDUEL

1. Numéro individuel (N01) ........................................ [0.1]

2. SI recopié du RP (chiffrer 1) [ ]

3. NOM, Prénoms
Écrivez le nom en capitales (exemple : ALLARD, épouse MAURIN, Françoise).

4. SEXE
Masculin [ ]
Féminin [ ]

5. DATE DE NAISSANCE

<table>
<thead>
<tr>
<th>Jour</th>
<th>Mois</th>
<th>Année</th>
</tr>
</thead>
</table>

6. ÉTAT MATRIMONIAL LÉGAUX
Une personne vivant en union libre coche le case correspondant à sa situation juridique préavus : si, par exemple, elle est célibataire, elle coche la 1ère case.
Une personne en instance de divorce ou séparée de son conjoint sépare coche la 2ème case.

7. NATIONALITÉ
« Français de naissance (y compris par naturalisation) » ............................................... [1]
« Devenu français par naturalisation, mariage, déclaration ou à votre majorité » ....................... [2]
Indiquez votre nationalité antérieure :

<table>
<thead>
<tr>
<th>Étranger</th>
</tr>
</thead>
</table>

8. INDIQUEZ VOTRE DIPLOMÉE LE PLUS ÉLEVÉ

- Certificat d'études primaires ........................................ [0]
- BEPC, brevet élémentaire, brevet des collèges .................. [0]
- CAP ........................................................................... [0]
- BEP ........................................................................... [0]
- BEP, brevet professionnel ou de technicien, autre brevet (BEP,BEC,BLE, etc.) ............... [0]
- Diplôme universitaire de 1er cycle, BTS, DUT, diplôme des professions libérales ou de la santé .............................................................. [0]
- Diplôme universitaire de 2e ou 3e cycle, diplôme d'ingénieur, d'une grande école, etc. ......... [0]
- Aucun diplôme ............................................................ [0]

9. ÉTES-VOUS ?
- Élève, étudiant, stagiaire non rémunéré ........................................ [0]
- Chômeur (brut ou net à l'ANPE) ....................................... [0]
- Retraité (ancien salarié) .................................................. [0]
- Retraité des affaires (ancien agriculteur, ancien artisan, ancien commerçant, etc.) .......... [0]
- Femme au foyer ........................................................... [0]
- Autre inactivité (y compris les personnes ne percevant qu'une pension de réversion) ..... [0]

10. AVEZ-VOUS DÉJÀ TRAVAILLÉ ?
OUI [1] — Quelle était votre profession principale ?
NON [2] .................................................................

11. CHERCHEZ-VOUS UN EMPLOI ?
- Vous ne cherchez pas d'emploi ........................................ [0]
- Vous cherchez un emploi depuis ........................................................................ [0]
- Moins de 3 mois .................................................................. [0]
- 3 mois à moins de 1 an .................................................... [0]
- 1 an à moins de 2 ans ..................................................... [0]
- 2 ans ou plus .................................................................. [0]

La collecte 141
12. **PROFESSION EXERCÉE ACTUELLEMENT**

Soyez très précis. Exemples : ouvrier électricien d'entou-

der, chef cuisinier, vendeur en electromé-

ner, agent de transport, casseur de libre-service, etc.

13. **AIDEZ-VOUS UN MEMBRE DE VOTRE FAMILLE DANS SON TRAVAIL ?**

| Exploitation agricole ou artisanale, non profession libérale | OUI | NON |

14. **OU TRAVAILLEZ-VOUS ?**

| Rue ou lieu-dit | Commune | Département |

15. **INDIQUEZ LA POSITION PROFESSIONNELLE DE VOTRE EMPLOI ACTUEL :**

<table>
<thead>
<tr>
<th>Position professionnelle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manutention, magasinage, transports</td>
</tr>
<tr>
<td>Secrétaire, secrétaire général, assistante administrative</td>
</tr>
<tr>
<td>Gestion, comptabilité, fonctions administratives</td>
</tr>
<tr>
<td>Communication, vente technique-commerciale</td>
</tr>
<tr>
<td>Renseignements, études, méthodes, informatique</td>
</tr>
<tr>
<td>Directeur général ou un de ses adjoints directs</td>
</tr>
<tr>
<td>Autre Case</td>
</tr>
</tbody>
</table>

**ETES-VOUS ?**

| Salarisé | ouvrier spécialisé | ouvrier spécialisé, artisan, commerçant, agriculteur, professionnel libéral, pour non salariés, etc. |

17. **TRAVAILLEZ-VOUS ?**

| à temps partiel | à temps complet |

18. **SI VOUS ETES A VOTRE COMPTE**

| Combien de salariés employés-employés-employés-employés | 1 |
| 1 ou 2 | 2 |
| 3 ou 4 | 3 |
| 5 ou plus | 4 |

**INDICEZ LES QUESTIONS 19 A 21 NE S'ADRESSENT QU'ÀUX SALARIÉS ET AUX STAGIAIRES RÉMUNÉRÉS.**

19. **ETES-VOUS DANS L'UNE DES SITUATIONS SUIVANTES ?**

| Salarisé sous contrat | 1 |
| Exerçant un travail d'utilité collective (TUC, etc.) | 2 |
| Sous contrat d'adaptation ou d'qualification | 3 |
| Stagiaire (principalement en entreprise : SIVP, etc.) | 4 |
| Stagiaire (principalement dans un centre de formation : PPA, etc.) | 5 |
| Placement par une agence d'interim | 6 |
| Sous contrat de travail à durée déterminée | 7 |

20. **QU'EST VOTRE FONCTION PRINCIPALE ?**

| Production, fabrication, chantier | 1 |
| Installation, entretien, réglage, réparation | 2 |
| Maintien, gestion, travail managéry | 3 |
| Manutention, magasinage, transports | 4 |
| Secrétaire, secrétaire, secrétaire général, assistant(e) | 5 |
| Gestion, comptabilité, fonctions administratives | 6 |
| Commerce, vente technique-commerciale | 7 |
| Renseignements, études, méthodes, informatique | 8 |
| Directeur général ou un de ses adjoints directs | 9 |
| Autre Case | 10 |

**LES QUESTIONS 19 À 21 NE S'ADRESSENT QU'ÀUX SALARIÉS ET AUX STAGIAIRES RÉMUNÉRÉS.**

**INSEE Méthodes n° 52 - 53**
RÉSERVÉ À LA DR
A REMPLIR PAR LA DR
ÉVENTUELLEMENT CONSIGNES À L'ENQUÉTEUR EN VUE DE LA 2E VAGUE D'ENQUÊTE

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A REMPLIR PAR L'ENQUÊTEUR

RÉSULTATS DE LA 2e VAGUE D'ENQUÊTE

(l'enquêteur doit notamment préciser toutes les modifications qu'il a apportées au questionnaire, suite à cette deuxième vague d'enquête)

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*La collecte*
**RÉSERVÉ À LA DR**

**APPARIEMENT DES INDIVIDUS À L'ADRESSE DES LOGEMENTS APPARTENANT À L'ÉCHANTILLON**

<table>
<thead>
<tr>
<th>RESULTAT DE L'APPARIEMENT</th>
<th>N° du logement correspondant à la résidence principale de la personne : 0 pour le logement enquêté, 1, etc. pour les autres résidences</th>
<th>OBSERVATIONS EVENTUELLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si M. est enquêté</td>
<td>Si M. est recensé et non enquêté</td>
<td></td>
</tr>
<tr>
<td>Résultat de l'appariement avec le RP</td>
<td>Si &quot;00 non recense&quot; en colonne précédente, raison pour laquelle M. n'a pas été recensée</td>
<td></td>
</tr>
<tr>
<td>cf. code C</td>
<td>cf. code D</td>
<td>Chiffret G</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>NOI</td>
<td>RA</td>
<td>RAINOR</td>
</tr>
</tbody>
</table>

| (a1) | | | | | |
| (a2) | | | | | |
| (a3) | | | | | |
| (a4) | | | | | |
| (a5) | | | | | |
| (a6) | | | | | |
| (a7) | | | | | |
| (a8) | | | | | |
| (a9) | | | | | |
| (a10) | | | | | |

* et recopier BI du recensement
**LPRM - Lien avec la personne de référence du ménage**

1. Personne de référence.
2. Conjoint de la personne de référence, marié ou de fait (la femme dans un couple).
5. Ascendant de la personne de référence ou de son conjoint: mère, père, beau-père, belle-mère, grand-père, grand-mère, etc.
6. Autre parent de la personne de référence ou de son conjoint: frère, sœur, neveu, nièce, cousin, cousine, beau-père, belle-mère, oncle, tante, etc.
7. Ami.
8.1. Pensionnaire, enfant en nourrice (sans lien de parenté avec la personne de référence ou son conjoint).
8.2. Étudiant logé (sans lien de parenté avec la personne de référence ou son conjoint).
8.3. Sous-locataire, locuteur.

**DÉTERMINATION DE LA PERSONNE DE RÉFÉRENCE DANS LE LOGEMENT**

| A | Il existe un ou plusieurs couples mariés ou de fait, avec ou sans enfants 1 couple | Personne référence = Le conjoint masculin marié ou de fait  
Personne référence = Permis les conjoints masculins mariés ou de fait, l'actif le plus âgé ou, à défaut d'actif, le plus âgé |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2 ou plusieurs couples</td>
<td>Personne référence = La personne active le plus âgé</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>À défaut de couple, il existe une famille monoparentale: une mère (ou un père) sans conjoint avec des enfants célibataires, ces enfants étant eux-mêmes sans conjoint et sans enfants 1 famille monoparentale</td>
<td>Personne référence = La mère (ou le père) des enfants</td>
</tr>
<tr>
<td>2 ou plusieurs familles monoparentales</td>
<td>Personne référence = Permis les mères (ou pères) des enfants, la personne active le plus âgé ou, à défaut de personne active, la personne la plus âgée</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>À défaut de couples et à défaut de famille monoparentale, toutes les personnes du logement sont considérées comme des personnes &quot;isolées&quot;</td>
<td>Personne référence = Permis toutes les personnes &quot;isolées&quot;, la personne active le plus âgé ou, à défaut de personne active, la personne la plus âgée</td>
</tr>
</tbody>
</table>

*Y compris les chômeurs

**Cas particuliers:**
- Les salariés logés chez l'employeur, même s'ils constituent un couple ou une famille monoparentale, ne seront jamais retenus comme personnes de référence. Choisir la personne de référence parmi les autres couple(s) ou autre(s) famille(s) monoparentale(s) ou autre(s) personne(s) "isolée(s)" du ménage.
- L'ascendant (grand-père ou grand-mère) sans conjoint élevant seul des petits-enfants ne constitue pas une famille monoparentale. Dans ce cas, l'ascendant et ses petits-enfants seront assimilés à des "isolés".
- En aucun cas, un enfant de moins de 15 ans ne peut être retenu comme personne de référence.
CODE A: CATÉGORIE DE LOGEMENT (ORDINAIRE)

pour la personne enquêtée

CAS PARTICULIERS

• Étudiant ayant une chambre en ville, ou logeant chez l’habitant. Classer son logement en ........................................ 11
• Salarié logé chez son employeur (les deux vivent dans le même logement): par exemple, jeune fille au pair, domestique, salarié agricole, apprenti logé chez son patron. Classer son logement chez l’employeur en ....... 12
• Enfant placé en pension ou en nourrice nuit et jour chez des particuliers. Classer son logement chez ces particuliers en ............................................................... 13

— Résidence principale
• Si la personne dispose de plusieurs résidences, résidence où elle vit la plus grande partie de l’année ou du moins celle où elle a vécu le plus longtemps depuis un an ................................................................. 14
• Si aux dates où la personne habitait (habitait) le logement, ce logement était (est) sa seule résidence ...... 15

— Logement utilisé occasionnellement pour des raisons professionnelles par une personne qui a sa résidence principale ailleurs ................................................................. 21
— Résidence secondaire (la personne y passe moins de temps que dans sa résidence principale) .......... 22
— Hébergement temporaire ou occasionnel chez des membres de la famille ........................................... 23
— Hébergement temporaire ou occasionnel chez des amis ........................................................................ 24
— Autre cas (éviter d’utiliser cette rubrique; préciser sur le questionnaire de quel type de résidence il s’agit) : 25

CODE B: CATÉGORIE DE COLLECTIVITÉ

• Caserne ................................................................. 1
• Internat, pension pour élèves .......................... 2
• Cité universitaire, foyer pour étudiants .............. 3
• Foyer de travailleurs ........................................... 4
• Hôpital, clinique, maternité, autre établissement de soins ... 5
• Maison de retraite, hospice ............................. 6
• Autre catégorie de collectivité (prison, etc.) .......... 7
2.4 - ÉLÉMENTS SUR LES TAUX DE NON RÉPONSE

Jean-Michel Durr

L'exploitation des recensements de la population met en œuvre des traitements visant à redresser les valeurs non déclarées ou incohérentes. La méthode la plus fréquemment utilisée est celle du hot-deck, qui consiste à attribuer la valeur de l'individu (ou du logement) précédent. Dans certains cas, on prend le précédent individu ayant certaines caractéristiques communes avec l'individu à redresser, le hot-deck est alors dit stratifié. Cette procédure, simple à mettre en œuvre surtout dans le cas d'exploitations de l'ampleur de celles d'un recensement, se justifie également statistiquement. En effet, le recensement procède géographiquement, par district, et les logements et leurs occupants proches sur le terrain sont également proches dans le fichier. Or, un certain nombre de caractéristiques, principalement concernant les logements, sont corrélées géographiquement. Si un logement possède un certain niveau de confort, les logements voisins ont souvent le même niveau. L'effet est naturellement beaucoup moins fort pour les individus, à l'exception de critères tels que la nationalité. Le hot-deck présente également l'intérêt de redresser au prorata et donc de ne pas perturber la structure. Ceci suppose une répartition uniforme des non-répondants. D'autres méthodes sont envisageables ou le deviendront à mesure des progrès de l'informatique.

Les tableaux ci-après présentent les taux de non déclaration des différentes variables du recensement. Lorsque la comparaison était possible, le taux correspondant du recensement de 82 a été rappelé. On constate une certaine stabilité des taux de non réponse, avec pour les logements une amélioration du taux de réponse pour les résidences principales (l'information est alors fournie par l'occupant) et une dégradation pour les logements vides (l'information doit être recueillie par l'agent recenseur). La comparaison des taux des variables individuelles est légèrement perturbée par le traitement effectué en 82 sur les bulletins considérés comme blancs. En effet, dès lors que l'année de naissance était non déclarée et que le type d'activité n'avait pu être chiffré, l'enregistrement initial était entièrement écrasé par celui d'un individu précédemment traité (et donc entièrement redressé). En supposant que les bulletins dits blancs en 82 aient été entièrement vierges, il faudrait ajouter aux différents chiffres de 82 l'effectif de ces 180 636 bulletins.

Le dernier tableau présente la répartition par département des bulletins "blancs" au sens de 82. Le nombre total de ces bulletins est passé à 222 176.

(1) Ce qui n'est pas le cas si la question est nouvelle, si elle avait été traitée à l'exhaustif au RP82...
### Tableau 1

**Compteurs de non déclaration - RP 90 sondage au 1/4 (y.c. panel)**

**France métropolitaine**

<table>
<thead>
<tr>
<th></th>
<th>État des logements</th>
<th>RP90</th>
<th>Dénominateur</th>
<th>%</th>
<th>% au RP82</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IMMEUBLES (TOTIMM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Époque d'achèvement</td>
<td>1 175 519</td>
<td>TOTIMM</td>
<td>7,14</td>
<td>7,30</td>
<td></td>
</tr>
<tr>
<td>Appartenance HLM</td>
<td>2 219 793</td>
<td>TOTIMM</td>
<td>13,49</td>
<td>(1) 8,51</td>
<td></td>
</tr>
<tr>
<td>Combustible</td>
<td>1 787 188</td>
<td>TOTIMM</td>
<td>10,86</td>
<td>(1) 6,59</td>
<td></td>
</tr>
<tr>
<td><strong>LOGEMENTS (TOTLOG)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logements occupés (LOGOCC)</td>
<td>26 462 279</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logements vides (LOGVID)</td>
<td>21 759 227</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catégorie</td>
<td>4 703 052</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>364 828</td>
<td>LOGOCC</td>
<td>1,68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>597 432</td>
<td>LOGVID</td>
<td>12,70</td>
<td></td>
</tr>
<tr>
<td>Nombre de pièces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>287 760</td>
<td>LOGOCC</td>
<td>1,32</td>
<td>1,80</td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>1 433 404</td>
<td>LOGVID</td>
<td>30,48</td>
<td>21,50</td>
</tr>
<tr>
<td>Statut d'occupation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occuèps</td>
<td>405 364</td>
<td>LOGOCC</td>
<td>1,86</td>
<td>2,64</td>
</tr>
<tr>
<td>Appartenance HLM</td>
<td>3 119 240</td>
<td>TOTLOG</td>
<td>11,79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Installations sanitaires</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>323 016</td>
<td>LOGOCC</td>
<td>1,48</td>
<td>2,52</td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>1 438 204</td>
<td>LOGVID</td>
<td>30,58</td>
<td>21,89</td>
</tr>
<tr>
<td>W.C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>559 428</td>
<td>LOGOCC</td>
<td>2,57</td>
<td>1,80</td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>1 259 792</td>
<td>LOGVID</td>
<td>26,79</td>
<td>19,50</td>
</tr>
<tr>
<td>Mode de chauffage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>400 288</td>
<td>LOGOCC</td>
<td>1,84</td>
<td>2,60</td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>1 637 796</td>
<td>LOGVID</td>
<td>34,82</td>
<td>27,30</td>
</tr>
<tr>
<td>Combustible</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>occupés</td>
<td>484 508</td>
<td>LOGOCC</td>
<td>2,23</td>
<td>3,74</td>
</tr>
<tr>
<td></td>
<td>vides</td>
<td>1 739 948</td>
<td>LOGVID</td>
<td>37,00</td>
<td>26,65</td>
</tr>
<tr>
<td>Nombre de voitures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>673 776</td>
<td>LOGOCC</td>
<td>3,10</td>
<td>4,40</td>
<td></td>
</tr>
<tr>
<td>Exploitation agricole</td>
<td>2 812 240</td>
<td>TOTLOG</td>
<td>10,63</td>
<td>(1) 23,13</td>
<td></td>
</tr>
<tr>
<td>Superficie agricole utilisée</td>
<td>70 424</td>
<td>LOGOCC</td>
<td>0,32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orientation des productions agricoles</td>
<td>94 540</td>
<td>LOGOCC</td>
<td>0,43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1) En 1982 le bordereau de maison (désormais limité aux immeubles collectifs) était généralisé. Or le taux de non réponse est nettement différent pour une variable selon qu'elle apparaît sur le dossier d'immeubles collectifs (ou sur le bordereau de maison en 1982) ou sur la feuille de logement.
<table>
<thead>
<tr>
<th>Élément</th>
<th>RP90</th>
<th>Dénominateur</th>
<th>%</th>
<th>% au RP82</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INDIVIDUS (TOTIND)</strong></td>
<td>57 152 019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individus de plus de 14 ans (PG14IND)</td>
<td>46 226 102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actifs ayant un emploi (ACTEMP)</td>
<td>22 554 559</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulletins blancs (1)</td>
<td>222 176</td>
<td>TOTIND</td>
<td>0,39</td>
<td>0,31</td>
</tr>
<tr>
<td>Jour de naissance</td>
<td>621 271</td>
<td>TOTIND</td>
<td>1,09</td>
<td>1,00</td>
</tr>
<tr>
<td>Mois de naissance</td>
<td>599 827</td>
<td>TOTIND</td>
<td>1,05</td>
<td>1,22</td>
</tr>
<tr>
<td>Année de naissance</td>
<td>337 666</td>
<td>TOTIND</td>
<td>0,59</td>
<td>0,20</td>
</tr>
<tr>
<td>état matrimonial (yc incohérent)</td>
<td>1 209 329</td>
<td>TOTIND</td>
<td>2,12</td>
<td>1,16</td>
</tr>
<tr>
<td>Indicateur de lieu de naissance</td>
<td>520 378</td>
<td>TOTIND</td>
<td>0,91</td>
<td>0,70</td>
</tr>
<tr>
<td>Indicateur de nationalité</td>
<td>488 038</td>
<td>TOTIND</td>
<td>0,85</td>
<td>0,60</td>
</tr>
<tr>
<td>Nationalité détaillée</td>
<td>710 003</td>
<td>TOTIND</td>
<td>1,24</td>
<td>1,70</td>
</tr>
<tr>
<td>Diplôme (2)</td>
<td>13 808 670</td>
<td>PG14IND</td>
<td>29,87</td>
<td></td>
</tr>
<tr>
<td>Résidence antérieure</td>
<td>1 285 417</td>
<td>TOTIND</td>
<td>2,25</td>
<td>2,09</td>
</tr>
<tr>
<td>Type d'activité</td>
<td>902 660</td>
<td>PG14IND</td>
<td>1,95</td>
<td></td>
</tr>
<tr>
<td>Profession</td>
<td>500 108</td>
<td>ACTEMP</td>
<td>2,22</td>
<td></td>
</tr>
<tr>
<td>Activité économique</td>
<td>889 570</td>
<td>ACTEMP</td>
<td>3,94</td>
<td>2,50</td>
</tr>
<tr>
<td>Statut</td>
<td>683 161</td>
<td>ACTEMP</td>
<td>3,03</td>
<td>1,00</td>
</tr>
<tr>
<td>Temps de travail</td>
<td>1 350 433</td>
<td>ACTEMP</td>
<td>5,99</td>
<td></td>
</tr>
<tr>
<td>Commune de lieu de travail</td>
<td>922 697</td>
<td>ACTEMP</td>
<td>4,09</td>
<td>1,85</td>
</tr>
<tr>
<td>Département de lieu de travail</td>
<td>876 336</td>
<td>ACTEMP</td>
<td>3,89</td>
<td>1,80</td>
</tr>
<tr>
<td>Arrondissement municipal de LT</td>
<td>68 029</td>
<td>ACTEMP</td>
<td>0,30</td>
<td></td>
</tr>
<tr>
<td>Position professionnelle</td>
<td>4 269 306</td>
<td>ACTEMP</td>
<td>18,93</td>
<td></td>
</tr>
</tbody>
</table>

(1) Les bulletins blancs sont estimés par le nombre de bulletins pour lesquels l'année de naissance et le type d'activité sont non déclarés. En 1990, aucun traitement spécifique n'a été réalisé pour ces bulletins, ils sont donc inclus dans les différents compteurs de non déclaration. En revanche, pour le RP82, on les redressait en bloc avant le redressement variable par variable (ce qui revient à peu près au même). Ils ne sont donc pas pris en compte dans les valeurs des compteurs de 82.

(2) La présentation de la question ne permet pas de distinguer entre absence de diplôme et non réponse. Voir chapitre : étude sur la question sur la formation.
<table>
<thead>
<tr>
<th>Département</th>
<th>Population</th>
<th>Bulletins blancs</th>
<th>%</th>
<th>Rang</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 AIN</td>
<td>471 019</td>
<td>388</td>
<td>0,1</td>
<td>57</td>
</tr>
<tr>
<td>02 AISNE</td>
<td>537 259</td>
<td>468</td>
<td>0,1</td>
<td>56</td>
</tr>
<tr>
<td>03 ALLIER</td>
<td>357 710</td>
<td>292</td>
<td>0,1</td>
<td>58</td>
</tr>
<tr>
<td>04 ALPES-DE-HAUTE-PROVENCE</td>
<td>130 883</td>
<td>260</td>
<td>0,2</td>
<td>29</td>
</tr>
<tr>
<td>05 HAUTES-ALPES</td>
<td>113 300</td>
<td>160</td>
<td>0,1</td>
<td>38</td>
</tr>
<tr>
<td>06 ALPES-MARITIMES</td>
<td>971 829</td>
<td>7 964</td>
<td>0,8</td>
<td>6</td>
</tr>
<tr>
<td>07 ARDÈCHE</td>
<td>277 581</td>
<td>444</td>
<td>0,2</td>
<td>34</td>
</tr>
<tr>
<td>08 ARDENNES</td>
<td>296 357</td>
<td>164</td>
<td>0,1</td>
<td>73</td>
</tr>
<tr>
<td>09 ARIÈGE</td>
<td>136 455</td>
<td>80</td>
<td>0,1</td>
<td>70</td>
</tr>
<tr>
<td>10 AUBE</td>
<td>289 207</td>
<td>300</td>
<td>0,1</td>
<td>48</td>
</tr>
<tr>
<td>11 AUDE</td>
<td>298 712</td>
<td>140</td>
<td>0,0</td>
<td>78</td>
</tr>
<tr>
<td>12 AVEYRON</td>
<td>270 141</td>
<td>180</td>
<td>0,1</td>
<td>64</td>
</tr>
<tr>
<td>13 BOUCHES-DU-RHÔNE</td>
<td>1 759 371</td>
<td>19 216</td>
<td>1,1</td>
<td>3</td>
</tr>
<tr>
<td>14 CALVADOS</td>
<td>618 478</td>
<td>392</td>
<td>0,1</td>
<td>69</td>
</tr>
<tr>
<td>15 CANTAL</td>
<td>158 723</td>
<td>240</td>
<td>0,2</td>
<td>36</td>
</tr>
<tr>
<td>16 CHARENTE</td>
<td>341 993</td>
<td>304</td>
<td>0,1</td>
<td>54</td>
</tr>
<tr>
<td>17 CHARENTE-MARITIME</td>
<td>527 146</td>
<td>544</td>
<td>0,1</td>
<td>49</td>
</tr>
<tr>
<td>18 CHER</td>
<td>321 559</td>
<td>184</td>
<td>0,1</td>
<td>72</td>
</tr>
<tr>
<td>19 CORRÈZE</td>
<td>237 908</td>
<td>100</td>
<td>0,0</td>
<td>83</td>
</tr>
<tr>
<td>2A CORSE-DU-SUD</td>
<td>118 174</td>
<td>244</td>
<td>0,2</td>
<td>27</td>
</tr>
<tr>
<td>2B HAUTE-CORSE</td>
<td>131 563</td>
<td>1 144</td>
<td>0,9</td>
<td>5</td>
</tr>
<tr>
<td>21 CÔTE-D'OR</td>
<td>493 866</td>
<td>680</td>
<td>0,1</td>
<td>40</td>
</tr>
<tr>
<td>22 CÔTES-D'ARMOR</td>
<td>538 395</td>
<td>484</td>
<td>0,1</td>
<td>52</td>
</tr>
<tr>
<td>23 CREUSE</td>
<td>131 349</td>
<td>60</td>
<td>0,0</td>
<td>79</td>
</tr>
<tr>
<td>24 DORDOGNE</td>
<td>386 365</td>
<td>436</td>
<td>0,1</td>
<td>45</td>
</tr>
<tr>
<td>25 DOUBS</td>
<td>484 770</td>
<td>368</td>
<td>0,1</td>
<td>60</td>
</tr>
<tr>
<td>26 DRÔME</td>
<td>414 072</td>
<td>1 100</td>
<td>0,3</td>
<td>22</td>
</tr>
<tr>
<td>27 EURE</td>
<td>513 818</td>
<td>332</td>
<td>0,1</td>
<td>66</td>
</tr>
<tr>
<td>28 EURE-ET-LOIR</td>
<td>396 073</td>
<td>120</td>
<td>0,0</td>
<td>88</td>
</tr>
<tr>
<td>29 FINISTÈRE</td>
<td>838 687</td>
<td>548</td>
<td>0,1</td>
<td>65</td>
</tr>
<tr>
<td>30 GARD</td>
<td>585 049</td>
<td>1 752</td>
<td>0,3</td>
<td>21</td>
</tr>
<tr>
<td>31 HAUTE-GARONNE</td>
<td>925 962</td>
<td>4 724</td>
<td>0,5</td>
<td>10</td>
</tr>
<tr>
<td>32 GERS</td>
<td>174 587</td>
<td>264</td>
<td>0,2</td>
<td>35</td>
</tr>
</tbody>
</table>
Tableau 2 (Suite)

<table>
<thead>
<tr>
<th>Département</th>
<th>Population</th>
<th>Bulletins blancs</th>
<th>%</th>
<th>Rang</th>
</tr>
</thead>
<tbody>
<tr>
<td>33 GIRONDE</td>
<td>1 213 499</td>
<td>6 584</td>
<td>0,5</td>
<td>9</td>
</tr>
<tr>
<td>34 HÉRAULT</td>
<td>794 603</td>
<td>3 528</td>
<td>0,4</td>
<td>13</td>
</tr>
<tr>
<td>35 ILLE-ET-VILAINE</td>
<td>798 718</td>
<td>248</td>
<td>0,0</td>
<td>87</td>
</tr>
<tr>
<td>36 INDRE</td>
<td>237 510</td>
<td>80</td>
<td>0,0</td>
<td>86</td>
</tr>
<tr>
<td>37 INDRE-ET-LOIRE</td>
<td>529 345</td>
<td>848</td>
<td>0,2</td>
<td>33</td>
</tr>
<tr>
<td>38 ISÈRE</td>
<td>1 016 229</td>
<td>1 168</td>
<td>0,1</td>
<td>44</td>
</tr>
<tr>
<td>39 JURA</td>
<td>248 759</td>
<td>304</td>
<td>0,1</td>
<td>41</td>
</tr>
<tr>
<td>40 LANDES</td>
<td>311 461</td>
<td>288</td>
<td>0,1</td>
<td>51</td>
</tr>
<tr>
<td>41 LOIR-ET-CHER</td>
<td>305 937</td>
<td>132</td>
<td>0,0</td>
<td>82</td>
</tr>
<tr>
<td>42 LOIRE</td>
<td>746 288</td>
<td>328</td>
<td>0,0</td>
<td>80</td>
</tr>
<tr>
<td>43 HAUTE-LOIRE</td>
<td>206 568</td>
<td>152</td>
<td>0,1</td>
<td>62</td>
</tr>
<tr>
<td>44 LOIRE-ATLANTIQUE</td>
<td>1 052 183</td>
<td>2 216</td>
<td>0,2</td>
<td>25</td>
</tr>
<tr>
<td>45 LOIRET</td>
<td>580 612</td>
<td>668</td>
<td>0,1</td>
<td>42</td>
</tr>
<tr>
<td>46 LOT</td>
<td>155 816</td>
<td>100</td>
<td>0,1</td>
<td>67</td>
</tr>
<tr>
<td>47 LOT-ET-GARONNE</td>
<td>305 989</td>
<td>144</td>
<td>0,0</td>
<td>77</td>
</tr>
<tr>
<td>48 LOZÈRE</td>
<td>72 825</td>
<td>64</td>
<td>0,1</td>
<td>55</td>
</tr>
<tr>
<td>49 MAINE-ET-LOIRE</td>
<td>705 882</td>
<td>1 060</td>
<td>0,2</td>
<td>37</td>
</tr>
<tr>
<td>50 MANCHE</td>
<td>479 636</td>
<td>360</td>
<td>0,1</td>
<td>61</td>
</tr>
<tr>
<td>51 MARNE</td>
<td>558 217</td>
<td>1 712</td>
<td>0,3</td>
<td>20</td>
</tr>
<tr>
<td>52 HAUTE-MARNE</td>
<td>204 067</td>
<td>40</td>
<td>0,0</td>
<td>95</td>
</tr>
<tr>
<td>53 MAYENNE</td>
<td>278 037</td>
<td>84</td>
<td>0,0</td>
<td>89</td>
</tr>
<tr>
<td>54 MEURTHE-ET-MOSELLE</td>
<td>711 822</td>
<td>1 304</td>
<td>0,2</td>
<td>30</td>
</tr>
<tr>
<td>55 MEUSE</td>
<td>196 344</td>
<td>104</td>
<td>0,1</td>
<td>74</td>
</tr>
<tr>
<td>56 MORBIHAN</td>
<td>619 838</td>
<td>180</td>
<td>0,0</td>
<td>90</td>
</tr>
<tr>
<td>57 MOSELLE</td>
<td>1 011 302</td>
<td>900</td>
<td>0,1</td>
<td>53</td>
</tr>
<tr>
<td>58 NIÈVRE</td>
<td>233 278</td>
<td>60</td>
<td>0,0</td>
<td>93</td>
</tr>
<tr>
<td>59 NORD</td>
<td>2 531 855</td>
<td>5 156</td>
<td>0,2</td>
<td>28</td>
</tr>
<tr>
<td>60 OISE</td>
<td>725 603</td>
<td>768</td>
<td>0,1</td>
<td>47</td>
</tr>
<tr>
<td>61 ORNE</td>
<td>293 204</td>
<td>84</td>
<td>0,0</td>
<td>91</td>
</tr>
<tr>
<td>62 PAS-DE-CALAIS</td>
<td>1 433 203</td>
<td>748</td>
<td>0,1</td>
<td>75</td>
</tr>
<tr>
<td>63 PUY-DE-DOME</td>
<td>596 213</td>
<td>1 448</td>
<td>0,2</td>
<td>23</td>
</tr>
<tr>
<td>64 PYRÉNÉES-ATLANTIQUES</td>
<td>578 516</td>
<td>420</td>
<td>0,1</td>
<td>63</td>
</tr>
<tr>
<td>65 HAUTES-PYRÉNÉES</td>
<td>224 759</td>
<td>80</td>
<td>0,0</td>
<td>85</td>
</tr>
<tr>
<td>66 PYRÉNÉES-ORIENTALES</td>
<td>363 796</td>
<td>624</td>
<td>0,2</td>
<td>31</td>
</tr>
</tbody>
</table>

La collecte
<table>
<thead>
<tr>
<th>Département</th>
<th>Population</th>
<th>Bulletins blancs</th>
<th>%</th>
<th>Rang</th>
</tr>
</thead>
<tbody>
<tr>
<td>67 BAS-RHIN</td>
<td>953 053</td>
<td>1 096</td>
<td>0,1</td>
<td>43</td>
</tr>
<tr>
<td>68 HAUT-RHIN</td>
<td>671 319</td>
<td>1 408</td>
<td>0,2</td>
<td>26</td>
</tr>
<tr>
<td>69 RHÔNE</td>
<td>1 508 966</td>
<td>7 140</td>
<td>0,5</td>
<td>11</td>
</tr>
<tr>
<td>70 HAUTE-SAÔNE</td>
<td>229 650</td>
<td>60</td>
<td>0,0</td>
<td>92</td>
</tr>
<tr>
<td>71 SAÔNE-ET-LOIRE</td>
<td>559 413</td>
<td>48</td>
<td>0,0</td>
<td>96</td>
</tr>
<tr>
<td>72 SARTHE</td>
<td>513 654</td>
<td>120</td>
<td>0,0</td>
<td>94</td>
</tr>
<tr>
<td>73 SAVOIE</td>
<td>348 261</td>
<td>284</td>
<td>0,1</td>
<td>59</td>
</tr>
<tr>
<td>74 HAUTE-SAVOIE</td>
<td>568 286</td>
<td>584</td>
<td>0,1</td>
<td>50</td>
</tr>
<tr>
<td>75 PARIS</td>
<td>2 152 423</td>
<td>65 504</td>
<td>3,0</td>
<td>1</td>
</tr>
<tr>
<td>76 SEINE-MARITIME</td>
<td>1 223 429</td>
<td>3 860</td>
<td>0,3</td>
<td>19</td>
</tr>
<tr>
<td>77 SEINE-ET-MARNE</td>
<td>1 078 166</td>
<td>3 440</td>
<td>0,3</td>
<td>18</td>
</tr>
<tr>
<td>78 YVEILINES</td>
<td>1 307 150</td>
<td>7 192</td>
<td>0,6</td>
<td>8</td>
</tr>
<tr>
<td>79 DEUX-SÈVRES</td>
<td>345 965</td>
<td>220</td>
<td>0,1</td>
<td>68</td>
</tr>
<tr>
<td>80 SOMME</td>
<td>547 825</td>
<td>1 800</td>
<td>0,3</td>
<td>17</td>
</tr>
<tr>
<td>81 TARN</td>
<td>342 723</td>
<td>200</td>
<td>0,1</td>
<td>71</td>
</tr>
<tr>
<td>82 TARN-ET-GARONNE</td>
<td>200 220</td>
<td>80</td>
<td>0,0</td>
<td>84</td>
</tr>
<tr>
<td>83 VAR</td>
<td>815 449</td>
<td>2 984</td>
<td>0,4</td>
<td>16</td>
</tr>
<tr>
<td>84 VAUCLUSE</td>
<td>467 075</td>
<td>1 004</td>
<td>0,2</td>
<td>24</td>
</tr>
<tr>
<td>85 VENDÉE</td>
<td>509 356</td>
<td>220</td>
<td>0,0</td>
<td>81</td>
</tr>
<tr>
<td>86 VIENNE</td>
<td>379 977</td>
<td>624</td>
<td>0,2</td>
<td>32</td>
</tr>
<tr>
<td>87 HAUTE-VIENNE</td>
<td>353 593</td>
<td>492</td>
<td>0,1</td>
<td>39</td>
</tr>
<tr>
<td>88 VOSGES</td>
<td>386 258</td>
<td>200</td>
<td>0,1</td>
<td>76</td>
</tr>
<tr>
<td>89 YONNE</td>
<td>323 096</td>
<td>360</td>
<td>0,1</td>
<td>46</td>
</tr>
<tr>
<td>90 TERRITOIRE-DE-BELFORT</td>
<td>134 097</td>
<td>0</td>
<td>0,0</td>
<td>97</td>
</tr>
<tr>
<td>91 ESSONNE</td>
<td>1 084 824</td>
<td>4 964</td>
<td>0,5</td>
<td>12</td>
</tr>
<tr>
<td>92 HAUTS-DE-SEINE</td>
<td>1 391 658</td>
<td>17 692</td>
<td>1,3</td>
<td>2</td>
</tr>
<tr>
<td>93 SEINE-SAINT-DENIS</td>
<td>1 381 197</td>
<td>12 104</td>
<td>0,9</td>
<td>4</td>
</tr>
<tr>
<td>94 VAL-DE-MARNE</td>
<td>1 215 538</td>
<td>7 772</td>
<td>0,6</td>
<td>7</td>
</tr>
<tr>
<td>95 VAL-D'OISE</td>
<td>1 049 598</td>
<td>4 644</td>
<td>0,4</td>
<td>14</td>
</tr>
<tr>
<td>FRANCE MÉTROPOLITAINE</td>
<td>56 614 493</td>
<td>222 176</td>
<td>0,4</td>
<td>15</td>
</tr>
</tbody>
</table>

Notes :

les bulletins individuels sont considérés comme blancs dès lors que l’année de naissance est non déclarée et que le type d’activité n’a pu être chiffré.

Le rang est calculé par ordre décroissant.
2.5 - LES ÉCARTS D’ESTIMATION DE LA POPULATION ACTIVE FRANSÇAISE
AU RECENSEMENT DE MARS 1990 ET À L’ENQUÊTE ANNUELLE SUR L’EMPLOI DE JANVIER 1990 : D’OÙ VIENNENT LES DIVERGENCES ?

Dominique Rouault


On dispose maintenant des résultats complets du recensement de mars 1990 ("sondage au quart" ou exploitation "lourde") et de l’enquête sur l’emploi de janvier 1990 (enquête annuelle par sondage) fournissant deux évaluations différentes à deux dates très proches de la population active française et de ses composantes. Le recensement (sondage au quart) dénombre ainsi 22 502 000 actifs occupés (y compris les militaires du contingent), 2 785 000 chômeurs et 20 588 000 inactifs de 15 ans ou plus (soit respectivement -107 000, + 52 000 et + 79 000 par rapport aux résultats de l’exploitation légère ou "exhaustif", écarts supérieurs à l’alea de sondage, car le traitement "lourd" du recensement opère des reclassements sur le code d’activité *). De son côté l’enquête dénombre selon la définition du Bureau International du Travail ("BIT") 22 322 000 actifs occupés, 2 254 000 chômeurs et 20 309 000 inactifs.

---

(1) La version provisoire de ce travail est parue en décembre 1992 dans la série des documents de travail de la Direction des Statistiques Démographiques et Sociales sous le N° F9221
Les évaluations brutes de la population active à l’enquête se situent donc nettement en dessous de celles du recensement, d’environ 180 000 actifs occupés par rapport à l’exploitation lourde (respectivement 287 000 par rapport à l’exploitation exhaustive légère) et 521 000 chômeurs (respectivement 479 000).

Les divergences entre les deux sources peuvent être imputées à cinq facteurs : une différence de champ d’investigation (l’enquête ne porte que sur la "population des ménages", c’est-à-dire sur les personnes résidant dans des logements ordinaires), une différence de concept dans le codage de l’activité (le codage au recensement ne satisfait pas aux critères précis du BIT qui nécessitent une batterie complexe de questions mise en œuvre dans l’enquête), un "effet protocole" à concept rendu équivalent (à l’enquête mode de questionnement, libellé des questions et enregistrement des réponses différent, et surtout intervient l’interaction enquêteur-enquêté), une différence de date de collecte (du 15 janvier au 15 février pour l’enquête, du 5 mars, date de référence, à la fin avril pour le recensement) et un alea de sondage à l’enquête (* l’intervalle de confiance à 95 % est de + (-) 130 000 sur la population active, + (-) 150 000 sur la population active occupée et + (-) 80 000 sur le chômage *), devant lequel l’alea dû au sondage au quart (* intervalles de confiance à 95 % respectivement + (-) 18 000, + (-) 7 000, + (-) 20 000 * ) est négligeable.

Deux premières approches [MARCHAND 19 et GUILLEMOT 20] effectuées au niveau "macro-social" après la disponibilité des résultats de l’exploitation légère (été 1991), avaient conclu que les différences de champ d’investigation, de concept d’activité, et de date de collecte ne semblaient pas suffire à rendre compte des divergences d’évaluation importantes constatées. Or ces données ont des enjeux importants, non seulement en termes de structure, notamment sur le taux d’activité des femmes ou le taux de chômage (cf tableau I en annexe), mais surtout en termes d’évolution récente et prévisible des ressources en main d’œuvre.

Il fut donc décidé fin novembre 1991 d’approfondir la comparaison et de procéder à un rapprochement "micro-social" des déclarations faites au Recensement (RP) et à l’enquête Emploi (EE) sur le "tiers sortant" de l’enquête (personnes enquêtées pour la dernière fois). Ce rapprochement s’est effectué en deux temps : établissement d’un "fichier de correspondance" magnétique au niveau des logements par appariement "à vue" des identifiants EE des 32767 locaux du "tiers sortant" avec les identifiants RP des mêmes logements retrouvés au recensement ( * pour des raisons de protection des personnes ces deux identifiants sont distincts et ne peuvent être rapprochés qu’au vu des documents de collecte * ), puis appariement informatique du "fichier de correspondance" et des fichiers magnétiques des deux sources, jusqu’au niveau des individus, identifiés à l’intérieur des logements par leur sexe, année et mois de naissance. L’apurement du fichier a été soumis à une stricte contrainte de délai (fin juin 92), mais la qualité globale de l’appariement est satisfaisante (environ 95 % des résidences principales retrouvées, et 91 % des individus enquêtés) et a permis de rapprocher environ 40 000 paires d’enregistrements d’adultes de 15 ans ou plus, dont 10 000 figurant dans le sondage au quart.
La présente étude vise à faire la part des différents facteurs, et notamment de l’"effet protocole", à partir du rapprochement "microsocial" des enregistrements des questionnaires individuels de l’enquête et des bulletins individuels du recensement pour 40 000 adultes.

Une première partie présente rapidement les protocoles de collecte et le codage de l’activité au recensement, puis explicite les choix faits lors de la simulation du concept de "type d’activité au sens du recensement" sur le questionnaire de l’enquête.

Dans une deuxième partie on opère une analyse statistique des personnes codées différemment aux deux sources, en termes de "lot de réponses" dans les deux protocoles (dans l’optique d’un contrôle de qualité des codages) et en termes de caractéristiques sociodémographiques (aux fins de désigner des sous populations à "risque objectif de divergence").

Dans une troisième partie on effectue une approche plus qualitative d’une sous-population ciblée par le précédent "cadrage", pour tenter de démêler les divergences tenant aux changements de situation des personnes, aux différences de codification d’une situation identique dans les deux chaînes de production, et au flou de déclaration d’une même situation par l’enquêté (recensé) lui même, cette approche débouchant sur l’interprétation sociologique.

Une dernière partie valide la représentativité des résultats obtenus par une tentative de "bouclage" décomposant les écarts bruts d’évaluation entre les deux sources selon les différents facteurs explicatifs avancés ci-dessus.

1 - Le concept d’"activité au sens du recensement"

Pour alléger l’exposé nous nous contenterons ici de rappeler les caractéristiques essentielles des deux protocoles de collecte et de traitement, en opérant au besoin quelques simplifications. Pour une description plus précise, on se reportera à la bibliographie en annexe. De même pour éviter de trop longs intitulés nous avons dû recourir à l’utilisation de codes (* les codes détaillés portent le nom qu’ils ont dans les fichiers diffusés pour faciliter leur identification immédiate aux utilisateurs habituels de ces fichiers. Nous avons opéré des regroupements, indicés en général par le suffixe C (condensé) ou pour les codes d’activité une codification en trois postes (actif occupé, chômeur, inactif) indicée par le suffixe 3. En cas de besoin l’exploitation exhaustive légère du recensement est distinguée de l’exploitation lourde (sondage au quart) par l’ajout d’un suffixe H.)
1.1. Présentation rapide des protocoles de collecte

Le questionnaire de l’enquête Emploi\(^1\) est administré annuellement en mars par un réseau d’enquêteurs qualifiés à l’ensemble des adultes de 15 ans au moins résidant à titre principal (ou éventuellement rattachés au ménage occupant) dans un échantillon de logements suivi trois années de suite et renouvelé par tiers. En 1990, année du recensement général de la population, cette enquête a été avancée en janvier et son questionnaire rénové. C’est l’enquêteur qui enregistre les réponses des personnes au fur et à mesure d’un cheminement rigoureux comportant de nombreux "filtres", et cet enquêteur est autorisé à remplir le questionnaire d’un adulte absent au moment de son passage, en interrogeant un autre adulte du ménage.

En revanche les bulletins du recensement\(^1\) ne comportent qu’un nombre limité de questions, formulées de façon assez sommaire, et ont été dans la plupart des cas remplis par les individus eux-mêmes, assez rarement en présence d’un agent recenseur.

1.2. Le codage du type d’activité au recensement

Le recensement utilise le verso du bulletin individuel (renseigné par les personnes qui travaillent) et la question 9 (renseignée par les autres), pour coder ce qu’on désignera ici par le terme d’"activité spontanée" (regroupée infra en cinq postes INAC5). Il exploite en complément la question 11 sur la "recherche d’emploi" (ANREC regroupé ici en trois modalités, "oui", "non", "blanc") pour coder définitivement le type d’activité (regroupé ici en trois postes TACT3 dans le sondage au quart, ou TACT3H dans l’exhaustif léger).

Les recensés notifiant (au verso du bulletin) qu’ils travaillent, sont codés actifs occupés. Pour ceux qui se classent spontanément "chômeur (inscrit ou non à l’ANPE) " à la question 9, on examine la réponse à la question 11 : s’ils indiquent explicitement ne pas chercher un emploi, ils sont recodés inactifs, sinon ils sont codés chômeurs (même en l’absence de réponse, donc "au bénéfice du doute").

Les "étudiant, élève, stagiaire non rémunéré" spontanés sont codés inactifs (du moins s’ils ne travaillent pas). Pour les autres cas ("retraité", "retiré des affaires", "femme au foyer" ou "autre inactif") on prend également en compte la réponse à la question 11 : ces inactifs spontanés sont recodés chômeurs s’ils indiquent explicitement rechercher un emploi (cela implique la précision d’une ancienneté de recherche) ; ils sont codés inactifs dans le cas contraire (y compris cette fois en l’absence de réponse).

---

\((1)\) Voir maquettes simplifiées en annexe
Le classement spontané (INAC5) des personnes donne donc lieu à une reventilation du code final "type d’activité" (TACT3 ou TACT3H), modifiant en particulier le partage chômage/inactivité, selon le tableau 1 :

Tableau 1

<table>
<thead>
<tr>
<th>INAC5 (activité spontanée)</th>
<th>ANREC (recherche d’emploi)</th>
<th>TACT3(ou TACT3H) (type d’activité)</th>
</tr>
</thead>
<tbody>
<tr>
<td>occupé (verso bulletin renseigné)</td>
<td>OUI ou BLANC</td>
<td>actif occupé</td>
</tr>
<tr>
<td>chômeur (inscrit ou non)</td>
<td>NON</td>
<td>chômeur</td>
</tr>
<tr>
<td>étudiant,élève,stagiaire non rémunéré</td>
<td>OUI</td>
<td>inactif</td>
</tr>
<tr>
<td>retraité,retiré des affaires</td>
<td>NON ou BLANC</td>
<td>inactif</td>
</tr>
<tr>
<td>femme au foyer,autre inactif</td>
<td>OUI</td>
<td>chômeur</td>
</tr>
</tbody>
</table>

Ajoutons pour être complet que les personnes recensées dans certains établissements mais "réintégrées" dans la résidence principale qu’ils indiquent, remplissent un autre type de bulletin, et ont un code INAC5 ("activité spontanée") à blanc les rendant à ce niveau indiscernables des "occupés spontanés". Ils sont en revanche directement ventilés selon le "type d’activité" (TACT3 ou TACT3H) : en actifs occupés pour les militaires (y compris du contingent) et en inactifs pour les élèves internes.

1.3. Simulation du concept d’"activité au sens du recensement" sur le questionnaire de l’enquête

La correction de concept ne peut qu’être approximative (elle vise à dégager un effet protocole "pur" en partie artificiel) et consiste à simuler le codage du type d’activité à l’enquête selon les conventions du recensement, et non plus celles du BIT.

Nous avons pour cette étude codé le type d’activité à l’enquête selon une logique la plus ajustée possible aux conventions du recensement, malgré une part d’arbitraire inévitable, en ne faisant intervenir que les questions dont la formulation est très proche de celles du bulletin du RP : d’abord deux questions posées en début de questionnaire à tous les enquêtés : la question Q2 (code regroupé FI5) sur l”’occupation actuelle” (dénommée ici "activité spontanée", et dont les rubriques sont très proches de son homologue au recensement INAC5), et la question Q4a sur la recherche d’emploi (code RECHE) ne comportant cette fois que deux réponses possibles (OUI ou NON) ; ensuite la question B8 sur l’ancienneté de recherche d’emploi posée sensiblement plus tard et seulement aux enquêtés susceptibles à divers titres d’une recherche (code regroupé
DREC en deux réponses possibles, OUI si une ancienneté est déclarée, NON dans le cas contraire, donc aussi si elle est non renseignée). Le classement spontané (FI5) des personnes donne donc lieu à une simulation du code final "type d’activité à l’enquête" (TACT3EE), selon le tableau 2 calqué sur le tableau 1 :

Ainsi, parmi les personnes se classant spontanément chômeurs ("chômeurs spontanés" à FI5) sont considérées "chômeurs au sens du recensement" (type d’activité TACT3EE) celles qui n’ont pas démenti immédiatement chercher un emploi en début d’enquête (question RECHE). En revanche les personnes se classant spontanément inactives ne sont codées "chômeurs au sens du recensement" que si elles déclarent immédiatement chercher un emploi (RECHE) et le confirment ensuite en précisant une ancienneté de recherche (question DREC). Enfin les personnes déclarant spontanément travailler et les militaires du contingent sont codés actifs occupés, et les "étudiant, élève en formation, en stage non rémunéré" sont codés inactifs sans examen des réponses à la recherche d’emploi.

1.4. Impact de la simulation retenue

Après correction de concept l’enquête emploi (tableau II en annexe en regard des résultats du sondage au quart) décompterait 22 361 000 actifs occupés (+ 39 000 par rapport au concept BIT), 2 436 000 chômeurs (+ 182 000) et 20 088 000 inactifs (- 221 000).

Concernant le partage chômage/inactivité, concerné au premier chef par cette correction, le solde est donc excédentaire de 182 000 chômeurs par rapport au concept BIT
(16 % des chômeurs "au sens du recensement" ne sont pas chômeurs BIT, et seulement 9 % des chômeurs BIT ne sont pas chômeurs "au sens du recensement").

En revanche l'effet de notre codage sur le classement spontané des individus revient à "déclasser" 13 % des "chômeurs spontanés" en inactifs, mais en revanche à réintégrer 1 % des "inactifs spontanés" en chômeurs. C'est donc une opération globalement déficitaire de 98 000 chômeurs.

Le recensement de son côté "déclasse" également 9 % des chômeurs spontanés, mais en contrepartie il réintègre 1 % d'inactifs, et l'opération est cette fois globalement bénéficiaire de 39 000 chômeurs environ.

Notre simulation se révélant empiriquement un peu plus exigeante que le traitement du RP, examinons les conséquences pour l'EE d'une extension du concept de "chômage au sens du recensement". En y incluant les personnes revenant sur leur première déclaration (question RECHE) à la question sur l'ancienneté de recherche d'emploi (question DREC), on définirait un "halo" de 80 000 personnes à l'enquête autour de notre définition (31 000 chômeurs spontanés et 48 000 inactifs spontanés), presqu'équivalent aux 98 000 chômeurs spontanés "perdus". Il convient cependant de noter que ce halo ne serait "récupéré" qu'à travers des questions supplémentaires de l'enquête (ainsi la question Q4B sur le souhait de travailler, la question Q7 sur l'inscription à l'ANPE ou la question B6 sur la situation au regard de la recherche d'emploi) qui constituent un préalable à la question sur l'ancienneté de recherche (DREC) et n'ont aucun équivalent sur le bulletin du recensement.

1.5. Restriction à la population des ménages

Le champ exact de l'enquête fait l'objet d'une certaine imprécision à cause des membres "rattachés" aux ménages. En s'en tenant aux chiffres publiés par l'INSEE (pour une approche plus précise, voir 4ème partie) on peut restreindre la comparaison des deux sources à la population des ménages au recensement et à celle qui lui est théoriquement assimilée à l'enquête : le tableau III en annexe fait alors apparaître 22 237 000 actifs occupés dans l'exploitation lourde du RP (sondage au quart) contre 22 340 000 à l'EE (103 000 de plus), 2 723 000 chômeurs au RP contre 2 433 000 à l'EE (290 000 de moins), 19 739 000 inactifs au RP contre 19 943 000 à l'EE (204 000 de plus). L'exploitation lourde du RP accroît les divergences avec l'enquête par rapport à l'exploitation exhaustive du RP sur les niveaux de l'emploi et du chômage. Si l'on néglige les variations conjoncturelles (de l'ordre de quelques dizaines de milliers cf. 4e partie) on voit que l'écart sur la population active occupée est acceptable compte tenu de l'alea de sondage de l'enquête, mais que les écarts sur le nombre d'inactifs et surtout le nombre de chômeurs sont nettement hors bornes de la marge habituellement tolérée (* intervalle de confiance à 95 % * ).
1.6. Tableau résumé

Les diverses évaluations évoquées supra sont résumées dans le tableau 3:

**Tableau 3**

*Population de 15 ans ou plus par type d'activité (en milliers)*

<table>
<thead>
<tr>
<th>Type activité</th>
<th>Recensement mars 1990</th>
<th>enquête-emploi janvier 1990</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Total</td>
</tr>
<tr>
<td>Exhaust.</td>
<td>45851</td>
<td>45875</td>
</tr>
<tr>
<td>Quart</td>
<td>22609</td>
<td>22502</td>
</tr>
<tr>
<td>Chomeurs</td>
<td>2733</td>
<td>2785</td>
</tr>
<tr>
<td>Inactifs</td>
<td>20509</td>
<td>20588</td>
</tr>
</tbody>
</table>

2 - Une analyse statistique des individus divergeant

2.1. Préliminaires

Le rapprochement a permis d'apparer 41917 adultes de 15 ans ou plus, soit 91 % du tiers de l'enquête, mais seulement 40117 (87 %) peuvent être considérés comme "sûrs", puisque de logement, sexe, année et mois de naissance identique, les 1795 autres étant "douteux" car appariés dans le même logement sur un critère de reconnaissance plus souple autorisant des erreurs de saisie sur la date de naissance.

L'absence des variables brutes de saisie dans les fichiers auxquels nous avons eu accès nous oblige à raisonner sur les variables redressées et complique parfois cette analyse statistique menée à un niveau "mésosocial", et limitée aux adultes "appariés sûrs" par sécurité.
2.2. Le déséquilibre porte sur le partage chômage/inactivité

La matrice de passage entre le type d'activité à l’enquête (TACT3EE) et à l’exhaustif (TACT3H) est la suivante sur les 40 117 adultes appariés "sûrs" :

**Tableau 4**

<table>
<thead>
<tr>
<th>Type d'activité a l'exhaustif (TACT3H)</th>
<th>A l’enquête (TACT3EE) (matrice observée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs occupés</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>19212</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>322</td>
</tr>
<tr>
<td>Inactifs</td>
<td>506</td>
</tr>
</tbody>
</table>

On relève immédiatement deux flux non diagonaux déséquilibrés : les "inactifs RP / chômeurs EE" ne compensent pas, et de loin, les "chômeurs RP/ inactifs EE". De fait les soldes sur l’échantillon sont déficitaires à l’enquête de 163 chômeurs et 53 actifs occupés, et excédentaires de 216 inactifs.

Case par case, on ne peut faire dans le tableau la part des flux saisonniers et du protocole de collecte et de traitement des données (* dont le redressement par "hot deck" opéré au recensement sur les 900 000 bulletins blancs ou incomplets pour coder le type d’activité, qui contribue à gonfler artificiellement les cases non diagonales si individuellement ce redressement s’est avéré malheureux * ). On peut cependant supposer en première approximation que les flux saisonniers sont équilibrés à quelques unités près, et que les redressements sont répartis au prorata des cases. Il se confirme alors que sur les 2372 cas de divergences situés hors diagonale la source essentielle du déséquilibre "mésosocial" réside effectivement dans les échanges entre chômage et inactivité, les autres cases hors diagonale se compensant relativement bien entre elles.

Le risque global de divergence est de 5,9 %. Il est plus élevé parmi les étrangers (10,1 %), et, plus curieusement, lorsque l’individu a répondu lui-même à l’enquête emploi (6,4 %).

Le risque de classement "Occupé RP/Inactif EE" est le plus élevé (1,36 %), un peu supérieur au risque de classement inverse "Inactif RP/Occupé EE" (1,26 %). Puis vient le risque "Chômeur RP/Inactif EE" (1,04 %), nettement plus élevé que "Inactif RP/Chômeur EE" (0,61 %). Les risques "Occupé RP/Chômeur EE" (0,83 %) et "Chômeur RP/Occupé EE" (0,80 %) sont presque équilibrés.
On ne dispose pas sur ces 40 117 adultes du code final de l'exploitation lourde (TACT3). Le "bouclage" de la 4\textsuperscript{e} partie montrera que l'échantillon des 9 979 d'entre eux figurant dans le sondage au quart, non représentatif, ne peut être utilisé directement, mais nous permettra aussi de reconstituer la matrice de passage entre le type d'activité à l'enquête (TACT3EE) et dans l'exploitation lourde (TACT3), matrice qui serait approximative-ment sur les 40 117 adultes :

**Tableau 5**

<table>
<thead>
<tr>
<th>Type d'activité au quart (TACT3)</th>
<th>A l' enquête (TACT3EE) (matrice reconstituée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs occupés</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>19004 (0,99)</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>387 (1,20)</td>
</tr>
<tr>
<td>Inactifs</td>
<td>649 (1,28)</td>
</tr>
</tbody>
</table>

(Entre parenthèses les rapports à la matrice supra réellement observée TACT3H/TACT3EE dans l'exploitation exhaustive : ainsi par exemple le risque "Chômeur RP/Occupé EE serait dans l'exploitation lourde de 0,80 * 1,20 = 0,96 : on voit que les reclassements entre l' exhaustif et le quart affectent peu le partage chômage/inactivité mais surtout les échanges avec l'emploi)

Les 40 117 individus de notre échantillon sont censés représenter la population de plus de quinze ans des ménages en janvier soit 44 680 000 personnes (cf 4\textsuperscript{e} partie infra). Un cas individuel doit donc à ce niveau être pondéré par un coefficient 1114.

**2.3. La logique du codage n'est pas en cause**

Examinons d'abord sur les 2 372 adultes "divergeant" dans l'échantillon le jeu de réponses aux deux protocoles. Cette analyse vise à dégager les sous populations "sensibles", soit parce qu'elles représentent un nombre de cas de divergence important, soit parce que, à nombre de cas seulement moyen, elles offrent un risque global de divergence important compte tenu de leur effectif. (* Il demeure que l'impossibilité de décider au niveau individuel si la divergence est due à un changement de la situation de l'intéressé entre janvier et mars obscurcit les conclusions menées à ce niveau "mésosocial", puisque certaines sous populations peuvent être caractérisées par des risques de mobilité plus élevés que d'autres. *)

Pour chaque sous-population caractérisée par un jeu de réponses : le type d'activité déclaré spontanément sur le bulletin du recensement (INAC5 : le code à blanc regroupe malheureusement les "vrais" occupés spontanés, les non réponses qui sont redressées ensuite et reventilées, et 26 élèves internes réintégrés qui ont rempli un autre bulletin),
son homologue sur le questionnaire de l’enquête (FI5), la déclaration éventuelle de recherche d’emploi au RP (ANREC regroupé en 3 postes) et les déclarations homologues au début (RECHE) puis en milieu d’enquête (DREC), le tableau IV en annexe fournit la proportion de cas de divergence représentée par la sous-population, le risque global de divergence de la sous-population, le type de risque premier et sa valeur.

A titre d’exemple servant de guide de lecture au tableau, examinons le cas des chômeurs spontanés au RP (INAC5) : ils représentent 29 % des cas de divergence, mais leur risque est très élevé (32 %) et porte d’abord sur le classement en inactif à l’enquête (13 %).

Parmi ceux qui s’étaient déclarés (F15) "autres inactifs" (ni retraités ni étudiants) à l’EE (ce double classement spontané est à l’origine de 5 % des cas de divergence et génère 59 % de risque essentiellement de type "chômeur RP/inactif EE"), plus de quatre sur cinq indiquent rechercher du travail au recensement (ANREC) et déclarent le contraire en début (RECHE) d’enquête (d’où logiquement 100 % de risque) : pour ces 120 000 personnes (entre 4 et 5 % du total des cas de divergence) les deux codages sont cohérents et la possibilité d’un changement de situation reste a priori indécidable.

À l’issue d’un examen complet on constate que le codage final des situations déclarées (sous réserve de bonne saisie) se déroule conformément à la logique prévue (* au dessus de la barre de 3 % des cas de divergence examinée dans le tableau IV seuls 45 chômeurs déclassés inactifs au RP pourraient apparaître douteux *).

2.4. Le traitement paraît cohérent pour les actifs occupés

Un bilan des stocks et des flux sur les déclarations spontanées montre (après correction pour tenir compte des élèves internes réintégrés) un déficit de 63 occupés (ou éventuellement non répondants) spontanés. Ce déséquilibre provient surtout de passages de l’emploi à la retraite (+ 71 nets) dont on ne retrouve pas la contrepartie parmi les flux en provenance des chômeurs (+ 6 nets), étudiants ou élèves (+ 6 nets), ou "autres inactifs" (+ 0 nets) spontanés.

Ce déficit d’emplois spontanés au recensement se transforme en excédent (53 cf supra) dans le codage final de l’exploitation exhaustive (grâce donc à l’indicateur d’emploi constitué par le remplissage au moins partiel du verso du bulletin individuel). Cette correction va dans le sens attendu par l’évolution conjoncturelle (cf 4e partie infra). Rappelons cependant que l’exploitation lourde recreée pratiquement le déficit.
2.5. L'"effet protocole" joue essentiellement sur la déclaration de recherche d'emploi

Le bilan montre en revanche un excédent de 39 chômeurs spontanés au recensement, dont l'origine majeure provient d'un flux d'anciens "autres inactifs" à l'enquête sans contrepartie équivalente (+ 47 nets). Le codage final accroît considérablement l'excédent (163 cf supra) : c'est donc la déclaration de recherche d'emploi servant d'aguillage pour reventiler les classements bruts spontanés, qui est surtout affectée par le protocole de collecte, comme on le pressent à travers l'exemple développé ci-dessus.

Devant la difficulté d'éliminer le parasitage des changements de situation éventuels, nous avons analysé cette question en isolant les deux catégories en cause de "chômeurs spontanés" et "autres inactifs spontanés", c'est à dire les 6 641 000 individus après pondération qui demeurent entre les deux sources dans une de ces deux catégories, ou qui passent d'une catégorie à l'autre. Ainsi les transitions avec l'emploi, la retraite ou la formation sont neutralisées, et les variations conjoncturelles a priori négligeables.

Le tableau V en annexe restitue le jeu de réponses sur cette population en croisant les situations spontanées déclarées aux deux dates. Les chômeurs passant du non au oui et les inactifs du oui au non à l'enquête entre le début (RECHE) et le milieu (DREC) de l'entretien sont comptabilisés en "hésitantes" (c'est le halo dont nous avons parlé supra en 1re partie).

Pour les chômeurs spontanés au recensement une constatation frappe immédiatement : au RP seuls 55 000 d'entre eux répondent négativement à la question sur la recherche d'emploi (ANREC) contre 227 000 qui s'abstiennent. Or 27 000 seulement avaient hésité à l'enquête (non à RECHE, oui à DREC). L'abstention au recensement n'est absolument pas symétrique de l'hésitation à l'enquête, mais joue pour une part un équivalent du "non" à l'enquête sur les "chômeurs spontanés" : 56 % des chômeurs RP s'abstenant avaient répondu "non" en début d'enquête (RECHE), un peu moins (40 %) pour les chômeurs aux deux dates.

Pour les inactifs spontanés au recensement les réponses négatives à la recherche d'emploi (ANREC) sont remises à blanc dans le fichier définitif et sont comptabilisées ici avec les abstentions : elles entraînent les mêmes conséquences de classement final en inactifs. En revanche 135 000 d'entre eux seulement ont indiqué explicitement rechercher un emploi dans les deux sources (RECHE et ANREC), 112 000 ne l'ont fait qu'au recensement et 178 000 qu'à l'enquête. Cette fois le "oui" semble s'exprimer plus facilement face à l'enquêteur.

S'opère cependant un glissement net de 51 000 personnes de la catégorie d'autre inactif spontané à celle de chômeur spontané entre l'enquête et le recensement, puisque la première catégorie compte 4 684 000 personnes au RP (contre 4 736 000 à l'EE) et la seconde 1 956 000 personnes au RP (contre 1 905 000 à l'EE) ; or ce glissement entraîne
mécaniquement un surplus de chômeurs confirmés au RP (TACT3H) parmi ceux qui s'abstiennent de répondre à la question sur la recherche d'emploi (ANRE).

Au total "joue" ainsi en faveur du chômage un solde net de 283 000 personnes, dont :

- sur les chômeurs spontanés aux deux sources 87 000 nets passant du non au oui (* 116 000-29 000 *), 66 000 passant du non à l'abstention, et 21 000 (* 17 000 + 4 000 * ) de l'hésitation au oui ou à l'abstention.

- sur les transitions spontanées "chômage/inactivité" un solde net de 86 000 personnes (* 119 000 + 2 000 + 57 000 + 1 000-90 000-3 000 *).

- sur les inactifs spontanés aux deux sources 23 000 nets passant du non ou de l'hésitation au oui (* 105 000 + 6 000-88 000 *).

Là réside en dernière instance l'"effet protocole" : dans le face à face avec l'enquêteur ne se retrouvent pas 280 000 recherches (au moins attentes) d'emploi dont 150 000 s'exprimeraient explicitement sur le bulletin du recensement, et 130 000 seraient seulement non démenties. En particulier 130 000 chômeurs spontanés à l'EE ayant répondu négativement à l'enquêteur en début d'enquête confirment explicitement leur recherche au RP (contre seulement 30 000 prenant l'attitude inverse).

Or on retrouve effectivement (tableau V bis) 300 000 personnes indiquant une ancien-neté de recherche supérieure à 3 mois au RP (donc excluant a priori un changement de leur situation) qui auraient démenti chercher un emploi face à l’enquêteur : 118 000 (* 102 000 + 16 000 * ) sont des chômeurs spontanés aux deux dates, 75 000 (* 72 000 + 3 000 * ) des inactifs spontanés aux deux dates et 101 000 (* 100 000 + 1 000 * ) des ex-inactifs se déclarant chômeurs. Sur le total 7 % seulement (* 20 000 * ) ont changé d'avis entre le début (question RECHE) et la fin (question DREC) de l'entretien et relèvent du "halo" défini en 1re partie.

Devant une telle constatation, qui donne cette fois une réponse "technique" imparable à l'énigme du départ (tous sont en définitive codés "chômeurs RP/inactifs EE" et, sans préjuger des divergences réelles sur les autres, une incohérence de cette taille suffit à rendre compte pour l'essentiel des écarts à expliquer), on ne peut que s'interroger sur les conduites des personnes concernées. Sont-elles fantaisistes ou de mauvaise foi ? Nous chercherons dans leurs caractéristiques sociodémographiques une clé pour l'interprétation.
2.6. Le flou du partage "chômage/inactivité" porte sur les populations démunies vis à vis du marché du travail

Dans l’exploitation exhaustive nous ne disposons pratiquement que du sexe et de l’âge des personnes. La catégorie socioprofessionnelle, disponible à l’enquête, est inutilisable isolément puisqu’elle dépend elle-même du type d’activité dans cette source ( * les inactifs et chômeurs n’ayant jamais travaillé ne sont pas dotés par l’INSEE d’un statut social * ). Nous avons donc retenu dans un premier temps le diplôme le plus élevé obtenu déclaré à l’enquête, indicateur social indépendant du type d’activité. Le tableau VI fournit, pour chaque catégorie représentant une proportion suffisamment élevée de cas de divergence, le risque global de classement différent, le "double classement premier", les deux risques "Chômeur RP/Inactif EE" et "Inactif RP/Chômeur EE" qui nous intéressent et leur écart ( * la matrice supra du paragraphe 22 montre que dans l’exploitation lourde ces deux risques seraient pratiquement identiques * ).

Les femmes ont une probabilité nettement plus forte que les hommes d’être classées diversément chômeuses ou inactives, et la balance penche vers le classement "Chômeur RP/Inactif EE" (1,5 % de risque, 0,6 % de plus que le risque inverse). Les tranches d’âge se répartissent en trois catégories : les "50-59 ans" et les "25-39" ans ont deux risques élevés et un solde sensible en faveur du chômage au RP ; les "15-24 ans" et les "60 ans ou plus" ont deux risques faibles s’équilibrant, les "40-59 ans" deux risques moyens s’équilibrant. Ce sont donc les femmes de 25 à 39 ans ou de 50 à 59 ans qui encouragent le risque de classement "Chômeur RP/Inactif EE" à la fois le plus élevé et le plus déséquilibré.

Quant au "capital culturel" détenu, il s’avère un critère hautement discriminant : le risque "Chômeur RP/Inactif EE" est pratiquement décroissant avec le diplôme, du certificat d’études (ou moins), à un diplôme "de niveau intermédiaire" (BEPC seul, BEP ou CAP), au baccalauréat et à un diplôme supérieur (BACC + 2 ou plus). Le risque inverse "Inactif RP/Chômeur EE" place les diplômes intermédiaires avant le certificat d’études, et le baccalauréat après le niveau "BACC + 2", mais le bilan fait bien apparaître une dichotomie : les non bacheliers présentent deux risques élevés et une balance favorable au classement en chômeurs (RP), les détenteurs du baccalauréat ou d’un diplôme supérieur offrent des risques et un déséquilibre plus faible.

Ainsi ce sont les femmes d’âge moyen, non bachelières et en particulier celles ne disposant que d’un bagage scolaire minimum qui font massivement "basculer" le bilan vers le chômage au recensement (autour de 1,5 point d’écart), le risque "Chômeur RP/Inactif EE" étant maximum chez celles de 25 à 39 ans n’ayant au mieux que le certificat d’études (3,6 %). On voit que l’"effet protocole" joue de façon privilégiée sur des populations bien particulières et disposant de peu d’atouts sur le marché du travail.

Pour pouvoir néanmoins préciser la position professionnelle de ces individus divergeant, on peut (avec précaution) se servir de l’échantillon des 9979 personnes figurant
dans le sondage au quart, en redressant les cases affectées par l’alea de sondage. La matrice reconstituée sur ces 9 979 personnes serait approximativement :

**Tableau 6**

<table>
<thead>
<tr>
<th>Type d’activité au quart (TACT3)</th>
<th>A l’enquête (TACT3EE) (matrice reconstituée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs occupés</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>4770 (0,99)</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>97 (1,23)</td>
</tr>
<tr>
<td>Inactifs</td>
<td>163 (1,37)</td>
</tr>
</tbody>
</table>

(Entre parenthèses les rapports à la matrice réellement observée dans l’exploitation lourde TACT3EE/TACT3 : ainsi par exemple le risque observé "Inactif RP/Chômeur EE" doit être redressé d’un coefficient 1,5)

Pour que ces 9 979 individus représentent la population de plus de quinze ans des ménages en janvier un cas individuel doit cette fois être pondéré par un coefficient 4479.

Quant à la position socioprofessionnelle nous l’avons reconstituée par mixage des deux sources (cette fois le quart en fournit un codage concurrent de celui de l’enquête), en privilégiant la catégorie socioprofessionnelle à l’enquête (* bâtie sur un processus en principe plus fiable *), sauf pour les "divers sans activité professionnelle" pour qui nous avons utilisé la CS au recensement si elle était différente.

Le tableau VII est donc construit sur les mêmes principes que le tableau VI, le risque "Chômeur EE/Inactif RP" étant redressé d’un coefficient 1,5 pour restituer la balance (* le risque "Chômeur RP/Inactif EE" est quant à lui représentatif au vu de la matrice *).

On pourra constater que les lignes à caractère purement "démographique", reproduites, donnent des chiffres en général concordants avec ceux du tableau VI, mais que les risques "Inactif EE/Chômeur RP" sont inférieurs pour les jeunes de moins de 25 ans et les adultes de 50 à 59 ans. Les résultats doivent donc être interprétés avec prudence, et seulement pour confirmer ou non dans ses grands traits le tableau d’ensemble dégagé à partir de l’exploitation exhaustive.

De fait le "capital économique" agit dans le sens attendu révélé par le "capital culturel" : d’un côté les catégories ouvrière et employée se distinguent par deux risques élevés, mais au profit de "Chômeur RP/Inactif EE", surtout chez les femmes (jusqu’à 4,5 % pour les ouvrières), et donc un bilan favorable au chômage au recensement (autour de 2 points) ; à l’opposé les professions intermédiaires ou supérieures offrent deux risques moyens ou faibles se compensant globalement ; les agriculteurs et indépendants (non agricoles) inflexissent eux aussi la balance vers le chômage (autour de 0,6 points) par un risque "Chômeur RP/Inactif EE" moyen et un risque inverse nul.
Les catégories socioprofessionnelles nombreuses les plus sujettes au flou de codage entre chômage et inactivité au profit du chômage au recensement sont les ouvriers non qualifiés de l'industrie, les employés administratifs d'entreprise, les personnels de services directs aux particuliers et les employés de commerce (toutes à risque "Chômeur RP/Inactif EE" supérieur à 3 %), catégories de main d'œuvre fortement ou au moins relativement féminisées et faiblement qualifiées.

A l’issue de cette analyse statistique mettant en lumière un "espace social" aussi nettement polarisé on est conduit à penser que les "conduites d’incohérence" relèvent de comportements sociologiques bien plus que de l’inconséquence des individus.

3 - Une analyse qualitative des incohérences

3.1. La méthode

Parallèlement à cette analyse statistique, nous avions édité dès la fin de constitution du fichier individuel, une liste exhaustive des individus appariés figurant dans le sondage au quart dont les codes "type d’activité" (TACT3 et TACT3EE) divergeaient à l’enquête et dans l’exploitation lourde du recensement avec leurs deux identifiants dans les deux sources, en leur attribuant un numéro d’ordre national. Nous avions demandé à notre équipe de retrouver les bulletins et questionnaires de ces personnes, de reporter les réponses aux questions essentielles sur deux maquettes simplifiées (qu’on trouvera en annexe), et de faire un diagnostic qualitatif “à vue” sur le motif de la divergence selon une grille que nous avions proposée. Cette fois le retour direct aux documents conduit à une approche véritablement "microsociale“ et évite les deux obstacles de l’analyse précédente, de l’utilisation de codes redressés et (en partie) du brouillage dû au décalage de dates.

Ce diagnostic (tableau VIII en annexe) porte sur 553 individus. ( * quelques "appariés sûrs", non retrouvés, ont été compensés par des "appariés douteux" s’avérant identiques ex post * ). La pondération individuelle permettant d’estimer des effectifs réels est de 4495, mais chaque colonne doit être redressée par son coefficient pour tenir compte du biais dans le sondage au quart (cf 2e partie supra).

3.2. Problèmes de codages, changements de situation et flous déclaratifs

Nous n’exploiterons pas ici l’aspect comptable de ces données qui est susceptible de contestations mais conforte pour l’essentiel les résultats tirés des études quantitatives.
Nous indiquerons seulement que les seuls chiffres vérifiables sont, après redressement et pondération, d'un ordre de grandeur "réaliste" compte tenu des limites de l'exercice.

Les causes de divergence peuvent être regroupées en trois grandes catégories :

- les problèmes de codage dus à l'INSEE sont, comme prévu, limités (un sixième des causes), minimes en particulier les erreurs de saisie sur les codes bruts spontanés (FI ou INAC), ou les contestations possibles du code simulé à l'enquête (TACT3EE) (* dans le cas d'une interprétation plus large de la recherche d'emploi par exemple à DREC *) ; les "hot-decks malheureux" sont somme toute modestes (6 % des causes), du même ordre que les contestations du codage complexe au recensement (TACT3). Sur ce contrôle de qualité nous renvoyons à [BADEYAN 15]).

- les changements de situation sont considérés comme probables lorsqu'on dispose de dates ou durées (notamment l'ancienneté de chômage sur le bulletin du recensement ou une durée de stage ou emploi à durée déterminée sur le questionnaire individuel de l'enquête etc.), comme possibles seulement lorsque rien ne permet d'en douter (diagnostic ultime après épuisement des autres causes). Ces changements de situation expliqueraient la moitié des causes de divergence.

- les "flous déclaratifs" de la part de l'enquêté, incluant les "double situation plausible" (étudiant travailleur, retraité chômeur etc.) et surtout les "incohérences de réponses" traduisant des contradictions flagrantes (du type déjà évoqué d'une ancienneté de recherche d'emploi au RP supérieure à 3 mois et non signalée à l'EE, ou d'une absence d'activité professionnelle antérieure sur le bulletin du recensement et d'une profession à l'EE etc.). Est confirmé l'importance du flou déclaratif (environ un tiers des causes de divergence) et notamment des incohérences (environ un quart).

3.3. "Faux chômeurs" ?

Nous nous sommes intéressés particulièrement aux 64 "Chômeurs RP/Inactifs EE" déclarés incohérents (représentatifs de 288 000 personnes sans redressement nécessaire) et nous avons nous même examiné les maquettes de ces 64 cas et de leurs 16 "vis à vis" classés "Inactifs RP/Chômeurs EE" ( * case non représentative, à redresser peut-être par 1,5, et envisagée seulement de manière illustrative * ). Globalement l'incohérence massive ne fait aucun doute : 63 ont indiqué rechercher un emploi (ANREC) au RP, 61 depuis au moins trois mois, et 45 depuis au moins un an. 50 ont répondu eux mêmes à l'enquêteur (trice). Or, si 44 se déclaraient spontanément chômeurs, 1 seul a déclaré une recherche en début d'entretien (RECHE). 18 ont dit cependant "souhaiter travailler" (SOUH) et 23 être inscrits à l'ANPE. 15 se déclaraient indisponibles pour travailler immédiatement, et 5 ont fini par donner une ancienneté de recherche (DREC) dont 2 avec des démarches. En outre leur exclusion du chômage BIT
(un seul rescapé) est en droit légitime (sauf dans un cas, l’enquêteur a respecté, sur le papier au moins, tous les filtres). D’ailleurs quelle définition extensive du chômage permettrait de les inclure majoritairement dans un quelconque halo ?

Pour tant nous n’éprouvons, sauf pour quelques cas suspects, guère de doute : ils n’inventent sûrement pas leur recherche (au moins attentive) d’emploi. Certes il faut pour risquer une interprétation, sortir du "château" délimitant les frontières de la statistique dans la métaphore topologique de [HERAN 39] et s’aventurer dans l’"espace non-pop- perien du raisonnement sociologique" [PASSERON 40]. Mais nous croyons pour notre part devoir admettre cette tension méthodologique entre la "contextualisation historique et le raisonnement expérimental" et les "ruptures de style" qui y sont liées, d’autant qu’en l’occurrence nous courrons, on le verra, peu de risques.

En effet le statisticien ne s’étonnera pas outre mesure des incohérences constatés ci-dessus sachant que "les agents sociaux (qu’il) classe sont producteurs non seulement d’actes classables, mais aussi d’actes de classement qui sont eux mêmes classés", actes "qui obéissent à la logique du parti-pris, qui, comme au tribunal, affronte non des jugements logiques, justiciables du seul critère de cohérence, mais des accusations et des défenses" [BOURDIEU 41], logique pratique dans laquelle se déroule le face à face entre l’enquêteur et l’enquêté.

Or notre analyse, en situant les enquêtés dans l’espace social, où "la vision que chaque agent a de l’espace, dépend de sa position dans cet espace" [BOURDIEU 55] nous livre leur "habitus, comme système de dispositions à la pratique" : il fait que "les agents qui en sont dotés se comporteront d’une certaine manière dans certaines circonstances", tout en obéissant " à une logique pratique, celle du flou, de l’à peu près qui définit le rapport ordinaire au monde" [BOURDIEU 43].

Ces 64 individus sont bien ( * comme leurs vis à vis à qui ils ressemblent à s’y méprendre, mais qui ont, eux, franchi "victorieusement" tous les filtres en présence de l’enquêteur, pour les "oublier" au recensement * ) la synthèse vivante de cette zone de l’espace social que nous avons cernée plus haut : des femmes surtout (49), en majorité âgées de 25 à 39 ans (26), des quinquagénaires (21), ouvriers (19) ou employés (31) très peu qualifiés : ouvrières dans la chaussure, la confection, l’alimentaire, manœuvres, magasinières ou manutentionnaires, aides ménagères, femmes de ménage ou agents de service, maçons ou plâtriers, caissières ou vendeuses, serveuses ou barman, auxiliaires, employées de bureau ou secrétaires... Et si çà et là surgissent une "éducatrice" ou une "professeur de lettres" ou une "informaticien", "musicien" ou "zootéchnicien", une "secrétaire de direction", une "styliste" ou une "dessinatrice de publicité", il s’agit parfois à l’évidence d’un habillement d’emprunt nous rappelant opportunément que l’"interaction enquêteur-enquêté" se déroule toute entière dans l’ambivalence entre la conscience et la représentation, entre la règle et le conflit.
Pour que le décor soit planté, reste à ajouter que l’enquêteur est en général une enquêtrice, dont la situation professionnelle certes précaire ne reflète pas en général le statut social, hérité davantage de celui de son conjoint, et situé la plupart du temps dans "les positions moyennes" du même espace.

La difficulté de la codification ne tient pas ici aux difficultés d’étiquetage de "ces inclassables de nos sociétés [...] qui ne sont ni chair ni poisson", ni aux luttes sociales autour de l’homologation, pour produire des "classements sociaux juridiquement garantis", pain quotidien des statisticiens publics, qui appartiennent à la "banque centrale qui produit tous les certificats". Non, d’une certaine façon les choses ici sont plus simples, et le dénouement relève de la formule de WEBER "Les agents sociaux obéissent à la règle quand l’intérêt à lui obéir l’emporte sur l’intérêt à lui désobéir".

Il suffit en effet d’avoir une expérience un peu prolongée de l’espace social non plus abstraitement construit mais géographiquement réel de ces enquêtés et d’avoir suivi quelques enquêtrices sur le terrain pendant le déroulement de l’enquête emploi, pour imaginer aisément que chacun des deux interlocuteurs a le sens à la fois de sa propre place et de celle de l’autre, qu’ils sont vraisemblablement tous deux pressés d’en finir, et comme "la logique de la pratique, c’est d’être logique jusqu’au point où être logique cesserait d’être pratique"...

4 - Un essai de "bouclage" quantitatif

Dans cette dernière partie, nous partirons des chiffres bruts de l’enquête pour retrouver les chiffres bruts du recensement. La correction de concept a été examinée en 1ère partie. Une correction de champ plus précise, et une correction de date permettront d’isoler l’"effet protocole" sur les répondants à l’EE, le "bouclage" s’achevant par une estimation du biais dû aux non réponses de l’enquête et la prise en compte de l’alea de sondage. Le bouclage avec l’exploitation exhaustive du RP se fait à l’intérieur de la marge admissible, ce qui valide ex post la matrice de passage observée sur l’échantillon. Le bouclage avec l’exploitation lourde est plus délicat, car la matrice directement observée sur l’échantillon de 9979 adultes est non représentative, et doit être reconstituée indirectement. Ce détour permet néanmoins d’achever le second bouclage dans des conditions satisfaisantes, de fournir des indications précieuses sur les reclassements opérés par l’exploitation lourde ("COLIBRI") et de calculer des coefficients de redressement pour chacune des 9 cases des matrices observées (cf supra 3e partie).
4.1. Correction de champ

L’enquête emploi fait l’objet d’un "calage démographique" (par sexe et âge détaillé) sur le recensement dans sa publication définitive. Dans la nouvelle série d’enquêtes depuis 1990, conformément aux recommandations de [COEFFIC 18] à la suite du précédent rapprochement, on considère les élèves internes, les militaires du contingent et les militaires de carrière "rattachés" aux ménages comme faisant partie de la "population des ménages au sens du recensement". Le "calage" consiste alors à assimiler l’ensemble des résidents à l’enquête et de ces trois catégories de "rattachés" à la population des ménages du recensement actualisée à la date de l’enquête. Les autres catégories de rattachés, théoriquement hors champ, ne font l’objet d’aucun calage. Les chiffres du tableau III que nous avons présenté en 1er partie s’appuient sur ce calage.

De fait les trois catégories précédentes de rattachés sont massivement bien appariées dans notre échantillon (à 85 % environ contre 91 % pour les résidents), ce qui justifie l’option prise pour le calage. Mais on constate en outre, si on compare l’équilibre du rapprochement à celui du calage (tableau IX en annexe), que 45 % des "hors calage" théoriques sont également appariés (* dans tous les cas bien entendu les pourcentages sont une estimation par défaut des proportions réelles, compte tenu de l’existence de logements non retrouvés manuellement mais pas nécessairement non recensés en réalité, et des déménagements entre les deux dates *).

Les 55 % restant sont presqu’exclusivement (à 97 %) des inactifs, principalement des étudiants en cité universitaire, et secondairement des personnes en établissements de soins. Nous avons donc opéré un calage personnalisé empirique, avec pour hypothèse que la population active totale à l’enquête peut être assimilée à la population active des ménages au recensement, et en ajustant la population inactive de 15 ans ou plus de l’enquête sur celle des ménages au recensement (avec une légère correction pour tenir compte de 16 000 décès entre les deux dates). Cette rectification a pour avantage de réduire l’écart global sur la population du champ à l’évolution démographique, et d’introduire une contrainte de compensation entre les trois postes.

Les nouvelles estimations à concept et champ comparables après arrondis au dix millième, sont donc pour l’enquête de 22 360 000 actifs occupés, 2 440 000 chômeurs et 19 890 000 (calage sur l’exhaustif) ou 19 920 000 (calage sur le sondage au quart) inactifs.

4.2. Correction de date

Même si les flux saisonniers entre catégories peuvent être très importants et très difficiles à estimer (* la tentative d’utiliser le calendrier mensuel de l’enquête emploi 1991 a révélé son manque de fiabilité *), les variations du niveau de chacune des
catégories entre fin décembre et fin février 1990 peuvent être assez correctement estimés à la dizaine de millier près à partir de sources conjoncturelles. Nous avons retenu un accroissement de 30 000 actifs occupés, et une diminution de 40 000 chômeurs et de 10 000 inactifs entre les deux dates ( * nous avons imputé entièrement aux inactifs le déficit démographique de 20 000 personnes dans le champ * ).

En définitive, les écarts d’estimations à concept, champ et date comparables sur la population active occupée sont totalement admissibles entre l’enquête et l’exploitation exhaustive du recensement ( + 50 000 à l’EE par rapport au RP) et à la rigueur acceptables par rapport à l’exploitation lourde ( + 150 000) compte tenu de l’alea de sondage de l’enquête. Mais les écarts sur le nombre d’inactifs (respectivement + 220 000 et + 170 000) et surtout le nombre de chômeurs ( - 270 000 et - 320 000) sont nettement hors de la marge d’incertitude admissible.

4.3. Les appariés entre l’enquête et le "RP exhaustif" et la mesure globale de l’"effet protocole"

Par application à la population totale de l’enquête, arrondi des cases au dix millième et introduction des décès, la matrice de passage ci-dessus du rapprochement "microsocial" des 40 117 individus, en partant des effectifs mesurés avec le codage simulé (vecteur TACT3EE), donne les effectifs qu’on aurait obtenus en faisant passer aux répondants le protocole du recensement exhaustif (vecteur TACT3H), en janvier et en mars en milliers :

Tableau 7

L’effet protocole "RP exhaustif" sur les répondants a l’EE

<table>
<thead>
<tr>
<th>Matrice observée pondérée (milliers)</th>
<th>Retropolation</th>
<th>Cible total RPExh</th>
</tr>
</thead>
<tbody>
<tr>
<td>TACT3EE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACT3H</td>
<td>occupés</td>
<td>Chômeurs</td>
</tr>
<tr>
<td>occupés</td>
<td>21440</td>
<td>400</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>360</td>
<td>1750</td>
</tr>
<tr>
<td>Inactifs</td>
<td>560</td>
<td>290</td>
</tr>
<tr>
<td>Total(EE)</td>
<td>22360</td>
<td>2440</td>
</tr>
</tbody>
</table>

(La dernière ligne des trois premières colonnes de chiffres donne les effectifs observés à l’enquête avec le codage TACT3EE ; la quatrième colonne de chiffres donne les effectifs avec le nouveau codage TACT3H sans correction de date, la cinquième les réactualise en janvier, et la dernière rappelle les résultats du recensement exhaustif. Les arrondis et les décès sont cause de déséquilibres comptables de +/-10 000 personnes)
L'écart entre l'avant-dernière colonne et la dernière (demi) ligne de ce tableau mesure l'effet purement dû au protocole (sur les répondants de l’enquête) : le codage à l’exploitation exhaustive du recensement (TACT3H) "produit" donc 50 000 occupés et 170 000 chômeurs en plus, 210 000 inactifs en moins que le codage supposé équivalent de l’enquête (TACT3EE).

L’écart entre la dernière colonne et la quatrième colonne de chiffres mesure le solde encore inexplicable entre les deux sources, soit un excédent net de 100 000 actifs occupés et un déficit net de 100 000 chômeurs à l’enquête par rapport au recensement. Ce dernier écart est encore un peu au delà de la marge admissible due à l'alea de sondage. L’écart global sur la population active est quasi-intégralement expliqué par l’effet de protocole sur les répondants, à champ, concept et date comparables.

4.4. Les non-appariés et les biais éventuels dus aux non-réponses

Une partie de l’écart peut tenir naturellement à des biais éventuels introduits par les non-répondants à l’enquête. Les 2410 adultes du rapprochement, recensés dans des résidences principales non enquêtées, auxquels on peut éventuellement ajouter les 943 recensés dans des résidences déclarées comme non principales par l’enquêteur (* à tort ou à raison, il est difficile d’en décider *), nous en fournissent un échantillon représentatif (cf tableau IX).

Or, contrairement aux conclusions du rapprochement de 1982 sur ce point [COEFFIC 18], la structure par type d’activité (TACT3) des 3353 non enquêtés, et plus encore celle des 2 410 non répondants stricts, s’écarte assez sensiblement de celle des enquêtés : parmi eux plus de chômeurs (respectivement 7,0 % et 7,2 % au lieu de 5,0 %), et plus d’actifs occupés (53,6 % et 55,7 % contre 45,4 %).

Nous avons cherché à évaluer le biais dû à la structure particulière des non-réponses. Compte tenu du mécanisme d’extrapolation de l’enquête à trois niveaux successifs (* REDECHA redressement de champ géographique, EXTRL redressement des logements non enquêtés, et EXTRI redressement des individus enquêtés par calage démographique *), il est pratiquement impossible de le faire sur les données calées qui corrigent assez largement ce biais car ces non enquêtés ont un profil démographique particulier (ce sont comme dans toutes les enquêtes des personnes vivant souvent seules, dans des grands centres urbains, et en général des moins de 35 ans).

En revanche on peut calculer le biais dû aux non-réponses avant calage démographique (* la méthode consiste à substituer les non-répondants aux répondants redressés entre les deux premiers niveaux *) : il serait de l’ordre de -100 000 actifs occupés, -50 000 chômeurs et + 150 000 inactifs à l’enquête si on base le calcul sur l’ensemble des non
enquêtés recensés, de -180 000 actifs occupés, -60 000 chômeurs et + 240 000 inactifs sur la base des seuls non répondants stricts.

Fort heureusement le calage démographique de l’enquête opère un lissage qui va "dans le bon sens" puisqu’il ajoute 333 000 actifs occupés et 23 000 chômeurs, et ôte 82 000 inactifs aux résultats non calés. Il subsisterait cependant un déficit sur les chômeurs d’au minimum 30 000 personnes (mais dans tous les cas inférieur à 60 000). Le biais résiduel après calage dû aux non réponses se situe, si on applique le même raisonnement, entre 0 et + 230 000 personnes pour les actifs occupés. Une étude plus détaillée de l’échantillon des 3 353 non enquêtés serait nécessaire pour affiner ces fourchettes d’estimation.

La prise en compte des biais dus aux non réponses permet d’achever le bouclage : le déficit final de chômeurs à l’enquête serait alors de 70 000 au maximum (encore réduit à 30 000 si le biais dû aux non réponses est estimé à -60 000), à l’intérieur de la marge admissible due aux alea de sondage. Pour les actifs occupés, quelle que soit l’hypothèse retenue sur le biais dû aux non réponses, les écarts (entre + 100 000 et -230 000) restent également dans la marge admissible (* intervalle de confiance à 95 %.*).

On trouvera dans le tableau X en annexe un résumé de notre proposition de bouclage entre l’enquête et l’exploitation exhaustive du RP (l’hypothèse retenue sur les biais dûs aux non réponses est de -30 000 chômeurs et de 0 actifs occupés).

4.5. Essai de bouclage entre l’enquête et l’exploitation lourde du recensement

L’absence de représentativité de la matrice de passage entre le type d’activité à l’enquête (TACT3EE) et le type d’activité dans l’exploitation lourde du recensement (TACT3) directement observée sur les 9979 adultes "appariés" sûrs" figurant dans le sondage au quart oblige à une estimation indirecte de cette matrice dont on trouvera un résumé dans l’encadré technique. Nous nous contenterons ici des résultats du bouclage.

Par application à la population totale de l’enquête, arrondi des cases au dix millième et introduction des décès, cette nouvelle matrice de passage en partant des effectifs mesurés avec le codage simulé (vecteur TACT3EE), donne les effectifs qu’on aurait obtenus en faisant passer aux répondants le protocole de l’exploitation lourde du recensement (vecteur TACT3), en janvier et en mars en milliers :

La collecte
**Tableau 8**

*L’effet protocole "RP lourd" sur les répondants à l’EE*

<table>
<thead>
<tr>
<th>Matrice reconstituée pondérée (milliers)</th>
<th>Retropolation</th>
<th>Cible total RPQAR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TACT3</strong></td>
<td><strong>occupés</strong></td>
<td><strong>Chômeurs</strong></td>
</tr>
<tr>
<td>Occupés</td>
<td>21210</td>
<td>420</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>430</td>
<td>1710</td>
</tr>
<tr>
<td>Inactifs</td>
<td>720</td>
<td>310</td>
</tr>
<tr>
<td><strong>Total (EE)</strong></td>
<td>22360</td>
<td>2440</td>
</tr>
</tbody>
</table>

(La dernière ligne des trois premières colonnes de chiffres donne les effectifs observés à l’enquête avec le codage TACT3EE ; la quatrième colonne de chiffres donne les effectifs avec le nouveau codage TACT3 sans correction de date, la cinquième les réactualise en janvier, et la dernière rappelle les résultats du recensement au quart. Les arrondis et les décès sont cause de déséquilibres comptables de + /-10 000 personnes)

L’écart entre l’avant-dernière colonne et la dernière (demi) ligne de ce tableau mesure l’effet purement dû au protocole (sur les répondants de l’enquête) : le codage dans l’exploitation lourde du recensement (TACT3) "produit" donc 200 000 chômeurs en plus, 50 000 occupés et 160 000 inactifs en moins que le codage supposé équivalent de l’enquête (TACT3EE).

L’écart entre la dernière colonne et la quatrième colonne de chiffres mesure le solde encore inexpliqué entre les deux sources, soit un déficit net de 100 000 actifs occupés et un excédent net de 120 000 chômeurs à l’enquête par rapport au recensement.

Si on maintient les fourchettes d’estimation des biais dus aux non réponses à l’enquête on peut achever le bouclage en laissant inexpliqués un déficit maximum de 90 000 chômeurs se situant à la borne inférieure de la marge d’incertitude admissible due aux alea cumulés de l’enquête et du quart ( * ce dernier modifie très peu la somme des variances, on peut admettre au risque de 5 % un écart de +/-90 000 chômeurs [DEVILLE et ROTH 14] * ). et des écarts sur les actifs occupés (identiques à ceux du bouclage précédent) à l’intérieur de cette marge.

En définitive, on trouvera dans le tableau X en annexe, au dessous du bouclage entre l’enquête et l’exhaustif, notre proposition de bouclage entre l’enquête et le quart.
Conclusion : un halo chargé de sens

Nous donnerons à nos conclusions une visée surtout opérationnelle. Nous restituerons les difficultés de codage dans le contexte socio-historique des années 1980, et ferons quelques suggestions aux fins de renforcer la cohérence entre les résultats obtenus par les deux protocoles différents de collecte.


En six ans la situation n’a cessé de renforcer la pertinence de ce constat.


Pour autant, puisque l’essentiel de notre analyse a porté le débat sur le partage chômage/inactivité, et montré que les chômeurs "exclus des statistiques", loin d’être aléatoirement distribués, se trouvent à un pôle identifiable de l’espace social, surtout parmi les femmes d’âge moyen peu diplômées et faiblement qualifiées, nous ne pensons pas qu’on puisse se résoudre à une dérive aussi ségrégative. Deux remarques suffiront à illustrer ce propos.

La collecte
Concernant le regard que les statisticiens portent à l'activité féminine, une étude historique [FOUQUET et CHARRAUD 35] a montré qu'elle leur révèle toujours des surprises, mais contribue aussi à "préciser et parfois corriger le regard que la société porte sur elle-même". Dans "Données Sociales 1990" (* on ne disposait pas alors des chiffres de l'année 1989 *) n'expliquait-on pas la stabilisation du chômage de 1985 à 1988 par une baisse probable de l'activité féminine [LACROIX 22] ? Or nous avons constaté au recensement de 1990, et c'est une nouveauté, une propension forte des femmes à déclarer rechercher un emploi. Il apparaît en effet à la lecture des seules traces laissées par le rapprochement de 1982 qu'en 1975 comme en 1982, les femmes qui ne s'étaient pas déclarées au chômage et cherchaient néanmoins du travail, d'après l'enquête emploi, étaient majoritairement classées inactives dans le recensement. La question sur la recherche d'emploi y était pourtant déjà présente (moins "visible" cependant). Une telle inflexion des comportements (ne concernant pas d'abord les plus jeunes), n'est-elle pas de nature à "alimenter le débat social" ?

Quant aux populations démunies, leur disparition des statistiques du chômage BIT par auto-éviction des intéressés ne risque-t-elle pas de donner de la structure du marché du travail une fausse appréciation et de masquer l'extension des phénomènes d'exclusion sociale dont beaucoup d'observateurs notent des signes tangibles autour de nous ? Ainsi le numéro-dossier cité avançait-il que le recul du chômage entre 1987 et 1990 s'était accompagné d'une réduction des disparités ayant bénéficié aux catégories les moins qualifiées ou les moins diplômées [DUMARTIN et MARCHAND 26].

Puisque le BIT ne va pas jusqu'à régler dans ses détails le protocole de collecte de l'enquête sur l'emploi laissé pour une part à l'appréciation nationale, ne peut-on au moins insister davantage dans la formation des enquêteurs sur les biais systématiques que comportent un passage trop rapide sur les questions de début d'enquête (avec les filtres irrémédiables que cela entraîne), et un abandon des non répondants (dont nous avons vu également supra qu'on trouvait davantage de chômeurs parmi eux) ? Enfin la restriction de l’enquête à la population des ménages est sans doute un problème (bien que de conséquences limitées) quand on connaît le lien entre l’absence d’emploi et la disposition d’un logement en propre.

Du côté du questionnaire du recensement nous suggérons une modification de présentation de la question 9 allant dans le sens d'une plus grande clarification de concepts que nous cherchons par ailleurs à vulgariser. On peut en effet regretter que la question 9 (INAC) du bulletin individuel ne précède pas le filtre dirigeant les personnes exerçant une activité professionnelle vers le verso du bulletin : ainsi cette question pourrait, comme son homologue à l'enquête (FI), comporter une modalité "travaille", et éviterait de regrouper les chômeurs avec des catégories d’inactifs en les isolant des actifs occupés. Avantage supplémentaire, une question demeurée "à blanc" serait identifiée sans ambiguïté comme relevant des non-réponses (* il subsisterait néanmoins in fine des "sans objet", à cause des réintégrés remplissant un autre bulletin * ).
Encadré technique sur l’estimation de la matrice de passage entre l’enquête et l’exploitation lourde du recensement

La matrice de passage entre le type d’activité à l’enquête (TACT3EE) et dans l’exploitation lourde (TACT3) observée directement sur les 9 979 adultes appariés "sûrs", figurant au sondage au quart est la suivante :

Tableau 9

<table>
<thead>
<tr>
<th>Type d’activité au quart (TACT3)</th>
<th>A l’enquête- (tact3ee) (matrice observée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs occupés</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>4832</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>79</td>
</tr>
<tr>
<td>Inactifs</td>
<td>119</td>
</tr>
</tbody>
</table>

Par rapport à la matrice de passage entre l’enquête et l’exhaustif (cf texte), la case "Inactifs RP/Chômeurs EE" surprend par son effectif de 42 personnes anormalement bas (la case correspondante à l’exhaustif est de 244 = 4x61), surtout si on considère qu’en regard la case "Chômeurs RP/Inactifs EE" est restée "intouchée" (105 au lieu de 419 = 4x105).

La démonstration de la non représentativité de la matrice directement observée sur les 9 979 personnes peut être faite inductivement. Nous disposons sur ces 9 979 personnes du croisement des trois codes (TACT3EE, TACT3H et TACT3). Or on ne retrouve dans l’échantillon que 44 des 244 "chômeurs RP/Inactifs EE" qui figuraient à l’exhaustif, conduisant après reclassements infimes à leurs 42 homologues dans notre matrice (TACT3EE/TACT3), rescapés du "passage à la trappe" de 200 personnes.

Nous pouvons cependant tourner cette difficulté, en estimant cette matrice indirectement : notre opération de rapprochement a nécessité l’extraction d’un échantillon de districts, "surzonage" des aires (ou grappes) du tiers sortant, représentant 297 135 adultes de la population des ménages dont 74 232 figurant au quart, qui peut être considéré comme un sondage du recensement représentatif France entière ( * en tenant compte, dans le cas des aires, d’une probabilité de tirage inégale des individus inversement proportionnelle à la taille des districts. * ) Nous disposons sur 74 221 adultes de la population des ménages des trois codes d’activité, spontané (INAC3), à l’exhaustif (TACT3H) et au quart (TACT3), ce qui est particulièrement précieux pour juger de la codification au RP. (* nous avons utilisé cet échantillon dans la 1ère partie du texte pour rendre compte du passage du classement spontané INAC3 au codes complexes TACT3H ou TACT3 * )
On obtient ainsi la matrice de passage du codage du type d'activité à l'exhaustif au code définitif de l'INSEE dans le sondage au quart, soit en pourcentages en colonnes :

**Tableau 10**

<table>
<thead>
<tr>
<th>Type d'activité au quart (TACT3)</th>
<th>A l'exhaustif (TACT3H) (matrice observée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs occupés</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>98,88</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>0,36</td>
</tr>
<tr>
<td>Inactifs</td>
<td>0,76</td>
</tr>
<tr>
<td>Total</td>
<td>100,00</td>
</tr>
</tbody>
</table>

Il n'est pas dans notre propos ici d'analyser en détail les recodifications opérées par l'exploitation lourde. On voit qu'elles sont très limitées en pourcentage et portent surtout sur les reclassements de chômeurs en actifs occupés (1,2 %). On peut estimer la matrice de passage entre l'enquête et le sondage au quart par composition de la matrice de passage entre l'enquête et l'exploitation exhaustive (observée sur 40 117 appariés "sûrs") et cette dernière matrice (observée sur 74 221 adultes de la population des ménages). On obtient en pourcentages en colonnes (entre parenthèses leurs homologues dans la matrice TACT3EE/TACT3H) :

**Tableau 11**

<table>
<thead>
<tr>
<th>Type d'activité au quart (TACT3)</th>
<th>A l'enquête- (TACT3EE) (matrice reconstituée)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actifs</td>
</tr>
<tr>
<td>Actifs occupés</td>
<td>94,83 (95,87)</td>
</tr>
<tr>
<td>Chômeurs</td>
<td>1,93 (1,61)</td>
</tr>
<tr>
<td>Inactifs</td>
<td>3,24 (2,52)</td>
</tr>
<tr>
<td>Total</td>
<td>100,00</td>
</tr>
</tbody>
</table>

On voit que la recodification ne semble pas avoir renforcé la cohérence avec l'enquête, bien au contraire (à l'exception de la case "Chômeurs RP/Inactifs EE" inchangée). Une fois traduite en vraie grandeur (cf fin du texte), notre matrice (contrairement à la matrice observée directement) rend bien compte des tendances observées entre le quart et l'exhaustif : écart accru avec l'enquête d'environ 100 000 personnes sur les actifs occupés, 30 000 sur les chômeurs, diminué d'environ 50 000 sur les inactifs.

182

*INSEE Méthodes n° 52 - 53*
Sur l’enquête Emploi de 1990 :

1) Conception de l’enquête, collecte et traitement des données :

in "Courrier des Statistiques" N°53, INSEE, 1990 :


2) Résultats publiés :


3) Précision des résultats :

in "Economie et Statistique" N°193-194, INSEE, 1986 "Emploi et chômage : l’éclatement" :


Sur le recensement de la population de 1990 :

1) Collecte et chiffrement-saisie des données :

2) Traitement des fichiers et codification :

[56] Recensement général de la population de 1990 ; Guide d’utilisation tome 3 ;
description des traitements exploitation exhaustive INSEE 1993 ;

[57] Recensement général de la population de 1990 ; Guide d’utilisation tome 5 ;
description des traitements exploitation lourde INSEE 1993 ;

3) Résultats publiés :

population de 1990" in "INSEE Premiere" N°149, 1991 ;

ments FRANCE", INSEE, 1991 (Résultats de l’exploitation exhaustive) ;

Population active" (Résultats de l’exploitation au 1/4) ; INSEE 1992 ;

4) Contrôle de qualité :

"Actes de journées de méthodologie statistique des 17 et 18 juin 1992" ; INSEE
Méthodes N°46-47-48 ;

1990" ;

[14] J-C. DEVILLE "Échantillonnages pour le contrôle de qualité du recensement de
1990" ;

[15] G.BADEYAN "Le contrôle de production des phases de saisie-chiffrement du
recensement de 1990" ;
Sur les récents rapprochements "Enquête Emploi-Recensement" :

1) Pour 1975 :

in "Archives et Documents" N°38, INSEE, 1981 "Les catégories socio-professionnelles et leur repérage dans les enquêtes ;

[16] M. CÉZARD "Le rapprochement de l’enquête Emploi et du recensement de la population de 1975 ; résultats généraux" ;

[17] M. CÉZARD "Les chômeurs selon le recensement et l’enquête Emploi de 1975 ; examen de bulletins individuels" ;

2) Pour 1982, en l’absence de publication :

[18] N. COEFFIC "Enquête Emploi : quelques points de méthodologie" Note interne INSEE Direction de la Production, Division collecte, 1er juillet 1988 ;

3) Pour 1990, en préalable au présent travail :


Sur les différents concepts d’emploi, de chômage et de population active, les notions de "halo" et de "continuum de situations", la problématique du "bouclage" et les projections de "ressources en main d’œuvre" dans la conjoncture de la fin des années 1980 :


[21] C. THELOT introduction ;
in "Données Sociales 1990" INSEE, 1990 :


in "Économie et statistique" N°243, INSEE, 1991

[23] D. BLANCHET et O. MARCHAND "Au delà de l’an 2000, s’adapter à une pénurie de main-d’œuvre" ;

in "Économie et Statistique" N°249, INSEE, 1991 "Marché du Travail" :


Sur les différentes phases du travail statistique, les notions d’observation, de nomenclatures, de codification, de données, de construction et d’objectivation, de qualité, de "redressement des non-réponses", d’analyse et d’interprétation des données, bref de la problématique de la "mesure" en statistiques sociales au sein d’un "institut public doté de légitimité" :

[27] M. VOLLE "Le métier de statisticien" Hachette Littérature, 1980 ;


Sur les rapports entre statistique et sociologie, la suspicion du statisticien vis à vis de l'interprétation, la nécessité de "contextualiser", le dépassement du dilemme objectivisme/subjectivisme, les classements du sujet objectivant et ceux des agents sociaux, la cohérence logique et la logique du parti-pris, l'espace social et l'habitus...

in "Économie et Statistique" N°168, INSEE, 1984 "Sociologie et statistique" :

[38] F. DE SINGLY "Les bons usages de la statistique dans la recherche sociologique" ;

[39] F. HERAN "L'assise statistique de la sociologie" ;

parmi les communications présentées à la journée d'études "Statistiques et sociologie" Société française de sociologie, INSEE, Paris 1982 :

[40] J-C. PASSERON "Ce que dit un tableau et ce qu'on en dit ; remarques sur le langage des variables et l'interprétation dans les sciences sociales" (repris dans "Le raisonnement sociologique" cf infra).
in "L’ENSAE 25 ans après : que devient l’école et que deviennent les anciens élèves ?" brochure ENSAE, 1986 :

[41] interview de P. BOURDIEU ;

in P. BOURDIEU "La distinction" Collection "Le sens commun" Editions de Minuit, 1979

[42] Conclusion : "Classes et classements" ;

in P. BOURDIEU "Choses dites" Collection "Le sens commun" Editions de minuit, 1987

[43] "La codification" ;

[44] "Objectiver le sujet objectivant" ;

[45] "De la règle aux stratégies" ;

[55] "Espace social et pouvoir symbolique" ;

in "Mélanges en l’honneur de C. DESABIE", INSEE, 1989 :

[37] C. SEIBEL "Statistiques sociales et recherche" ;

### Tableau I

**COMPARAISON DES ESTIMATIONS DE POPULATIONS AUX DEUX SOURCES**

**RECENSEMENT MARS 1990** | **ENQUETE EMPLOI JANVIER 1990**

#### POPULATION DES PLUS DE 15 ANS (CHAMP MOINS ETENDU POUR L’ENQUETE)

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>EXHAUSTIF TOTAL</th>
<th>CALEE CHAMP TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>HOMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>4212579</td>
<td>2154700</td>
</tr>
<tr>
<td>20 A 24</td>
<td>4281745</td>
<td>2166156</td>
</tr>
<tr>
<td>25 A 29</td>
<td>4309311</td>
<td>2153018</td>
</tr>
<tr>
<td>30 A 39</td>
<td>8548215</td>
<td>4268131</td>
</tr>
<tr>
<td>40 A 49</td>
<td>7317343</td>
<td>2153018</td>
</tr>
<tr>
<td>50 A 54</td>
<td>3011167</td>
<td>1468370</td>
</tr>
<tr>
<td>60 OU +</td>
<td>11299373</td>
<td>4706241</td>
</tr>
<tr>
<td></td>
<td>45850910</td>
<td>22056711</td>
</tr>
</tbody>
</table>

#### POPULATION ACTIVE (CHAMP MOINS ETENDU ET DEP. BIT POUR L’ENQUETE)

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>EXHAUSTIF TOTAL</th>
<th>CALEE CHAMP TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>HOMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>514410</td>
<td>320966</td>
</tr>
<tr>
<td>20 A 24</td>
<td>2732100</td>
<td>1463926</td>
</tr>
<tr>
<td>25 A 29</td>
<td>3748363</td>
<td>2013391</td>
</tr>
<tr>
<td>30 A 39</td>
<td>6258284</td>
<td>3580890</td>
</tr>
<tr>
<td>40 A 49</td>
<td>2212572</td>
<td>1304007</td>
</tr>
<tr>
<td>50 A 54</td>
<td>1749157</td>
<td>1027859</td>
</tr>
<tr>
<td>60 OU +</td>
<td>721285</td>
<td>385451</td>
</tr>
<tr>
<td></td>
<td>25341950</td>
<td>14235934</td>
</tr>
</tbody>
</table>

#### CATEGORIE

| TOTAL      | 25341950        | 14235934          | 11106016 |

#### ACT.00CU.

| TOTAL      | 22379569        | 12875480          | 9504089  |

#### MIL.CONT.

| TOTAL      | 229810          | 228258            | 1552     |

#### CHOMEUR

| TOTAL      | 2732571         | 1132196           | 1600375  |

#### TAUX D'ACTIVITE (CHAMP MOINS ETENDU ET DEP. BIT POUR L’ENQUETE)

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>EXHAUSTIF TOTAL</th>
<th>CALEE CHAMP TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>HOMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>55,27%</td>
<td>64,54%</td>
</tr>
<tr>
<td>20 A 24</td>
<td>63,81%</td>
<td>67,58%</td>
</tr>
<tr>
<td>25 A 29</td>
<td>86,90%</td>
<td>93,51%</td>
</tr>
<tr>
<td>30 A 39</td>
<td>86,64%</td>
<td>96,98%</td>
</tr>
<tr>
<td>40 A 49</td>
<td>85,53%</td>
<td>96,59%</td>
</tr>
<tr>
<td>50 A 54</td>
<td>77,06%</td>
<td>91,01%</td>
</tr>
<tr>
<td>60 OU +</td>
<td>58,09%</td>
<td>70,00%</td>
</tr>
<tr>
<td></td>
<td>55,27%</td>
<td>64,54%</td>
</tr>
</tbody>
</table>

#### RATIO DE CHOMAGE (CHAMP MOINS ETENDU ET DEP. BIT POUR L’ENQUETE)

| TOTAL      | 10,78%          | 7,95%             | 14,41%  |

La collecte 189
### Tableau II

**COMPARAISON DES ESTIMATIONS DE POPULATIONS AUX DEUX SOURCES**

**RECENSEMENT MARS 1990**

**ENQUETE EMPLOI JANVIER 1990**

**POPULATION DES PLUS DE 15 ANS (CHAMP MOINS ETENDU POUR L'ENQUETE)**

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>TOTAL</th>
<th>CHAMP</th>
<th>CALÉE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>4124939</td>
<td>2155461</td>
<td>2059478</td>
<td>2155461</td>
</tr>
<tr>
<td>20 A 24</td>
<td>4282196</td>
<td>2168190</td>
<td>2114006</td>
<td>2168190</td>
</tr>
<tr>
<td>25 A 29</td>
<td>4306887</td>
<td>2147717</td>
<td>2159168</td>
<td>2147717</td>
</tr>
<tr>
<td>30 A 39</td>
<td>8559763</td>
<td>4270349</td>
<td>4289414</td>
<td>4270349</td>
</tr>
<tr>
<td>40 A 49</td>
<td>7326802</td>
<td>3711163</td>
<td>3615639</td>
<td>3711163</td>
</tr>
<tr>
<td>50 A 59</td>
<td>2872736</td>
<td>1434130</td>
<td>1438606</td>
<td>1434130</td>
</tr>
<tr>
<td>60 OU +</td>
<td>11301449</td>
<td>4700623</td>
<td>6600826</td>
<td>4700623</td>
</tr>
</tbody>
</table>

**TOTAL**

| 45875366 | 22053634 | 23821732 | 44885312 |

**POPULATION ACTIVE (CHAMP MOINS ETENDU ET REF. RP POUR L'ENQUETE)**

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>TOTAL</th>
<th>CHAMP</th>
<th>CALÉE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>496681</td>
<td>311030</td>
<td>185651</td>
<td>311030</td>
</tr>
<tr>
<td>20 A 24</td>
<td>2725246</td>
<td>1461881</td>
<td>1263365</td>
<td>1461881</td>
</tr>
<tr>
<td>25 A 29</td>
<td>3740190</td>
<td>2008747</td>
<td>1731443</td>
<td>2008747</td>
</tr>
<tr>
<td>30 A 39</td>
<td>6256120</td>
<td>3586225</td>
<td>3257820</td>
<td>3586225</td>
</tr>
<tr>
<td>40 A 49</td>
<td>2206946</td>
<td>1304413</td>
<td>902533</td>
<td>1304413</td>
</tr>
<tr>
<td>50 A 59</td>
<td>1736937</td>
<td>1021785</td>
<td>715152</td>
<td>1021785</td>
</tr>
<tr>
<td>60 OU +</td>
<td>724195</td>
<td>393522</td>
<td>330673</td>
<td>393522</td>
</tr>
</tbody>
</table>

**TOTAL**

| 25286711 | 14230179 | 11056532 | 24797318 |

**CATÉGORIE**

<table>
<thead>
<tr>
<th>TOTAL</th>
<th>CHAMP</th>
<th>CALÉE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT.OCCU.</td>
<td>25286711</td>
<td>14230179</td>
</tr>
<tr>
<td>MIL.CONT.</td>
<td>22270218</td>
<td>12834629</td>
</tr>
<tr>
<td>CHOMEUR</td>
<td>2785018</td>
<td>1165451</td>
</tr>
</tbody>
</table>

**TAUX D'ACTIVITE (CHAMP MOINS ETENDU ET REF. RP POUR L'ENQUETE)**

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>TOTAL</th>
<th>CHAMP</th>
<th>CALÉE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>55,12%</td>
<td>64,53%</td>
<td>46,41%</td>
<td>64,53%</td>
</tr>
<tr>
<td>20 A 24</td>
<td>63,64%</td>
<td>67,42%</td>
<td>59,76%</td>
<td>67,42%</td>
</tr>
<tr>
<td>25 A 29</td>
<td>86,84%</td>
<td>93,53%</td>
<td>80,19%</td>
<td>93,53%</td>
</tr>
<tr>
<td>30 A 39</td>
<td>86,46%</td>
<td>97,01%</td>
<td>75,95%</td>
<td>97,01%</td>
</tr>
<tr>
<td>40 A 49</td>
<td>85,39%</td>
<td>96,63%</td>
<td>73,84%</td>
<td>96,63%</td>
</tr>
<tr>
<td>50 A 54</td>
<td>76,82%</td>
<td>90,96%</td>
<td>62,74%</td>
<td>90,96%</td>
</tr>
<tr>
<td>55 A 59</td>
<td>57,69%</td>
<td>69,70%</td>
<td>46,30%</td>
<td>69,70%</td>
</tr>
<tr>
<td>60 OU +</td>
<td>6,41%</td>
<td>8,37%</td>
<td>5,01%</td>
<td>8,37%</td>
</tr>
</tbody>
</table>

**TOTAL**

| 55,12% | 64,53% | 46,41% |

**RATIO DE CHOMAGE (CHAMP MOINS ETENDU ET REF. RP POUR L'ENQUETE)**

<table>
<thead>
<tr>
<th>GROUPE AGE</th>
<th>TOTAL</th>
<th>CHAMP</th>
<th>CALÉE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOMMES</td>
<td>FEMMES</td>
<td>HOMMES</td>
<td>FEMMES</td>
</tr>
<tr>
<td>15 A 19</td>
<td>11,78%</td>
<td>14,43%</td>
<td>9,01%</td>
<td>14,43%</td>
</tr>
<tr>
<td>20 A 24</td>
<td>63,64%</td>
<td>67,42%</td>
<td>59,76%</td>
<td>67,42%</td>
</tr>
<tr>
<td>25 A 29</td>
<td>86,84%</td>
<td>93,53%</td>
<td>80,19%</td>
<td>93,53%</td>
</tr>
<tr>
<td>30 A 39</td>
<td>86,46%</td>
<td>97,01%</td>
<td>75,95%</td>
<td>97,01%</td>
</tr>
<tr>
<td>40 A 49</td>
<td>85,39%</td>
<td>96,63%</td>
<td>73,84%</td>
<td>96,63%</td>
</tr>
<tr>
<td>50 A 54</td>
<td>76,82%</td>
<td>90,96%</td>
<td>62,74%</td>
<td>90,96%</td>
</tr>
<tr>
<td>55 A 59</td>
<td>57,69%</td>
<td>69,70%</td>
<td>46,30%</td>
<td>69,70%</td>
</tr>
<tr>
<td>60 OU +</td>
<td>6,41%</td>
<td>8,37%</td>
<td>5,01%</td>
<td>8,37%</td>
</tr>
</tbody>
</table>

**TOTAL**

| 11,78% | 14,43% | 9,01% |

**INSEE Méthodes n° 52 - 53**
### Tableau III

**COMPARAISON DES ESTIMATIONS DE POPULATIONS AUX DEUX SOURCES**

**RECENTEMENT MARS 1990**  
**ENQUETE EMPLOI JANVIER 1990**

<table>
<thead>
<tr>
<th>GROUPE ÂGE</th>
<th>TOTAL</th>
<th>HOMMES</th>
<th>FEMMES</th>
<th>CALEE</th>
<th>CHAMP</th>
<th>ASSIMILE</th>
<th>HOMMES</th>
<th>FEMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>44699388</td>
<td>21456448</td>
<td>23242940</td>
<td>44715822</td>
<td>21482972</td>
<td>23232850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 A 19</td>
<td>4138084</td>
<td>2115640</td>
<td>2022444</td>
<td>4135036</td>
<td>2114727</td>
<td>2020309</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 A 24</td>
<td>4086060</td>
<td>2048280</td>
<td>2037880</td>
<td>4085789</td>
<td>2046918</td>
<td>2038871</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 A 29</td>
<td>4213600</td>
<td>2081780</td>
<td>2131820</td>
<td>4214442</td>
<td>2086219</td>
<td>2129223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 A 39</td>
<td>8438992</td>
<td>4182240</td>
<td>4256752</td>
<td>8427947</td>
<td>4181160</td>
<td>4246787</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 A 49</td>
<td>7292888</td>
<td>3631600</td>
<td>3661288</td>
<td>7273995</td>
<td>3632746</td>
<td>3660250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 A 54</td>
<td>2828316</td>
<td>1401120</td>
<td>1427196</td>
<td>2827172</td>
<td>1400472</td>
<td>1426700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55 A 59</td>
<td>2965944</td>
<td>1435032</td>
<td>1530912</td>
<td>2967591</td>
<td>1438309</td>
<td>1529282</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60 OU +</td>
<td>10799104</td>
<td>4556256</td>
<td>6242848</td>
<td>10836609</td>
<td>4581421</td>
<td>6255188</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>POPULATION ACTIVE DES MENAGES (DEF. RP POUR L’ENQUETE)</th>
<th>QUART</th>
<th>CALEE</th>
<th>CHAMP</th>
<th>ASSIMILE</th>
<th>HOMMES</th>
<th>FEMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>24960856</td>
<td>13980312</td>
<td>10980544</td>
<td>24773444</td>
<td>13923657</td>
<td>10849787</td>
</tr>
<tr>
<td>15 A 19</td>
<td>482940</td>
<td>301156</td>
<td>181784</td>
<td>479325</td>
<td>310563</td>
<td>168762</td>
</tr>
<tr>
<td>20 A 24</td>
<td>2659456</td>
<td>1413184</td>
<td>1246272</td>
<td>2587238</td>
<td>1392940</td>
<td>1217609</td>
</tr>
<tr>
<td>25 A 29</td>
<td>3692444</td>
<td>1973464</td>
<td>1718980</td>
<td>3656123</td>
<td>1971661</td>
<td>1684642</td>
</tr>
<tr>
<td>30 A 39</td>
<td>7330776</td>
<td>4088144</td>
<td>3242632</td>
<td>7273995</td>
<td>4078334</td>
<td>3195271</td>
</tr>
<tr>
<td>40 A 49</td>
<td>6191952</td>
<td>3532036</td>
<td>2659916</td>
<td>6175665</td>
<td>3528854</td>
<td>2646811</td>
</tr>
<tr>
<td>50 A 54</td>
<td>2179104</td>
<td>1281848</td>
<td>897256</td>
<td>2186929</td>
<td>1282672</td>
<td>904257</td>
</tr>
<tr>
<td>55 A 59</td>
<td>1715380</td>
<td>1005392</td>
<td>709988</td>
<td>1707293</td>
<td>996207</td>
<td>706586</td>
</tr>
<tr>
<td>60 OU +</td>
<td>708804</td>
<td>385088</td>
<td>323716</td>
<td>711766</td>
<td>385737</td>
<td>326029</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORIE</th>
<th>TOTAL</th>
<th>HOMMES</th>
<th>FEMMES</th>
<th>TOTAL</th>
<th>HOMMES</th>
<th>FEMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT.000U.</td>
<td>22017744</td>
<td>12645848</td>
<td>9371896</td>
<td>22114376</td>
<td>12682386</td>
<td>9431990</td>
</tr>
<tr>
<td>MIL.CONT.</td>
<td>219684</td>
<td>218480</td>
<td>1204</td>
<td>226017</td>
<td>224345</td>
<td>1672</td>
</tr>
<tr>
<td>CHOMEUR</td>
<td>2723428</td>
<td>1115984</td>
<td>1607444</td>
<td>2433051</td>
<td>1016926</td>
<td>1416125</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Taux d'Activité de la Population des Menages (DEF. RP POUR L'EE)</th>
<th>QUART</th>
<th>CALEE</th>
<th>CHAMP</th>
<th>ASSIMILE</th>
<th>HOMMES</th>
<th>FEMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>55,84%</td>
<td>65,16%</td>
<td>47,24%</td>
<td>55,40%</td>
<td>64,81%</td>
<td>46,70%</td>
</tr>
<tr>
<td>15 A 19</td>
<td>11,67%</td>
<td>14,23%</td>
<td>8,99%</td>
<td>11,59%</td>
<td>14,69%</td>
<td>8,35%</td>
</tr>
<tr>
<td>20 A 24</td>
<td>65,09%</td>
<td>60,99%</td>
<td>61,16%</td>
<td>63,32%</td>
<td>66,91%</td>
<td>59,72%</td>
</tr>
<tr>
<td>25 A 29</td>
<td>87,63%</td>
<td>94,80%</td>
<td>80,63%</td>
<td>86,75%</td>
<td>94,51%</td>
<td>79,15%</td>
</tr>
<tr>
<td>30 A 39</td>
<td>86,87%</td>
<td>97,75%</td>
<td>76,18%</td>
<td>86,30%</td>
<td>97,54%</td>
<td>75,24%</td>
</tr>
<tr>
<td>40 A 49</td>
<td>85,65%</td>
<td>97,14%</td>
<td>74,03%</td>
<td>85,52%</td>
<td>97,11%</td>
<td>73,78%</td>
</tr>
<tr>
<td>50 A 54</td>
<td>77,05%</td>
<td>91,49%</td>
<td>62,87%</td>
<td>77,35%</td>
<td>91,59%</td>
<td>63,38%</td>
</tr>
<tr>
<td>55 A 59</td>
<td>57,84%</td>
<td>70,06%</td>
<td>46,38%</td>
<td>57,38%</td>
<td>69,26%</td>
<td>46,20%</td>
</tr>
<tr>
<td>60 OU +</td>
<td>6,16%</td>
<td>8,45%</td>
<td>5,19%</td>
<td>6,57%</td>
<td>8,42%</td>
<td>5,21%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>55,84%</td>
<td>65,16%</td>
<td>47,24%</td>
<td>55,40%</td>
<td>64,81%</td>
<td>46,70%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ratio de Chomage de la Population des Menages (DEF. RP POUR L'EE)</th>
<th>QUART</th>
<th>CALEE</th>
<th>CHAMP</th>
<th>ASSIMILE</th>
<th>HOMMES</th>
<th>FEMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>10,91%</td>
<td>7,98%</td>
<td>14,64%</td>
<td>9,82%</td>
<td>7,30%</td>
<td>13,05%</td>
</tr>
</tbody>
</table>

La collecte
### Tableau IV

<table>
<thead>
<tr>
<th>CAS ET RISQUES DE DIVERGENCE ENTRE LES DEUX SOURCES SELON LE JEU DE REPONSES (EE/EXH; ECH:4017 ADU.SURS; CAS:2372)</th>
<th>SPO.RP</th>
<th>SPO.EE</th>
<th>REC.RP</th>
<th>REC1EE</th>
<th>REC2EE</th>
<th>4 CAS</th>
<th>RSQ.D1</th>
<th>RSQ.1</th>
<th>RS1.D1</th>
</tr>
</thead>
<tbody>
<tr>
<td>INAC5 F15 ANREC RECHE DREC RP/EE</td>
<td>TOTAL</td>
<td>OCCUPE</td>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100,0</td>
<td>5,9</td>
<td>OCU/INA</td>
<td>1,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL OCCUPE</td>
<td>34,9</td>
<td>4,1</td>
<td>INA/OCU</td>
<td>2,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC</td>
<td>33,3</td>
<td>3,9</td>
<td>OCU/INA</td>
<td>2,1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHOM. TOTAL</td>
<td>29,3</td>
<td>32,0</td>
<td>CHO/INA</td>
<td>13,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM.</td>
<td>27,7</td>
<td>31,1</td>
<td>OCU/CHOO</td>
<td>15,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>27,7</td>
<td>12,4</td>
<td>INA/OCU</td>
<td>6,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC CHOM.</td>
<td>13,0</td>
<td>96,0</td>
<td>OCU/CHOO</td>
<td>82,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC CHOM.</td>
<td>13,2</td>
<td>10,0</td>
<td>OCU/CHOO</td>
<td>86,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.INA. OCCUPE</td>
<td>12,9</td>
<td>77,4</td>
<td>INA/OCU</td>
<td>70,9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. BLANC</td>
<td>12,1</td>
<td>75,9</td>
<td>OCU/INA</td>
<td>68,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC CHOM.</td>
<td>11,6</td>
<td>100,0</td>
<td>OCU/CHO</td>
<td>97,1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. OCCUPE</td>
<td>11,5</td>
<td>100,0</td>
<td>OCU/CHO</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC A.INA. BLANC</td>
<td>10,7</td>
<td>74,7</td>
<td>OCU/INA</td>
<td>74,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. OCCUPE</td>
<td>10,6</td>
<td>76,1</td>
<td>OCU/INA</td>
<td>76,1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC A.INA. BLANC</td>
<td>10,6</td>
<td>74,3</td>
<td>OCU/INA</td>
<td>74,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC A.INA. BLANC</td>
<td>10,4</td>
<td>74,2</td>
<td>OCU/INA</td>
<td>74,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. A.INA. TOTAL</td>
<td>9,7</td>
<td>5,7</td>
<td>CHO/INA</td>
<td>2,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. CHOM.</td>
<td>9,2</td>
<td>14,2</td>
<td>CHO/INA</td>
<td>8,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC CHOM. BLANC</td>
<td>9,1</td>
<td>100,0</td>
<td>OCU/CHO</td>
<td>96,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. OCCUPE</td>
<td>7,3</td>
<td>100,0</td>
<td>OCU/CHO</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL OUI CHO/INA</td>
<td>6,5</td>
<td>1,8</td>
<td>INA/OCU</td>
<td>1,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL ETUD. TOTAL</td>
<td>6,2</td>
<td>3,2</td>
<td>INA/INA</td>
<td>2,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. BLANC</td>
<td>5,6</td>
<td>3,4</td>
<td>INA/CHO</td>
<td>1,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. A.INA. TOTAL</td>
<td>5,4</td>
<td>58,9</td>
<td>CHO/INA</td>
<td>53,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL ETUD.</td>
<td>5,4</td>
<td>3,1</td>
<td>INA/OCU</td>
<td>1,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM.</td>
<td>5,0</td>
<td>8,8</td>
<td>CHO/INA</td>
<td>8,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM.</td>
<td>5,0</td>
<td>100,0</td>
<td>CHO/INA</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL OUI</td>
<td>4,6</td>
<td>71,2</td>
<td>CHO/INA</td>
<td>71,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>4,5</td>
<td>100,0</td>
<td>CHO/INA</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL OUI</td>
<td>4,4</td>
<td>100,0</td>
<td>CHO/INA</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL RETR. OCCUPE</td>
<td>4,4</td>
<td>84,6</td>
<td>INA/OCU</td>
<td>82,1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC ETUD.</td>
<td>4,3</td>
<td>17,3</td>
<td>OCU/INA</td>
<td>16,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL RETR. OCCUPE</td>
<td>4,3</td>
<td>84,2</td>
<td>OCU/INA</td>
<td>84,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. OCCUPE</td>
<td>4,2</td>
<td>100,0</td>
<td>CHO/OCU</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA. OUI</td>
<td>4,2</td>
<td>100,0</td>
<td>CHO/OCU</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>4,2</td>
<td>65,6</td>
<td>CHO/INA</td>
<td>65,6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC</td>
<td>4,2</td>
<td>16,9</td>
<td>OCU/INA</td>
<td>16,9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>4,0</td>
<td>100,0</td>
<td>CHO/INA</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>3,9</td>
<td>100,0</td>
<td>CHO/INA</td>
<td>100,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. CHOM.</td>
<td>3,8</td>
<td>60,7</td>
<td>INA/CHO</td>
<td>30,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC ETUD.</td>
<td>3,8</td>
<td>15,7</td>
<td>OCU/INA</td>
<td>15,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL BLANC ETUD.</td>
<td>3,8</td>
<td>15,7</td>
<td>OCU/INA</td>
<td>15,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. CHOM.</td>
<td>3,7</td>
<td>100,0</td>
<td>INA/CHO</td>
<td>51,7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. OCCUPE</td>
<td>3,5</td>
<td>1,0</td>
<td>OCU/INA</td>
<td>0,6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>3,5</td>
<td>47,4</td>
<td>INA/CHO</td>
<td>43,9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>3,4</td>
<td>78,7</td>
<td>INA/CHO</td>
<td>73,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>3,4</td>
<td>100,0</td>
<td>INA/CHO</td>
<td>93,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL ETUD. OCCUPE</td>
<td>3,2</td>
<td>58,6</td>
<td>OCU/INA</td>
<td>58,6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL CHOM. CHOM.</td>
<td>3,1</td>
<td>100,0</td>
<td>OCU/CHO</td>
<td>50,0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL A.INA.</td>
<td>2,8</td>
<td>84,8</td>
<td>INA/CHO</td>
<td>83,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

INSEE Méthodes n° 52 - 53
La collecte
Tableau Vbis

L'EFFET PROTOCOLE PUR SUR LES CAS DIVERGEANT ENTRE EE ET RP (EXH. SUR LES CHOMEURS SPONTANÉS ET LES "AUTRES INACTIFS" SPONTANÉS SELON DÉCLARATIONS AUX QUESTIONS ANALOGUES (ECH:2372 ADU."SURS")

1) "DIVERGEANT" A ANC.RECH.RP(ANRE) >= 3 MOIS NON PONDERES

<table>
<thead>
<tr>
<th>CLAS.SPON.</th>
<th>RECENSEMENT</th>
<th>ENQUETE EMPLOI</th>
<th>INAC5/FI5&quot;CHOMEUR&quot;</th>
<th>&quot;AUTRE INACTIF&quot;</th>
<th>&quot;CHOMEUR+AUT.INA.&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECH/DREC</td>
<td>RECH/DREC</td>
<td></td>
<td>N/O NON</td>
<td>O/N NON</td>
<td>HESI NON</td>
</tr>
<tr>
<td>OUI</td>
<td>14 92</td>
<td>1 90</td>
<td>15 182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;AUTRE INACTIF&quot;</td>
<td>0 1</td>
<td>3 65</td>
<td>3 66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;CHOMEUR&quot; + &quot;AUTRE INACTIF&quot;</td>
<td>14 93</td>
<td>4 155</td>
<td>18 248</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2) "DIVERGEANT" A ANC.RECH.RP(ANRE) >= 3 MOIS PONDERES EN MILLIERS

<table>
<thead>
<tr>
<th>CLAS.SPON.</th>
<th>RECENSEMENT</th>
<th>ENQUETE EMPLOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECH/DREC</td>
<td>RECH/DREC</td>
<td></td>
</tr>
<tr>
<td>OUI</td>
<td>16 102</td>
<td>1 100</td>
</tr>
<tr>
<td>&quot;AUTRE INACTIF&quot;</td>
<td>0 1</td>
<td>3 72</td>
</tr>
<tr>
<td>&quot;CHOMEUR&quot; + &quot;AUTRE INACTIF&quot;</td>
<td>16 104</td>
<td>4 173</td>
</tr>
</tbody>
</table>

194

INSEE Méthodes n° 52 - 53
<table>
<thead>
<tr>
<th>SEXE</th>
<th>AGE</th>
<th>DIPLOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOMME</td>
<td>25A39</td>
<td>TOTAL</td>
</tr>
<tr>
<td>15A24</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>50A59</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>TOTAL</td>
<td>TOTAL</td>
</tr>
</tbody>
</table>
**Tableau VII**

**CAS ET RISQUES DE DIVERGENCE ENTRE LES DEUX SOURCES**

Selon sexe, âge, cat. socio. (EE/QAR: ECH:9979 ADU.SURS; CAS:549)

<table>
<thead>
<tr>
<th>Sexe</th>
<th>TR.AGE</th>
<th>CA.SOC.</th>
<th>% Cas</th>
<th>RSO.DI.</th>
<th>RSO.1 RISQUE RSO.COR ECARTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S (EE) AG5(EE)CS8EERP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RP/EE CHO/INA/INA/CHO DE RSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(* CSEERP)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>TOTAL</strong></td>
<td><strong>TOTAL</strong></td>
<td><strong>TOTAL</strong></td>
<td><strong>TOTAL</strong></td>
<td><strong>TOTAL</strong></td>
</tr>
<tr>
<td>Femme</td>
<td>100,0</td>
<td>5,5 OCU/INA</td>
<td>1,1</td>
<td>0,6</td>
<td>0,5</td>
</tr>
<tr>
<td>Homme</td>
<td>63,2</td>
<td>6,6 OCU/INA</td>
<td>1,5</td>
<td>0,9</td>
<td>0,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>25A39</td>
<td>34,6</td>
<td>6,7 OCU/INA</td>
<td>1,5</td>
<td>0,8</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>30,6</td>
<td>10,3 OCU/INA</td>
<td>2,5</td>
<td>0,8</td>
<td>1,8</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>28,2</td>
<td>9,1 OCU/INA</td>
<td>2,0</td>
<td>0,9</td>
<td>1,1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>26,8</td>
<td>11,8 OCU/INA</td>
<td>3,0</td>
<td>0,8</td>
<td>2,3</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>26,8</td>
<td>10,2 OCU/INA</td>
<td>2,3</td>
<td>1,2</td>
<td>1,1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>23,9</td>
<td>7,6 OCU/INA</td>
<td>0,9</td>
<td>0,3</td>
<td>0,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>17,7</td>
<td>7,5 OCU/INA</td>
<td>1,2</td>
<td>0,3</td>
<td>0,9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>17,1</td>
<td>5,7 INA/OCU</td>
<td>1,0</td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>16,8</td>
<td>6,7 CHO/INA</td>
<td>1,9</td>
<td>1,2</td>
<td>0,7</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>13,7</td>
<td>9,7 OCU/INA</td>
<td>2,2</td>
<td>0,6</td>
<td>1,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>13,5</td>
<td>8,8 OCU/INA</td>
<td>0,5</td>
<td>0,2</td>
<td>0,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>12,8</td>
<td>11,7 OCU/INA</td>
<td>2,1</td>
<td>0,8</td>
<td>2,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>12,2</td>
<td>8,2 OCU/INA</td>
<td>1,6</td>
<td>1,5</td>
<td>0,1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>12,2</td>
<td>6,2 OCU/INA</td>
<td>1,5</td>
<td>0,3</td>
<td>0,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>11,7</td>
<td>13,5 CHO/INA</td>
<td>3,4</td>
<td>2,6</td>
<td>0,9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>10,6</td>
<td>14,4 CHO/INA</td>
<td>4,5</td>
<td>2,6</td>
<td>2,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>10,4</td>
<td>6,5 CHO/INA</td>
<td>1,3</td>
<td>0,5</td>
<td>0,9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>10,4</td>
<td>9,9 CHO/INA</td>
<td>3,0</td>
<td>0,8</td>
<td>2,3</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>10,2</td>
<td>8,1 CHO/INA</td>
<td>2,3</td>
<td>1,5</td>
<td>0,8</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>9,7</td>
<td>20,2 OCU/CHO</td>
<td>0,4</td>
<td>0,0</td>
<td>0,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>8,9</td>
<td>2,3 INA/CHO</td>
<td>0,7</td>
<td>1,1</td>
<td>-0,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>8,2</td>
<td>5,8 CHO/INA</td>
<td>1,7</td>
<td>0,2</td>
<td>1,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7,8</td>
<td>3,1 OCU/CHO</td>
<td>0,6</td>
<td>0,2</td>
<td>0,5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7,7</td>
<td>1,7 INA/OCU</td>
<td>0,3</td>
<td>0,2</td>
<td>0,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7,5</td>
<td>8,6 INA/OCU</td>
<td>0,6</td>
<td>0,6</td>
<td>0,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7,3</td>
<td>19,5 OCU/CHO</td>
<td>0,0</td>
<td>0,0</td>
<td>0,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7,3</td>
<td>15,1 INA/OCU</td>
<td>3,4</td>
<td>1,2</td>
<td>2,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>6,6</td>
<td>5,4 CHO/INA</td>
<td>1,5</td>
<td>0,9</td>
<td>0,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>6,4</td>
<td>16,3 CHO/INA</td>
<td>3,7</td>
<td>2,1</td>
<td>1,6</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>6,2</td>
<td>14,0 OCU/CHO</td>
<td>2,3</td>
<td>0,0</td>
<td>2,5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>6,0</td>
<td>5,5 OCU/INA</td>
<td>0,3</td>
<td>0,5</td>
<td>-0,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,6</td>
<td>8,2 INA/OCU</td>
<td>0,5</td>
<td>0,0</td>
<td>0,5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,5</td>
<td>13,6 CHO/INA</td>
<td>3,6</td>
<td>2,7</td>
<td>0,9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,5</td>
<td>7,6 INA/CHO</td>
<td>1,5</td>
<td>2,7</td>
<td>-1,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,5</td>
<td>10,2 CHO/OCU</td>
<td>1,0</td>
<td>0,0</td>
<td>1,0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,5</td>
<td>3,8 INA/CHO</td>
<td>1,3</td>
<td>2,4</td>
<td>-1,1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,3</td>
<td>1,9 CHO/INA</td>
<td>0,8</td>
<td>1,2</td>
<td>-0,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,3</td>
<td>6,2 OCU/INA</td>
<td>1,1</td>
<td>0,0</td>
<td>1,1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>5,1</td>
<td>5,4 OCU/INA</td>
<td>0,4</td>
<td>0,0</td>
<td>0,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,9</td>
<td>10,1 INA/OCU</td>
<td>0,7</td>
<td>0,0</td>
<td>0,7</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,9</td>
<td>3,3 CHO/OCU</td>
<td>0,4</td>
<td>0,6</td>
<td>-0,2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,9</td>
<td>14,1 OCU/INA</td>
<td>1,0</td>
<td>0,8</td>
<td>0,3</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,7</td>
<td>15,6 CHO/INA</td>
<td>4,2</td>
<td>1,8</td>
<td>2,4</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,7</td>
<td>4,3 OCU/OCU</td>
<td>0,3</td>
<td>0,0</td>
<td>0,3</td>
</tr>
</tbody>
</table>
### Tableau VIII

**Analyse qualitative des cas de divergence entre EE et RP (quart)**

Selon type divergence et diagnostic (EE/QAR; ECH:553 ADU. IDENT.)

Chiffres non redressés

<table>
<thead>
<tr>
<th>Type de divergence RPQUART/EE</th>
<th>Diverg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCU/CHO</td>
<td>OCU/INA</td>
</tr>
<tr>
<td>CHO/OCU</td>
<td>CHO/INA</td>
</tr>
<tr>
<td>INA/OCU</td>
<td>INA/CHO</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnostic</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PB.CODAGE</td>
<td>14</td>
<td>35</td>
<td>2</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>INSEE</td>
<td>2.5</td>
<td>6.3</td>
<td>0.4</td>
<td>3.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Saisie</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Fi</td>
<td>0.0</td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Saisie</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>INAC</td>
<td>0.2</td>
<td>0.9</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Codage</td>
<td>5</td>
<td>11</td>
<td>1</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>TACT3</td>
<td>0.9</td>
<td>2.0</td>
<td>0.2</td>
<td>1.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Codage</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TACT3EE</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Redres.BI</td>
<td>8</td>
<td>13</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Hot deck</td>
<td>1.5</td>
<td>2.4</td>
<td>0.0</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>CHANGEMT.</td>
<td>60</td>
<td>69</td>
<td>62</td>
<td>20</td>
<td>53</td>
</tr>
<tr>
<td>SITUATION</td>
<td>10.9</td>
<td>12.5</td>
<td>11.2</td>
<td>3.6</td>
<td>9.6</td>
</tr>
<tr>
<td>Changemt.</td>
<td>27</td>
<td>11</td>
<td>50</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>probable</td>
<td>4.9</td>
<td>2.0</td>
<td>9.0</td>
<td>2.4</td>
<td>2.0</td>
</tr>
<tr>
<td>Changemt.</td>
<td>33</td>
<td>58</td>
<td>12</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>possible</td>
<td>6.0</td>
<td>10.5</td>
<td>15.2</td>
<td>1.3</td>
<td>7.6</td>
</tr>
<tr>
<td>FLOU</td>
<td>5</td>
<td>23</td>
<td>15</td>
<td>68</td>
<td>52</td>
</tr>
<tr>
<td>DECLARAT.</td>
<td>0.9</td>
<td>4.2</td>
<td>2.7</td>
<td>12.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Doub. sit</td>
<td>4</td>
<td>14</td>
<td>1</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>plausible</td>
<td>0.7</td>
<td>2.5</td>
<td>0.2</td>
<td>0.7</td>
<td>4.0</td>
</tr>
<tr>
<td>Incohere.</td>
<td>1</td>
<td>9</td>
<td>14</td>
<td>54</td>
<td>30</td>
</tr>
<tr>
<td>réponses</td>
<td>0.2</td>
<td>1.6</td>
<td>2.5</td>
<td>11.6</td>
<td>5.4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>79</td>
<td>127</td>
<td>79</td>
<td>105</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>14.3</td>
<td>23.0</td>
<td>14.3</td>
<td>19.0</td>
<td>21.2</td>
</tr>
</tbody>
</table>

La collecte
### Tableau IX

**LE CALAGE DE L’ENQUETE EMPLOI 1990 SUR LE RECENSEMENT**

**CATEGORIES DE POPULATION RECENSEES ET CHAMP DE L’ENQUETE**

<table>
<thead>
<tr>
<th>CATEGORIE</th>
<th>ENQUETE EMPLOI</th>
<th>CATEGORIE</th>
<th>RECENSEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>POPULATION</td>
<td>JANVIER 1990</td>
<td>POPULATION</td>
<td>MARS 1990</td>
</tr>
<tr>
<td>AVANT</td>
<td>APRES</td>
<td>EXHAUSTIF</td>
<td></td>
</tr>
<tr>
<td>CALAGE</td>
<td>CALAGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>44612003</td>
<td>44885609</td>
<td>TOTAL</td>
</tr>
<tr>
<td>POP.CALEE</td>
<td>44441965</td>
<td>44716119</td>
<td>POP. MENAGES</td>
</tr>
<tr>
<td>RESIDENTS</td>
<td>44013100</td>
<td>44284683</td>
<td>RESIDENTS</td>
</tr>
<tr>
<td>ELEVES INTERNES</td>
<td>177893</td>
<td>176649</td>
<td>INTERNES REINT.</td>
</tr>
<tr>
<td>MILI.CONTINGENT</td>
<td>222657</td>
<td>226022</td>
<td>MILITAIR.REINT.</td>
</tr>
<tr>
<td>MILI.CARRIERE</td>
<td>28315</td>
<td>28765</td>
<td></td>
</tr>
<tr>
<td>POP. NON CALEE</td>
<td>170038</td>
<td>169490</td>
<td>POP. COLLECTIV.</td>
</tr>
<tr>
<td>FOY.JEUN.TRAV.</td>
<td>12417</td>
<td>12372</td>
<td>FOYERS TRAVAIL.</td>
</tr>
<tr>
<td>ETUD.CITE,FOY.</td>
<td>108860</td>
<td>108514</td>
<td>FOYERS ETUDIANT</td>
</tr>
<tr>
<td>MAISON RETRAITE</td>
<td>5180</td>
<td>5338</td>
<td>MAISON RETRAITE</td>
</tr>
<tr>
<td>ETAB.SOINS,CONV</td>
<td>34338</td>
<td>34422</td>
<td>HOSPI.LONG.DUR.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>COMMUN.RELIGIE.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CENT.HEBERGEMT.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AUTRES COLL.</td>
</tr>
<tr>
<td>CHANTIER TEMPO.</td>
<td>5072</td>
<td>4949</td>
<td>POP.HAB.MOBILES</td>
</tr>
<tr>
<td>DETENUS</td>
<td>4171</td>
<td>3895</td>
<td>POP.NON REINTEG</td>
</tr>
<tr>
<td>POP.RATTACHEE</td>
<td>598903</td>
<td>600926</td>
<td></td>
</tr>
</tbody>
</table>

**L’EQUILIBRE DU RAPPROCHEMENT: POPULATION DES 15 ANS OU PLUS MEMBRES ORDINAIRES OU RATTACHES DES LOGEMENTS 1/3 SOR. EE90**

<table>
<thead>
<tr>
<th>ENQUETE EMPLOI</th>
<th>JANVIER 1990</th>
<th>POPULATION</th>
<th>MARS 1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>40117 APPARIES &quot;SURS&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1773 APPARIES &quot;DOUTEUX&quot; ADULTES A EE ET RP</td>
<td>1773</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 APPARIES &quot;DOUTEUX&quot; ENFANTS AU RP</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36 APPARIES &quot;DOUTEUX&quot; ENFANTS A EE</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2396 ENQUETES NON APPARIES DS LOGEMENT APPARIE RECENSES NON APPARIES DS LOGEMENT ENQUETE</td>
<td>2274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1889 ENQUETES DANS LOGEMENTS NON APPARIES RECENSES DS RES. DE MENAGE NON REPONDANT</td>
<td>2410</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RECENSES DS RES. NON PRINCIPALES A EE</td>
<td>943</td>
<td></td>
</tr>
<tr>
<td>46202 TOTAL POPULATION DES 15 ANS OU PLUS</td>
<td>47553</td>
<td></td>
<td></td>
</tr>
<tr>
<td>46032 POPULATION DE L’EE CALEE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77 HORS CALAGE A L’ENQUETE APPARIES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93 HORS CALAGE A L’ENQUETE NON APPARIES</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

198 INSEE Méthodes n° 52 - 53
Tableau X

POP. DES 15 ANS OU PLUS DANS LES MENAGES PAR TYPE ACTIVITE
ESSAI DE BOUCLAGE ENTRE LE RECENSEMENT ET L'ENQUETE EMPLOI
RP MARS 1990 EXHAUSTIF, EE JANV.1990 (CHIFFRES EN MILLIERS)
MESURE EFFET PROTOCOLE : MAT. TACT3H<TACT3EE (40117 AD.SURS)

<table>
<thead>
<tr>
<th>OCCUPES</th>
<th>CHOMEURS</th>
<th>INACTIFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EE CALEÉE CHAMP TOTAL CON.BIT</td>
<td>22322</td>
<td>2254</td>
</tr>
<tr>
<td>CORRECTION DE CONCEPT Bit&gt;RP</td>
<td>39</td>
<td>182</td>
</tr>
<tr>
<td>EE CALEÉE CHAMP TOTAL CON.RP</td>
<td>22361</td>
<td>2436</td>
</tr>
<tr>
<td>CORRECTION DE CHAMP POP.MEN.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EE CON.RP, CH.MEN., ARROND.PERSO</td>
<td>22360</td>
<td>2440</td>
</tr>
<tr>
<td>CORRECTION DE DATE JANV.&gt;MARS</td>
<td>30</td>
<td>-40</td>
</tr>
<tr>
<td>EFFET PROTOCOLE SUR REPONDANTS</td>
<td>50</td>
<td>170</td>
</tr>
<tr>
<td>EE CON.RP, PRO.RP, CH.MEN, MARS</td>
<td>22440</td>
<td>2570</td>
</tr>
<tr>
<td>BIAIS DU AUX NON REPONSES EE</td>
<td>(?)0</td>
<td>30</td>
</tr>
<tr>
<td>ALEA DE SONDAGE EE</td>
<td>-100</td>
<td>70</td>
</tr>
<tr>
<td>RP POP. MENAGES EXHAUSTIF</td>
<td>22340</td>
<td>2670</td>
</tr>
</tbody>
</table>

La collecte
### Extrait du "Bulletin individuel de recensement"

**ANNEXE I**

**BULLETIN INDIVIDUEL DE RECENSEMENT**

<table>
<thead>
<tr>
<th>NOM, Prénoms</th>
<th>Écrivez le nom en capitales (exemple: ALLARD, épouse MAURIN, Françoise)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEXE</td>
<td>Masculin ☐</td>
</tr>
<tr>
<td>DATE ET LIEU DE NAISSANCE</td>
<td>Nom, prénom, année et numéro de civilité</td>
</tr>
<tr>
<td>ÉTES-VOUS?</td>
<td>Si vous êtes agent de l'État, d'une collectivité locale, d'un hôpital public, d'un service public (EDF, SNCF, etc.) ou militaire de carrière, précisez votre classification (corps, grade, etc.)</td>
</tr>
<tr>
<td>TRAVAILLEZ-VOUS?</td>
<td>À temps partiel ☐</td>
</tr>
<tr>
<td>AVEZ-VOUS DÉJÀ TRAVAILLÉ ?</td>
<td>OUI ☒</td>
</tr>
<tr>
<td>CHERCHEZ-VOUS UN EMPLOI ?</td>
<td>OUI ☒</td>
</tr>
<tr>
<td>PROFESSON EXERCE ACTUELLEMENT</td>
<td>Manœuvre ou ouvrier spécialisé (OS1, OS2, OS3, etc.) ☐</td>
</tr>
<tr>
<td>AIDEZ-VOUS UN MEMBRE DE VOTRE FAMILLE DANS SON TRAVAIL ?</td>
<td>OUI ☒</td>
</tr>
</tbody>
</table>

---

**La collecte**

201
Extrait du "Questionnaire de l'enquête emploi 1990"

ANNEXE II
ENQUÊTE SUR L'EMPLOI 1990

Mouvement de l'individu par rapport à l'enquête précédente (à remplir directement par l'enquêteur à partir du code MOB de la feuille de local)

Personne déjà enquêtée l'année précédente
Personne non enquêtée l'année précédente

Quelle est l'occupation actuelle de M... ?

Travaille
1. Exerce une profession, à son compte ou comme salarié, même à temps partiel.
   Aide un membre de sa famille dans son travail même sans être rémunéré.
   Est apprenti, stagiaire rémunéré, éleve-fonctionnaire, intérimaire, TUC, etc.
   • Y compris congés de maladie, de maternité, congés annuels, dispenses d'activité, congés de conversion, etc.
   • Exclure les personnes au service national, en disponibilité, en pré-retraite, en invalidité

Ne travaille pas (ou plus) actuellement
2. Chômeur (inscrit ou non à l'ANPE)
3. Étudiant, élève, en formation, en stage non rémunéré
4. Militaire du contingent
5. Retire des affaires (ancien agriculteur, ancien artisan, ancien commerçant)
6. Femme au foyer
7. Autre inactif (y compris les personnes ne touchant qu'une pension de reversion et les personnels invalides)

POUR CEUX QUI NE TRAVAILLENT PAS (OU PLUS) ACTUELLEMENT

La semaine dernière, du lundi au dimanche, M... a-t-il cependant travaillé ?
Oui
Non
• Ne fait ce qu'une heure, ne durée qu'un travail occasionnel ou exceptionnel.
• Exclure une activité non rémunérée en aidant un membre de sa famille dans son travail, par exemple.
• Les activités déshéritées ne sont pas prises en compte.

Depuis mars de l'année dernière, M... a-t-il exercé une activité professionnelle ?
Oui
Non
• Exclure ici les activités purement occasionnelles, par exemple les jobs d'étudiant pendant les vacances.

Si NON, M... a-t-il déjà exercé une activité professionnelle même s'il y a longtemps ?
Oui
Non
• Exclure ici les activités purement occasionnelles, par exemple les jobs d'étudiant pendant les vacances.

M... cherche-t-il un emploi, une situation ?
Oui
Non
• Si NON, M... souhaiterait-il cependant travailler ?
   Classez sans objet les personnes déjà pourvues d'un emploi.
   Oui
   Non
   Sans objet

INSEE Méthodes n° 52 - 53
**Extrait du "Questionnaire de l'enquête emploi 1990"**

*Suite et fin*

**8.** Depuis combien de temps M... cherche-t-il un emploi ou un autre emploi ?

<table>
<thead>
<tr>
<th>Nombre de mois</th>
<th>Indiquer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mois à moins de 3 mois</td>
<td>1</td>
</tr>
<tr>
<td>2 mois à moins de 6 mois</td>
<td>2</td>
</tr>
<tr>
<td>3 mois à moins de 1 an</td>
<td>3</td>
</tr>
<tr>
<td>4 mois à moins de 2 ans</td>
<td>4</td>
</tr>
<tr>
<td>5 mois à moins de 3 ans</td>
<td>5</td>
</tr>
<tr>
<td>6 mois à moins de 1 an et demi</td>
<td>6</td>
</tr>
<tr>
<td>7 mois à moins de 2 ans</td>
<td>7</td>
</tr>
<tr>
<td>8 mois à moins de 3 ans</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Durée</th>
<th>Indiquer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 an à moins de 1 an et demi</td>
<td>9</td>
</tr>
<tr>
<td>1 an et demi à moins de 2 ans</td>
<td>10</td>
</tr>
<tr>
<td>2 ans à moins de 3 ans</td>
<td>11</td>
</tr>
<tr>
<td>3 ans et plus</td>
<td>12</td>
</tr>
</tbody>
</table>

**9.** Nature de l'emploi recherché.

- M... cherche-t-il :
  1. Un emploi salarié
  2. Une situation à son compte
  3. Indifféremment l'un ou l'autre

- M... cherche-t-il un emploi :
  1. A temps complet (à l'exclusion du temps partiel)
  2. A temps complet, mais à défaut accepterait à temps partiel
  3. A temps partiel, mais à défaut accepterait à temps complet
  4. A temps partiel (à l'exclusion du temps complet)

- M... cherche-t-il :
  1. Un emploi permanent (à l'exclusion d'un emploi à durée limitée)
  2. Un emploi permanent, mais à défaut accepterait un emploi à durée limitée
  3. Un emploi pour une durée limitée (à l'exclusion d'un emploi permanent)
  4. Un emploi pour une durée limitée, mais à défaut accepterait un emploi permanent

- Profession (et qualification) cherchée :
  1. Cherche une profession précise : indiquer laquelle
  2. Cherche plusieurs professions : indiquer la principale
  3. Cherche dans une gamme large d'emplois : indiquer des précisions éventuelles
  4. Cherche n'importe quoi : indiquer des précisions éventuelles
Partie 3

L'EXPLOITATION
3.1 - LE DÉNOMBREMENT DU RP90
(Bilan chiffré)

Chantal Madinier

Les conclusions tirées après l'étude de comparaison des dénombrements issus au RP82 de l'établissement de la population légale (population sans doubles comptes) d'une part et de l'exploitation exhaustive de l'autre faite par Geneviève LEBOUTEUX ont conduit à utiliser très largement la saisie exhaustive pour vérifier la population légale. Cette vérification est d'ailleurs réciproque car aucune des deux sources n'est a priori plus juste que l'autre. Les résultats que l'on peut tirer en 1990 de la comparaison population légale (population sans doubles comptes) - exhaustif montrent que le bilan est très positif. 23 330 communes ont la même population au dénombrement et à l'exhaustif contre 20 070 en 1982. De plus, la somme des valeurs absolues des écarts communaux entre les deux chiffres est passée de 112 672 en 1982 à 47 059 en 1990.

1 – La procédure utilisée en 1990

Chaque Direction Régionale a constitué des unités de traitement (UT) : chacune de ces UT contenait entre 100 000 et 250 000 bulletins individuels (BI). Avant l'expédition chez les façonniers, une saisie des chiffres des districts était effectuée à partir des bordereaux n° 6 (M6). On avait donc pour chaque district sa population et son nombre de feuilles de logements (FL).

Après saisie de l’UT, la bande saisie par le façonnier permettait de calculer la population et le nombre de logements. Le contrôle d’exhaustivité était le suivant :

pour chaque district, la différence entre les deux sources (exhaustif et population légale) devait respecter les inégalités suivantes :

\[ |\text{nb BI (M6)} - \text{nb BI (exh)}| < 1 + \sqrt{\frac{\text{nb BI (M6)}}{10}} \]
Pour chaque commune, on effectuait les mêmes comparaisons avec une différence autorisée un peu plus grande :

\[
\left| \text{nb FL (M6)} - \text{nb FL (exh)} \right| < 1 + \sqrt{\frac{\text{nb FL (M6)}}{10}}
\]

\[
\left| \text{nb BI (M6)} - \text{nb BI (exh)} \right| < 1 + \sqrt{\frac{\text{nb BI (M6)}}{5}}
\]

\[
\left| \text{nb FL (M6)} - \text{nb FL (exh)} \right| < 1 + \sqrt{\frac{\text{nb FL (M6)}}{5}}
\]

Les identifiants des districts et des communes ne respectant pas la fourchette d'écart autorisée étaient édités et les Directions régionales (DR) cherchaient l'origine de l'écart entre les deux dénombrements pour soit corriger la valeur calculée en population légale, soit signaler au façonnier une erreur qu'il aurait commise (documents oubliés ou comptés plusieurs fois par exemple).

Cette procédure a pu être utilisée pour toutes les UT traitées à l'exploitation exhaustive suffisamment tôt pour entrer dans le cadre des délais réglementaires d'établissement de la population légale (c'est-à-dire avant le 1er janvier de l'année qui suit le recensement).

Cette comparaison des deux sources a été un très bon outil pour la qualité du RP90.

2 — Les résultats

Les résultats présentés dans cet article ne concernent que la France métropolitaine sans la Corse. En effet, les chiffres pour 1982 ne sont disponibles que sur ce champ car il n'y a pas eu d'exploitation exhaustive pour la Corse au RP 82.

2.1. Résultats nationaux et départementaux

La différence entre la population obtenue lors de l'exploitation exhaustive (appelée par la suite population exhaustive) et celle du dénombrement (population sans doubles
comptes) est de 10 513 habitants soit 0,02 %. Par rapport à 1982, cette différence a légèrement augmenté puisqu’elle n’était que de 6 742 (0,01 %).

A l’échelle départementale, cette différence peut être un peu plus importante en valeur relative. Elle varie en valeur absolue de − 710 habitants pour la Loire-Atlantique à + 734 pour le Nord en passant par 0 dans le Loiret. En valeur relative, elle est maximum pour l’Ariège (+ 0,1 %). Viennent ensuite le Var et le Lot (+ 0,08 %). Dans 18 départements seulement, la population exhaustive est inférieure à la population sans doubles comptes (maximum − 0,07 % pour le Maine-et-Loire et la Loire-Atlantique).

Par rapport à 1982, l’amélioration est certaine puisque cet écart (en valeur absolue) dépassait 0,1 % pour 17 départements. La somme des valeurs absolues des écarts départementaux est passée de 46 384 en 1982 à 16 137 en 1990.

Si on considère la somme des valeurs absolues des écarts communaux, l’amélioration est nette : cet indicateur est passé de 112 672 en 1982 à 47 059 en 1990.

2.2. Résultats communaux

Sur les 36 210 communes du champ étudié, 23 330 (soit 64 %) présentent le même chiffre pour la population exhaustive et la population sans doubles comptes. L’écart (en valeur absolue) n’excède pas deux habitants pour 8 770 communes (24 %).

La différence en valeur absolue dépasse 10 pour 595 communes. Elle excède plus ou moins 2 % pour 330 d’entre elles (Tableaux 1 et 2).

En 1982, la différence était nulle pour 20 070 communes soit 55 %. Elle dépassait (en valeur absolue) 10 habitants pour 1 315 communes, et 2 % en plus ou en moins pour 616 d’entre elles (Tableaux 3 et 4).

329 communes présentent des écarts importants entre la population exhaustive et la population sans doubles comptes. Ces communes à écarts importants étaient au nombre de 775 en 1982. Ces écarts importants sont définis ainsi : la valeur absolue de l’écart est supérieure ou égale à 5 et en même temps représente au moins 1 % de la population sans doubles comptes.

Ces communes à gros écarts sont des petites communes : 199 (soit 60 %) ont moins de 500 habitants et 93 (28 %) ont entre 500 et 999 habitants. Ces petites communes sont surreprésentées par rapport à leur poids au niveau national. Aucune commune de plus de 50 000 habitants ne présente un gros écart. La somme des valeurs absolues des écarts pour ces 329 communes est de 5 541 (Tableau 5).
En 1982, la situation était très différente : la moitié des communes à gros écarts avait plus de 500 habitants. A partir de 5 000 habitants plus les communes étaient grandes, plus des divergences étaient fréquentes. La somme des valeurs absolues des écarts était de 31 365 (Tableau 6).

On constate donc une très nette amélioration de la qualité en 1990. Les écarts entre ces deux populations sont surtout le fait des petites communes et le nombre d’habitants concerné est très faible.

Référence des notes :


877/RP90 du 13 juillet 1989 : Contrôles informatiques au cours de la vérification de la population légale.
Tableau 1
Répartition des communes en 1990 selon l’écart entre la population sans doubles comptes et la population exhaustive

<table>
<thead>
<tr>
<th>Taille (habitants)</th>
<th>-11 ou moins</th>
<th>-10 à -6</th>
<th>-5 à -3</th>
<th>-2 à -1</th>
<th>0</th>
<th>1 à 2</th>
<th>3 à 10</th>
<th>11 et plus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>5</td>
<td>43</td>
<td>266</td>
<td>2 163</td>
<td>17 251</td>
<td>1 392</td>
<td>264</td>
<td>14</td>
<td>21 418</td>
</tr>
<tr>
<td>500 à 999</td>
<td>10</td>
<td>74</td>
<td>355</td>
<td>1 337</td>
<td>3 717</td>
<td>849</td>
<td>198</td>
<td>27</td>
<td>6 567</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>28</td>
<td>126</td>
<td>373</td>
<td>967</td>
<td>1 535</td>
<td>615</td>
<td>248</td>
<td>29</td>
<td>3 921</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>74</td>
<td>209</td>
<td>404</td>
<td>611</td>
<td>634</td>
<td>397</td>
<td>239</td>
<td>25</td>
<td>2 593</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>73</td>
<td>102</td>
<td>153</td>
<td>153</td>
<td>134</td>
<td>126</td>
<td>105</td>
<td>33</td>
<td>879</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>68</td>
<td>73</td>
<td>62</td>
<td>54</td>
<td>41</td>
<td>42</td>
<td>67</td>
<td>20</td>
<td>427</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>83</td>
<td>42</td>
<td>24</td>
<td>31</td>
<td>18</td>
<td>26</td>
<td>42</td>
<td>31</td>
<td>297</td>
</tr>
<tr>
<td>50 000 à 99 999</td>
<td>25</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>17</td>
<td>64</td>
</tr>
<tr>
<td>100 000 et plus</td>
<td>16</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>15</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>384</td>
<td>674</td>
<td>1 667</td>
<td>5 319</td>
<td>23 330</td>
<td>3 451</td>
<td>1 174</td>
<td>211</td>
<td>36 210</td>
</tr>
</tbody>
</table>

(Champ : France métropolitaine hors Corse)

Tableau 2
Répartition des communes en 1990 selon l’écart entre la population sans doubles comptes et la population exhaustive

<table>
<thead>
<tr>
<th>Taille (habitants)</th>
<th>-2 et moins</th>
<th>-1,99 à -1</th>
<th>-0,99 à -0,5</th>
<th>-0,49 à -0,2</th>
<th>-0,19 à -0,001</th>
<th>0</th>
<th>0,001 à 0,199</th>
<th>0,2 à 0,99</th>
<th>1 à 1,99</th>
<th>2 et plus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>133</td>
<td>330</td>
<td>784</td>
<td>1 250</td>
<td>0</td>
<td>17 251</td>
<td>0</td>
<td>1240</td>
<td>270</td>
<td>160</td>
<td>21 418</td>
</tr>
<tr>
<td>500 à 999</td>
<td>6</td>
<td>37</td>
<td>205</td>
<td>541</td>
<td>987</td>
<td>3 717</td>
<td>613</td>
<td>410</td>
<td>31</td>
<td>20</td>
<td>6 567</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>3</td>
<td>6</td>
<td>79</td>
<td>378</td>
<td>1 028</td>
<td>1 535</td>
<td>659</td>
<td>217</td>
<td>10</td>
<td>6</td>
<td>3 921</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>0</td>
<td>4</td>
<td>27</td>
<td>198</td>
<td>1 069</td>
<td>634</td>
<td>576</td>
<td>80</td>
<td>4</td>
<td>1</td>
<td>2 593</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>431</td>
<td>134</td>
<td>245</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>879</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>243</td>
<td>41</td>
<td>123</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>427</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>174</td>
<td>18</td>
<td>96</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>297</td>
</tr>
<tr>
<td>50 000 à 99 999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>36</td>
<td>0</td>
<td>25</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>100 000 et plus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>377</td>
<td>1 105</td>
<td>2 428</td>
<td>3 992</td>
<td>23 330</td>
<td>2 356</td>
<td>1 975</td>
<td>317</td>
<td>188</td>
<td>36 210</td>
</tr>
</tbody>
</table>

(Champ : France métropolitaine hors Corse)

L’exploitation 211
Tableau 3
Répartition des communes en 1982 selon l’écart entre la population sans doubles comptes et la population exhaustive

<table>
<thead>
<tr>
<th>Taille (habitants)</th>
<th>-11 ou moins</th>
<th>-10 à -6</th>
<th>-5 à -3</th>
<th>-2 à -1</th>
<th>0</th>
<th>1 à 2</th>
<th>3 à 10</th>
<th>11 et plus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>51</td>
<td>107</td>
<td>479</td>
<td>3 163</td>
<td>15 909</td>
<td>2 018</td>
<td>353</td>
<td>36</td>
<td>22 116</td>
</tr>
<tr>
<td>500 à 999</td>
<td>62</td>
<td>136</td>
<td>468</td>
<td>1 541</td>
<td>2 762</td>
<td>1 095</td>
<td>333</td>
<td>40</td>
<td>6 437</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>79</td>
<td>184</td>
<td>439</td>
<td>970</td>
<td>964</td>
<td>721</td>
<td>311</td>
<td>51</td>
<td>3 719</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>168</td>
<td>217</td>
<td>339</td>
<td>463</td>
<td>353</td>
<td>387</td>
<td>338</td>
<td>86</td>
<td>2 351</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>142</td>
<td>100</td>
<td>101</td>
<td>88</td>
<td>61</td>
<td>84</td>
<td>137</td>
<td>84</td>
<td>797</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>118</td>
<td>51</td>
<td>44</td>
<td>23</td>
<td>17</td>
<td>15</td>
<td>48</td>
<td>76</td>
<td>392</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>146</td>
<td>14</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>9</td>
<td>21</td>
<td>75</td>
<td>286</td>
</tr>
<tr>
<td>50 000 à 99 999</td>
<td>35</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>21</td>
<td>67</td>
</tr>
<tr>
<td>100 000 et plus</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>824</td>
<td>811</td>
<td>1 883</td>
<td>6 254</td>
<td>20 070</td>
<td>4 331</td>
<td>1 546</td>
<td>491</td>
<td>36 210</td>
</tr>
</tbody>
</table>

(Champ : France métropolitaine hors Corse)

Tableau 4
Répartition des communes en 1982 selon l’écart entre la population sans doubles comptes et la population exhaustive

<table>
<thead>
<tr>
<th>Taille (habitants)</th>
<th>-2 et moins</th>
<th>-1,99 à -1</th>
<th>-0,99 à -0,5</th>
<th>-0,49 à -0,2</th>
<th>-0,19 à 0</th>
<th>0</th>
<th>0,001 à 0,199</th>
<th>0,2 à 0,99</th>
<th>1 à 1,99</th>
<th>2 et plus</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>293</td>
<td>579</td>
<td>1 222</td>
<td>1 706</td>
<td>0</td>
<td>15 909</td>
<td>0</td>
<td>1 857</td>
<td>353</td>
<td>197</td>
<td>22 116</td>
</tr>
<tr>
<td>500 à 999</td>
<td>41</td>
<td>99</td>
<td>317</td>
<td>706</td>
<td>1 044</td>
<td>2 762</td>
<td>790</td>
<td>596</td>
<td>55</td>
<td>27</td>
<td>6 437</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>11</td>
<td>28</td>
<td>139</td>
<td>439</td>
<td>1 055</td>
<td>964</td>
<td>757</td>
<td>292</td>
<td>18</td>
<td>16</td>
<td>3 719</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>10</td>
<td>12</td>
<td>64</td>
<td>257</td>
<td>844</td>
<td>353</td>
<td>650</td>
<td>135</td>
<td>17</td>
<td>9</td>
<td>2 351</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>81</td>
<td>331</td>
<td>61</td>
<td>235</td>
<td>59</td>
<td>7</td>
<td>4</td>
<td>797</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>33</td>
<td>192</td>
<td>17</td>
<td>104</td>
<td>27</td>
<td>6</td>
<td>2</td>
<td>392</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>34</td>
<td>136</td>
<td>3</td>
<td>75</td>
<td>23</td>
<td>4</td>
<td>3</td>
<td>286</td>
</tr>
<tr>
<td>50 000 à 99 999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>100 000 et plus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>Total</td>
<td>357</td>
<td>721</td>
<td>1 776</td>
<td>3 266</td>
<td>3 652</td>
<td>20 070</td>
<td>2 641</td>
<td>3 007</td>
<td>461</td>
<td>259</td>
<td>36 210</td>
</tr>
</tbody>
</table>

(Champ : France métropolitaine hors Corse)
### Tableau 5

**Communes à écart important en 1990 entre la population exhaustive et la population sans doubles comptes**

<table>
<thead>
<tr>
<th>Taille</th>
<th>Nombre de communes à écart important</th>
<th>Nombre de communes du champ</th>
<th>communes à écart</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>199 (60,5)</td>
<td>21 418 (59,2)</td>
<td>0,9</td>
</tr>
<tr>
<td>500 à 999</td>
<td>93 (28,3)</td>
<td>6 567 (18,1)</td>
<td>1,4</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>25 (7,6)</td>
<td>3 921 (10,8)</td>
<td>0,6</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>9 (2,7)</td>
<td>2 593 (7,2)</td>
<td>0,3</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>1 (0,3)</td>
<td>879 (2,4)</td>
<td>0,1</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>1 (0,3)</td>
<td>427 (1,2)</td>
<td>0,2</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>1 (0,3)</td>
<td>297 (0,8)</td>
<td>0,3</td>
</tr>
<tr>
<td>plus de 50 000 hab.</td>
<td>0 (0,0)</td>
<td>110 (0,3)</td>
<td>0,0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>329 (100,0)</td>
<td>36 210 (100,0)</td>
<td>0,9</td>
</tr>
</tbody>
</table>

*(Champ : France métropolitaine hors Corse)*

### Tableau 6

**Communes à écart important en 1982 entre la population exhaustive et la population sans doubles comptes**

<table>
<thead>
<tr>
<th>Taille</th>
<th>Nombre de communes à écart important</th>
<th>Nombre de communes du champ</th>
<th>communes à écart</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 499</td>
<td>399 (51,5)</td>
<td>22 116 (61,0)</td>
<td>1,8</td>
</tr>
<tr>
<td>500 à 999</td>
<td>222 (28,6)</td>
<td>6 437 (17,8)</td>
<td>3,4</td>
</tr>
<tr>
<td>1 000 à 1 999</td>
<td>73 (9,4)</td>
<td>3 719 (10,3)</td>
<td>2,0</td>
</tr>
<tr>
<td>2 000 à 4 999</td>
<td>48 (6,2)</td>
<td>2 351 (6,5)</td>
<td>2,0</td>
</tr>
<tr>
<td>5 000 à 9 999</td>
<td>13 (1,7)</td>
<td>797 (2,2)</td>
<td>1,6</td>
</tr>
<tr>
<td>10 000 à 19 999</td>
<td>9 (1,2)</td>
<td>392 (1,1)</td>
<td>2,3</td>
</tr>
<tr>
<td>20 000 à 49 999</td>
<td>8 (1,2)</td>
<td>392 (1,1)</td>
<td>2,3</td>
</tr>
<tr>
<td>50 000 à 99 999</td>
<td>2 (0,3)</td>
<td>67 (0,2)</td>
<td>3,0</td>
</tr>
<tr>
<td>100 000 et plus</td>
<td>1 (0,1)</td>
<td>45 (0,1)</td>
<td>2,2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>775 (100,0)</td>
<td>36 210 (100,0)</td>
<td>2,1</td>
</tr>
</tbody>
</table>

*(Champ : France métropolitaine hors Corse)*

L’exploitation 213
3.2 - ÉCHANTILLONNAGES POUR LE CONTRÔLE DE QUALITÉ DU RECENSEMENT DE 1990

Jean-Claude Deville

Introduction digressive

Bien que cela puisse sembler paradoxal, le Recensement de la Population (RP) engendre un nombre considérable d’opérations de sondage au cours de sa réalisation.


L’exploitation du recensement fait elle-même appel à la technique des sondages. En effet, pour réduire les coûts et les délais de la publication, un quart seulement des bulletins sont codifiés entièrement, notamment en ce qui concerne la détermination de la catégorie socioprofessionnelle (CSP), de l’activité économique et des liens entre personnes d’un même ménage. L’échantillon codifié résulte d’une stratification par commune et par taille de ménage. Il est utilisé pour produire des résultats détaillés au niveau d’unités comptant au moins 5 000 habitants.

Cet échantillon au quart est précédé par la publication des résultats sur échantillon au 1/20 de la population.

L’utilité de ce dernier sondage est d’obtenir rapidement des résultats valides au niveau de la France entière et éventuellement au niveau des régions. Il est réalisé en tirant avec probabilités égales un district sur cinq, dont on codifie le quart des bulletins. Ainsi l’échantillon au 1/20 apparaît-il comme le premier cinquième de l’échantillon au quart.

Rappelons qu’un district est une unité de collecte du RP correspondant, en ville, à un pâté de maisons et, à la campagne, à un village ou une réunion de hameaux. Il peut comporter de quelques logements à plus de mille, et compter une population pouvant...
varier de 0 à environ 2 000 habitants. La moyenne est de 150 logements pour 350 habitants environ.

En raison de cette dispersion de taille, le tirage d’un cinquième des districts pouvait produire un sondage très imprécis. On a pu établir (Chartier, 1979) qu’un sondage aléatoire simple de districts multipliait par un facteur variant de deux à cinq environ la variance d’un sondage aléatoire simple de logements. Une technique d’échantillonnage équilibré (Deville, et Alii, 1988) a permis d’atténuer très largement cet effet.

**Contrôle de qualité**

La technique des sondages a aussi été très utile dans le processus de contrôle de fabrication du RP.

Résumons le processus qui a été suivi. À mesure de l’achèvement de la collecte et de sa vérification, les différents bulletins du recensement, notamment les bulletins individuels (BI) et les feuilles de logement (FL), sont soigneusement comptés pour chaque district. Les données récapitulatives des districts sont saisies sur support informatique, tandis que les bulletins groupés dans une chemise de district, partent pour la saisie.

Celle-ci se réalise en deux étapes :

- dans la première étape, dite de l’”exhaustif”, des ensembles de districts groupant environ 100 000 logements sont constitués. Ce sont les unités de traitement (UT). Chaque UT est saisie par une entreprise à façon pour le compte de l’Insee. L’Insee, le "client" en termes de théorie du contrôle, vérifie la qualité du travail de chaque façonnier en contrôlant par sondage un certain nombre de bulletins dans chaque UT. Les principes permettant d’optimiser le coût de ce contrôle sont exposés dans cet article ;

- la seconde étape d’élaboration des données est dite opération COLIBRI (pour Codification en Ligne des Bulletins du Recensement des Individus). Recevant des bulletins toujours groupés en districts, les opérateurs et opératrices des Directions Régionales (DR) de l’Insee, procédaient à leur codification pour constituer le sondage au quart.

Physiquement, chaque opérateur(trice) travaille devant un écran qui lui indique l’identifiant du prochain logement à inclure dans l’échantillon au quart dont il doit coder tous les BI.

Le contrôle de la qualité de la codification est également réalisé par sondage. L’unité de contrôle est l’ensemble du travail réalisé en une semaine dans une Direction Régionale. L’opération dure un peu plus d’un an dans les 22 Directions Régionales soit plus de mille sondages. L’unité à contrôler est le ménage (c’est-à-dire l’ensemble
des BI d’un ménage tiré pour figurer dans l’échantillon de contrôle). L’objectif est d’estimer la proportion de bulletins comportant une erreur. Pour cela, on détecte automatiquement ceux pour lesquels apparaît une divergence entre les deux codifications. Une opération de réconciliation permet de chiffrer le nombre d’erreurs. La modalité pratique et les enseignements tirés de ces contrôles sont détaillés dans G. Badeyan (1992).

Données générales sur les contrôles exhaustif et Colibri : la fonction de coût

Il se trouve que dans les deux problèmes la fonction de coût est la même, exprimée en temps de travail. Le contrôle consiste à aller chercher un district là où il est rangé, à l’amener au poste de travail de l’opérateur (trice) chargé du contrôle. Celui-ci (ou celle-ci) recherche les bulletins à contrôler dont il (elle) dispose des identifiants, en saisit leur codification. Les bulletins sont remis à leur place dans la chemise de district. Celle-ci est ramenée dans l’étagère où elle est habituellement stockée.

Globalement, donc, le temps passé au contrôle d’un district se décompose en un temps de manipulation de la chemise de district, qui ne dépend pas du district (taille ou localisation), et un temps de traitement de chaque bulletin à contrôler dans celui-ci. Ce temps de traitement ne dépend que du type de bulletin (BI ou FL), mais est considéré, à type de bulletin donné, comme indépendant du bulletin.

La fonction de coût du contrôle peut donc être approximée par une formule simple. Supposons que le contrôle porte sur \( m \) districts et regroupe \( n_g \) bulletins du type \( g \) au total. Si \( C_o \) est le coût en temps associé à la manipulation d’une chemise de district et \( C_g \) le coût en temps associé au contrôle d’un bulletin de type \( g \), le coût du contrôle vaudra :

\[
C_t = m \cdot C_o + \sum_{g} n_g \cdot C_g
\]

Cette expression est une généralisation naturelle d’une fonction de coût couramment utilisée décomposant les coûts d’une collecte en un coût d’approche \( C_o \) et un coût unitaire \( C_f \) (voir par exemple Desabie (1965) ou Cochran (1977)).
Le contrôle de l’exhaustif

**Première forme : on ne contrôle que les bulletins individuels**

Les districts \( k \) d’une UT (notée \( U \)) comportent chacun un nombre connu \( N_k \) de BI. Parmi ceux-ci \( D_k \) comportent une "erreur" (assimilée à une différence entre ce qui est codé à l’Insee et ce qui a été codé chez le façonnier). Le but est donc d’estimer :

\[
P = \frac{\sum U D_k}{\sum U N_k}
\]

Le sondage consistera à tirer un échantillon \( s \) de districts avec des probabilités d’inclusion \( \pi_k \) au premier ordre et \( \pi_{kl} \) au second ordre à déterminer. Ensuite, si le district \( k \) est tiré dans \( s \) on vérifiera \( n_k \) BI tirés par sondage aléatoire simple sans remise (SASSR). Soit \( d_k \) le nombre de BI erronés qu’on relèvera.

L’estimateur \( \hat{P}_k \) de \( P_k = D_k/N_k \) sera \( \hat{P}_k = d_k/n_k \) et \( \hat{D}_k = N_k \hat{P}_k \) estimera \( D_k \) sans biais. L’estimateur de \( P \) sera :

\[
\hat{P} = \frac{\sum_k \hat{D}_k}{\pi_k} = \frac{\sum_k \hat{D}_k}{\sum_k N_k/\pi_k}
\]

(4-1)

C’est le ratio des estimateurs sans biais de \( D \) et de \( N \), le nombre total de bulletins. Bien que ce nombre soit connu, il est bien évident que l’estimateur (4-1) est plus précis que \( \frac{1}{N} \sum_s \hat{P}_k / \pi_k \).

On a :

\[
\text{VAR}(\hat{P}) = E \text{Var}(\hat{P}/s) + \text{Var}(E \hat{P}/s)
\]

(4-2)

Or :

\[
\text{Var}(\hat{P}/s) = \hat{N}^{-2} \sum_s \frac{N_k^2}{\pi_k^2} \frac{P_k (1 - P_k) N_k (1 - \frac{1}{N_k})}{N_k - 1}
\]

avec \( \hat{N} = \sum_s N_k / \pi_k \)

D’où :

\[
E \text{Var}(\hat{P}/s) = \hat{N}^{-2} \sum_U \frac{N_k^2}{\pi_k} \frac{P_k (1 - P_k) N_k [1 - \frac{1}{N_k}]}{N_k - 1} \frac{\sum_k D_k / \pi_k}{\sum_s N_k / \pi_k}
\]

(4-3)

Par ailleurs :

\[
E(\hat{P}/s) = \frac{\sum_s N_k / \pi_k}{\sum_s N_k / \pi_k}
\]

218 INSEE Méthodes n°52-53
La variance de cette quantité s'obtient par linéarisation en introduisant la variable
\[ Z_k = D_k - PN_k = N_k (P_k - P). \]

On obtient :
\[ \text{Var} E (\hat{P} / s) = N^{-2} \text{Var} \left( \sum_s \frac{Z_k}{\pi_k} \right) \]

Soit, compte tenu de ce que \( \sum U Z_k = 0 \) :
\[
\text{Var} E (\hat{P} / s) = N^{-2} \left( \sum_k \frac{Z_k^2}{\pi_k} + \sum_{k \neq 1} \sum \frac{z_k z_{1,k}}{\pi_k \pi_{1,k}} \right) \quad (4 - 4)
\]

La somme des quantités (4-3) et (4-4) nous donne la variance de l'estimateur (4-1).

**Introduction d'un modèle**

La variance de \( \hat{P} \) est difficile à manipuler et, de plus, contient des paramètres inconnus. On se tire de la difficulté en faisant de nécessaires hypothèses qui se traduisent par un modèle de superpopulation. On verra plus loin que les paramètres de ce modèle sont susceptibles d'être estimés à partir d'un essai préliminaire de recensement portant sur une toute petite partie du territoire. On note \( E \xi \) l'espérance sous le modèle (resp \( \text{Var} \xi \) pour la variance) dont tous les aléas sont supposés indépendants du processus d'échantillonnage.

Le modèle suit les spécifications suivantes :

a) \( D_k \) suit une loi binomiale \( (N_k, p_k) \). \( P_k \) est donc, sous le modèle, un estimateur de \( p_k \).

b) \( p_k \) est lui-même aléatoire. On suppose les \( p_k \) indépendantes et de même loi avec :
\[ E \xi p_k = P \]
\[ \text{Var} \xi p_k = \sigma^2 \]

pour tout \( k \), quelle que soit, en particulier, la valeur de \( N_k \).

En conditionnant, dans le modèle, par les \( p_k \) on a évidemment :
\[ E \xi (D_k \mid p_k) = N_k p_k \]
\[ \text{Var} \xi D_k = N_k p_k (1 - p_k) \]

La variance anticipée de \( \hat{P} \) est la quantité \( E \xi \text{Var} \hat{P} \). C'est à elle que nous allons nous intéresser désormais. Pour l'évaluer on remarque que :
\[
\alpha - E \xi (P_k - P)^2 = E \xi (E \xi p_k + p_k - p_k - P)^2 \mid p_k) = \frac{P (1 - P) - \sigma^2}{N_k} + \sigma^2
\]

L'exploitation
\[
\begin{align*}
  b - E \xi_p k (1 - P_k) &= E \xi_p (E \xi_\xi ((p_k - P_k)^2 \mid p_k)) = E \xi_p p_k (1 - p_k) \frac{N_k - 1}{N_k} \\
  &= (P (1 - P) - \sigma^2) \frac{N_k - 1}{N_k}
\end{align*}
\]

\( c - E \xi Z_k Z_l = 0 \) à cause de l'indépendance des \( Z_k \) et des \( Z_l \), ce qui nous débarrasse d'un terme bien encombrant en même temps que des \( \pi_k \).

En recollant tous les morceaux de (4-3) et (4-4) un petit miracle algébrique se produit et nous avons l'expression :

\[
E \xi \text{ Var } \hat{P} = N^{-2} \sum_U \frac{N_k^2}{\pi_k} \left( \sigma^2 + \frac{\tau^2}{n_k} \right)
\]

avec \( \tau^2 = P (1 - P) - \sigma^2 \geq 0 \)

**Remarque :**

Le miracle algébrique s'explique bien si on ne cherche pas à obtenir la variance sous le plan de sondage uniquement. Elle est d'ailleurs la conséquence d'un modèle un peu plus général que celui que nous avons posé.

Supposons que nous voulions estimer le total \( N \bar{Y} = \sum_U Y_i \) d'une variable \( Y \) et que pour cela nous réalisions un tirage à deux degrés : un premier degré où des unités primaires (UP) \( k \) sont tirées avec des probabilités \( \pi_k \), un second où \( n_k \) unités finales sont tirées par sondage aléatoire simple.

Nous posons un modèle où :

\[
Y_i = \bar{Y} + \alpha_k + \varepsilon_i
\]

avec \( \alpha_k \) variable liée à l'UP d'indice \( k \). Les \( \alpha_k \) sont indépendantes de même loi d'espérance nulle de variance \( \sigma^2 \).

Les \( \varepsilon_i \) sont également indépendantes centrées de variance égale à \( \tau^2 \). Avec \( \pi^*_i = n_k / N_k \) (\( N_k \) taille de l'UP numéro \( k \)), l'estimateur de Horvitz-Thompson du total vaut \( \hat{Y} = \sum \hat{Y}_i / \pi^*_i \) la somme étant étendue à l'échantillon.
Sous le modèle, et conditionnellement à l'échantillon on a :

\[
\text{Var}_k \hat{Y} = \sum_s \frac{N_k^2}{\pi_k^2} \left( \sigma^2 + \frac{\tau^2}{n_k} \right)
\]

L'espérance sous le plan de cette expression redonne la formule

(5-1).

**Optimisation du sondage**

La variance maximum de \( \hat{P} \) est fixée par les critères retenus pour le contrôle de qualité. Le sondage étant répété pour chacune des unités de traitement il est tout à fait naturel de chercher à minimiser l'espérance du coût du sondage donné en (3-1) soit :

\[
E \sum_s (C_o + n_k \ C_1) = \sum_U \pi_k (C_o + n_k \ C_1)
\]

avec \( C_o \) temps de manipulation des chemises de district et \( C_f \) temps requis pour la codification d'un BI.

Le problème d'optimisation s'écrit donc :

Minimiser \( \sum_U \pi_k (C_o + n_k \ C_1) \)

sous les contraintes :

\[
N^{-2} \sum_U \frac{N_k^2}{\pi_k} \left( \sigma^2 + \frac{\tau^2}{n_k} \right) \leq V_o
\]

et pour tout \( k, n_k \leq N_k \).

Associons un multiplicateur de Lagrange \( \lambda \) à la première contrainte - qui sera évidemment saturée - et des multiplicateurs \( \mu_k \) aux autres. On obtient les solutions :

\[
C_o + n_k \ C_1 = \lambda \frac{N_k^2}{\pi_k} \left( \sigma^2 + \frac{\tau^2}{n_k} \right)
\]

et, pour tout \( k \) :

\[
\pi_k = \frac{N_k^2}{\pi_k} \left( \frac{\tau^2}{n_k^2} + \mu_k \right)
\]

avec \( \mu_k = 0 \) si \( n_k < N_k \) et \( \mu_k \geq 0 \) - roman size 10 si~ n sub k~−~N sub k~

Pour tous les districts où \( \mu_k = 0 \) (les plus gros) on obtient :

\[
n_k = \frac{\tau}{\sigma} \left( \frac{C_o}{C_1} \right)^\frac{1}{2} = n^*
\]

Chaque district reçoit donc la même allocation, ce qui correspond à l'idée qu'on a besoin de la même précision dans chacun d'eux. Retournons à l'équation (6-3). On constate alors que, toujours pour ces districts, les probabilités d'inclusion \( \pi_k \) doivent être proportionnelles aux tailles \( N_k \) soit :

---

L'exploitation
\[
\pi_k = \lambda^{\frac{1}{2}} C^1 \left( \frac{\tau}{n^*} \right) N_k
\]  
(6 - 5)

C’est la justification habituelle d’un sondage autopondéré avec un premier degré tiré avec des probabilités proportionnelles à une mesure de taille.

Comme \( n_k \) ne dépend pas de \( N_k \) on ne pourra avoir \( n_k = N_k \) et \( \mu_k > 0 \) que si \( N_k \leq n^* \). L’équation (6-1) nous permet alors d’obtenir les probabilités d’inclusion à un facteur près :

\[
n_k = \lambda^{\frac{1}{2}} N_k \left( \frac{\sigma^2 + \tau^2/N_k}{C_o + C_1 N_k} \right)^{\frac{1}{2}} = \lambda^{\frac{1}{2}} N_k^{\frac{1}{2}} \left( \frac{N_k \sigma^2 + \tau^2}{N_k C_1 + C_o} \right)^{\frac{1}{2}}
\]  
(6 - 6)

Les relations (6-5), valide si \( N_k \geq n^* \) et (6-6) valide si \( N_k \leq n^* \), établissent que \( \pi_k \) est proportionnelle à une variable connue \( T_k = f(N_k) \) dont le graphique est donné à la figure 1.

Pour spécifier entièrement le sondage il reste à trouver le nombre d’unités primaires à tirer. Or, \( T = \sum U T_k \) est aussi une quantité connue.

En se restreignant à un échantillonnage de taille fixe on aura donc \( \pi_k = m T_k/T \). On trouve \( m \) en portant cette valeur dans la contrainte de variance soit :

\[
N^2 \ V_o \ m = T \sum U \ \frac{N_k^2}{T_k} \left( \sigma^2 + \tau^2/n_k \right)
\]

Si, en première approximation, on prend \( T_k = N_k \), on obtient la formule simplifiée :

\[
m \ V_o = \sigma^2 + \tau^2/n^*
\]

Les données recueillies sur le test de recensement ont permis de retenir les valeurs numériques approximatives suivantes :

\[
\tau^2 = p^2 = 14 \cdot 10^{-4} \quad \sigma^2 = p = 4 \cdot 10^{-2}
\]

D’autre part, des mesures faites dans les ateliers ont permis d’estimer à 5 minutes le temps de manipulation d’une chemise de district et à 30 secondes le temps de saisie d’un BI. Le rapport \( C_o/C_1 \) vaut donc environ 10 d’où on tire \( n^* = 16 \) puis \( m = 40 \). Le contrôle porte donc sur 40 districts dont on extrait au total 640 bulletins.
Figure 1 : graphique de $\pi_k$ en fonction de $N_k$

Le contrôle de l’exhaustif, deuxième forme

Ou comment optimiser un sondage à deux degrés où les unités primaires sont elles mêmes stratifiées.

La dure réalité des choses nous amène à compliquer un peu le problème. En fait, il s’agit de contrôler deux types de documents : les bulletins individuels (BI) et les feuilles de logement (FL). On avait été conduit à négliger ces dernières, en première approximation, parce qu’elles sont moins susceptibles de receler des erreurs et que leur temps de codification est plus court (la moitié environ) que celui nécessaire pour un BI. Toutefois, dans certains districts, par exemple dans les communes très touristiques, on trouve une forte majorité de résidences secondaires et donc beaucoup de FL pour très peu de BI. Cette situation demande une étude particulière.

Il s’avère qu’elle correspond, de plus, à un problème assez général qui est le suivant.

Pour chaque unité primaire (ici les districts d’une UT) on connaît les effectifs $N_{kg}$ d’unités secondaires appartenant à G groupes (ici $g = 1, 2$) selon qu’on s’intéresse aux BI ou aux FL). La "population" de l’UP numéro $k$ vaut $N_{i+} = \sum N_{ig}$. celle du groupe $g$ vaut $N_{+g} = \sum_k N_{k,g}$. Comme dans ce qui précède on cherche avec quelle probabilité

L’exploitation

223
d’inclusion $\pi_k$ échantillonner l’UP numéro $k$, le nombre d’UP à tirer et l’allocation $n_{kg}$ de l’échantillon parmi les différents groupes dans l’UP $k$, sachant que ces $n_{kg}$ unités sont tirées par un SASSR parmi les $N_{kg}$ unités tirables.

**Optimisation sous modèle**

On postule, dans chacun des groupes, un modèle identique à celui formulé à la section précédente (ou sous une forme plus générale dans la remarque qui la termine).

Pour $g = 1$ à $G$ on aura donc :

$$V_g = E_V \text{Var}(\hat{P}_g) = N^{-2} \sum_U \frac{N_{kg}^2}{\pi_k} (\sigma_g^2 + \tau_g^2 / n_{kg})$$  \hspace{1cm} (8 - 1)

La fonction de coût est donnée par la forme générale (3-1). On va chercher à minimiser l’espérance du coût de sondage :

$$C_T = \sum \pi_k \left( c_0 + \sum_g n_{kg} C_g \right)$$  \hspace{1cm} (8 - 2)

sous les contraintes $V_g \leq v_g$ où les quantités $v_g$ sont fixées de façon extérieure, par exemple par la qualité des données qu’on veut obtenir et la rigueur du contrôle.

Sous cette forme, le problème peut s’avérer assez complexe. Nous allons écrire un "Lagrangien" général :

$$L = \lambda C_T + \sum \lambda_g V_g$$

Le problème posé fixe $\lambda = 1$ et les $\lambda_g$ sont des multiplicateurs à déterminer. Une variante simple consiste à fixer les $\lambda_g$ : on désire alors minimiser une combinaison linéaire donnée des variances sous une contrainte de coût. Dans toutes les hypothèses, on obtient par dérivation par rapport aux $n_{kg}$ (considérées comme des variables réelles) :

$$\lambda \pi_k^2 C_g = \lambda_g N_{kg}^2 \pi_g^2 \tau_g^2 / n_{kg}^2$$  \hspace{1cm} (8 - 3)

Les $\pi_k$ étant, pour l’instant, destinées à être connues à un facteur près, on peut écrire :

$$\pi_k n_{kg} = \left( \frac{\lambda_g C_g}{N_{kg}} \right)^{1/2} \frac{N_{kg}}{N_{+g}}$$  \hspace{1cm} (8 - 4)
Par sommation sur $k$ on en déduit que :

$$E n_{+g} = \sum_{U} \pi_k n_{kg} = \left( \frac{\lambda_g}{C_g} \right)^{1/2} \tau_g$$  \hspace{1cm} (8-5)

La taille totale de l'échantillon dans chaque groupe est donc directement liée au multiplicateur $\lambda_g$.

La dérivation du Lagrangien par rapport aux $\pi_k$ nous donne de nouvelles relations qui se simplifient miraculeusement si on utilise aussi (8-4). On obtient :

$$C_o = \sum_{g} C_g \left( \frac{\sigma_g}{\tau_g} \right)^2 n_{kg}$$  \hspace{1cm} (8-6)

où encore, si on introduit les nombres

$$n^*_g = \left( \frac{C_o}{C_g} \right)^{1/2} \frac{\tau_g}{\sigma_g} \text{ on écrit :}$$

$$\sum_{g} \left( \frac{n_{kg}}{n^*_g} \right)^2 = 1$$  \hspace{1cm} (8-7)

Les $n^*_g$ sont les nombres d'unités secondaires à tirer par $UP$ s'il n'y avait qu'un seul groupe ; $n_{kg}$ sera toujours inférieur à $n^*_g$.

De (8-4), (8-5) et (8-7) on tire les relations :

$$\pi_k^2 = \frac{1}{C_o} \sum_{g} \lambda_g \frac{\tau_g^2}{\left( N_{kg} / N_{+g} \right)^2}$$

Ainsi, les $\pi_k$ sont proportionnelles aux quantités $T_k$ telles que $T_k^2 = \sum_{g} \lambda_g \frac{N_{kg}^2}{N_{+g}^2}$ qui apparaissent comme la mesure de taille adéquate. Les relations (8-4) montrent que, à $k$ fixé, les $n_{kg}$ sont proportionnelles à $n^*_g \lambda_g^{1/2} \sigma_g N_{kg} / N_{+g}$, ce qui, compte tenu de (8-7) conduit à :

$$n_{kg} = n^*_g \lambda_g^{1/2} \sigma_g N_{kg} / N_{+g} T_k^{-1}$$

Des solutions possibles au problème précédent

a) Si les $\lambda_g$ étaient connus, c'est-à-dire si on minimisait $\sum_g \lambda_g V_g$ sous une contrainte de coût, alors (8-8) nous permettrait de calculer les $T_k$. 

L'exploitation

225
En reportant $\pi_k = m \frac{T_k}{T}$ ($T = \sum U_k$, $m$ nombre de districts à tirer)
dans la contrainte de budget $C_T \leq \sum C_T^g$, on trouve :

$$C_T^* = \frac{m}{T} \left( C_0 \sum_{k} T_k + \sum_{g} C_g^* \lambda_g^* \frac{1}{\sqrt{2 \pi}} \sigma_g \frac{N_k}{N_g} \right)$$

soit : $m = \frac{C_T^*}{\left( C_0 + \sum_{g} \frac{\lambda_g^*}{\sqrt{\sigma_g^2}} \frac{N_k}{N_g} \right)}$

Si un seul des $\lambda_g$ est différent de zéro, on trouve avec satisfaction le résultat donné à la fin de la section 6.

b) Le problème initial ($\text{min } C_T$ sous $V_g \leq \gamma_g$) se résoud assez facilement dans deux cas particuliers :

b1 - Dispersion maximale des groupes. Pour toute $UP_k$, on a $N_{kg} = N_k$ pour un certain $k$. Le problème est décomposé en $G$ problèmes distincts, chacun d'eux étant du type étudié aux sections 4, 5 et 6.

b2 - Dispersion minimale : la répartition est la même dans toutes les $UP$ ; autrement dit on a pour tout $k$ et $g$

$$N_{kg} = N_{kg}^{*} = \frac{N_k}{N} \left( \text{avec } N = \sum_{g} N_{kg} \right)$$

$T_k$ est alors proportionnelle à $N_k$, et les $N_{kg}$ sont des quantités indépendantes de $k$.

Avec $\pi_k = m N_{kg}^{*}/N$, on obtient en écrivant $V_g = \gamma_g$ :

$$m \gamma_g = \sigma_g^2 + \tau_g^2 / n_{kg}^* u_g$$

soit : $m = \frac{\gamma_g}{\sigma_g^2 + \tau_g^2 / n_{kg}^*}$

On obtient ainsi $G-1$ relations linéaires entre les $u_g^{-1}$ ce qui permet, en principe, de résoudre complètement le problème sachant que la somme des $u_g^2$ vaut 1.

Pour $G = 2$ (les BI, $g = 1$, les FL, $g = 2$) on obtient $m = 73$, $n_1^* u_1 = 15$ BI, $n_2^* u_2 = 1$ FL pour les données :

- $P_1 = 0,04$
- $P_2 = 0,01$
- $C_0 = 10 C_1 + 20 C_2$
- $\gamma_1 = (0,0075)^2$

$\sigma_1 = P_1$  
$\sigma_2 = P_2$  
$\tau_1 = P_1 (1 - P) - \tau_1^2 = P_1 - 2 P_1^2$  
$\tau_2 = P_2 - 2 P_2^2$  
$\gamma_2 = (0,0150)^2$
c) Une résolution numérique itérative du problème peut se faire de la façon suivante :

**Étape 1** : On fixe une allocation approximative de l'échantillon dans chaque groupe, soit $n_{+g}$ unités dans le groupe $g$. Pour y arriver on peut, par exemple, se servir de la solution approximative avec les hypothèses du point a) ou du point b).

**Étape 2** : La valeur des $\lambda_g$ est déterminée par les relations (8-5) :

$$\lambda_g = C_g \frac{n^2_{+g}}{\tau^2_g}$$

**Étape 3** : Les $\pi_k$ sont déterminées par les relations (8-8). La somme des $\pi_k$ fixe, en particulier, le nombre d'UP à tirer.

**Étape 4** : Les $n_{gk}$ sont déterminés par les relations (8-4). On peut ensuite itérer par retour à l'étape 2 en espérant que cet algorithme converge vers la solution d'optimisation.

d) **Remarque** : La probabilité de tirer une unité de type $g$ vaut :

$$\pi_k \frac{n_{kg}}{N_{kg}} = \left( \frac{\lambda_k}{C_g} \right)^{\frac{1}{2}} \frac{\tau_k}{N_{+g}}$$

Elle est donc uniforme et on en déduit la taille $n_g^+$ de l'échantillon. Pratiquement, il arrive que l'on fixe "autoritairement" les tailles des échantillons. Ceci revient à déterminer les $\lambda_g$ ou, implicitement, des variances $\gamma_g$. Ce résultat est assez naturel lui aussi.

**Sondage pour le contrôle de qualité "Colibri"**

**Position du problème**

Il s'agit d'estimer la proportion de bulletins présentant une erreur de codification dans l'"univers" $U$ de tous les bulletins codifiés une semaine donnée dans une direction régionale. Le caractère particulier du problème est le suivant : tous les bulletins $i$ sont déjà précédimentés ce qui permet, grâce à des informations tirées de l'essai de recensement, d'attribuer à chacun d'eux une variable numérique positive $X_i$ qui traduit sa "difficulté". Cette variable a été calibrée de façon à ce que $Y_i$ (qui vaut 1 en cas d'erreur et 0 sinon) ait une "espérance" proportionnelle à $X_i$.

Toujours pour les mêmes raisons de coût du contrôle, on est amené à envisager un sondage à deux degrés :
- au premier degré de sondage on tirera un échantillon $s_I$ de districts $k$ à probabilités inéguales $\pi_k$ à déterminer. On notera $\pi_{k1}$ les probabilités d'inclusion double pour cet échantillonnage,
au second degré de sondage, on tirera un échantillon \( s_k \) des bulletins (BI) dans le district échantillon \( k \). On notera \( \pi_{ijk} \) la probabilité d’inclusion du bulletin \( i \) dans le district \( k \), \( \pi_{ij} \) la probabilité d’inclusion du couple \((i, j)\) dans les districts

\[
s = \bigcup_{k \in s_1} s_k \text{ l'échantillon de BI.}
\]

On notera \( X_k = \sum_{i \in k} X_i \) le total des \( X_i \) dans le district \( k \),

\[
X = \sum_{k \in U} X_k = \sum_{i \in U} X_i \quad \text{et on adoptera des notations analogues pour toutes les variables.}
\]

\[
\sum_{i \in s_k} Y_i
\]

Le but est d’estimer une quantité de la forme \( R = \frac{\sum W_i}{\sum U} \) où \( W_i \) est une variable connue pour chaque bulletin. Cela pourra être \( W_i = 1 \) ou \( W_i = X_i \) selon la mesure qui semble la plus adéquate du taux d’erreur.

**Choix d’estimateur et variance**

a) Au niveau d’un district (UP numéro \( k \)) il est naturel d’estimer le total \( Y_k \) des \( Y_i \) pour \( i \in k \) par le ratio :

\[
\hat{Y}_k = X_k \left( \sum_{i \in s_k} \frac{Y_i}{\pi_i k} \right) / \left( \sum_{i \in s_k} \frac{X_i}{\pi_i k} \right) = X_k \hat{a}_k
\]

Ici \( \hat{a}_k \) estime \( a_k = Y_k/X_k \) avec un faible biais.

b) Pour estimer le ratio \( Y/X \) on utilisera :

\[
\hat{a} = \frac{\sum s_i \hat{Y}_k / \pi_k}{\sum s_i X_k / \pi_k} = \frac{\sum s_i \hat{a}_k X_k / \pi_k}{\sum s_i X_k / \pi_k}
\]
c) Si on veut estimer $R$, on remarquera que :

$$R = \frac{Y \cdot X}{X \cdot W}$$

où $X$ et $W$ sont des totaux connus (difficulté totale et nombre total de bulletins, par exemple). Comme la variable $X_i$ a été choisie pour sa bonne corrélation avec $Y_i$, un estimateur a priori intéressant de $R$ sera :

$$\hat{R} = \hat{a} \frac{X}{W}$$
de sorte que la véritable question semble porter sur l’estimation de $a = \sum_{k} a_k X_k / X$

d) on aura :

$$\text{Var}(\hat{a}) = \text{Var}(E \hat{a} \mid s_1) + E \text{Var}(\hat{a} \mid s_1)$$

Pour le premier terme, compte tenu du fait que $\hat{a}_{k}$ estime (à peu près) sans biais $a_k$, on peut écrire :

$$\text{Var}(E \hat{a} \mid s_1) = \frac{1}{X^2} \text{Var} \left( \sum_{s_1} \frac{(a_k - a) X_k}{\pi_k} \right)$$

$$= \frac{1}{X^2} \left( \sum_{k} \frac{(a_k - a)^2 X_k^2}{\pi_k^2} + \sum_{k \neq 1} (a_k - a) (a_1 - a) \frac{X_k X_1 \pi_k \pi_1}{\pi_k \pi_1} \right)$$

(11-1)

Pour le second terme, on a, conditionnellement à $s_1$ :

$$\text{Var} \left( \sum_{s_1} \frac{\hat{a}_k X_k / \pi_k}{\sum_{s_1} X_k / \pi_k} \right) = \left( \sum_{s_1} \frac{X_k / \pi_k}{s_1} \right)^{-2} \sum_{s_1} \text{Var} \left( \hat{a}_k \right) \frac{X_k^2}{\pi_k^2}$$

L’espérance de cette quantité vaut approximativement :

$$X^{-2} \sum_{k} E \text{Var}(\hat{a}_k \mid s_1) \frac{X_k^2}{\pi_k}$$

(11-2)

avec :

$$\text{Var}(\hat{a}_k \mid s_1) = \text{Var} \left( \sum_{s_1} \frac{Y_i / \pi_{i \mid k}}{X_i / \pi_{i \mid k}} \right) = \frac{1}{X_k^2} \text{Var} \sum_{s_1} \frac{Y_i - a_k X_i}{\pi_{i \mid k}}$$

$$= \frac{1}{X_k^2} \left( \sum_{i \in k} \frac{(Y_i - a_k X_i)^2}{\pi_{i \mid k}} + \sum_{i \neq j} \frac{(Y_i - a_k X_i)(Y_j - a_k X_j) \pi_{i \mid k} \pi_{j \mid k}}{\pi_{i \mid k} \pi_{j \mid k}} \right)$$

Comme dans les parties précédentes, nous arrivons à des formules complexes et, finalement, inutilisables. Un modèle va nous simplifier un peu l’existence.
Intervention d’un modèle

Il aura la même structure que ceux qui ont déjà servi antérieurement :

a) Les \( a_k \) seront des variables aléatoires indépendantes de même espérance et de même variance :
\[
E a_k = a \quad \text{Var}(a_k) = \sigma^2
\]
La variance prend en compte l’influence de l’opérateur ou de l’opératrice, qu’on renonce à isoler, mais aussi celle du jour de la semaine, de l’heure dans la journée, de certains jours du mois etc.

En revanche, par référence au modèle utilisé pour l’exhaustif, il ne contient plus l’effet lié à la difficulté du district. Celui-ci est pris en compte, du moins on l’espère, par la variable auxiliaire \( X_k \) qui intervient au point suivant.

b) Conditionnellement à \( a_k \), les \( Y_i \) du district \( k \) sont des variables de Bernoulli indépendantes avec \( E(Y_i \mid k) = a_k X_k \) et donc :
\[
\text{Var}(Y_i \mid k) = a_k X_k - a_k^2 X_i^2
\]

Remarque : La variable \( X_i \) n’a pas de véritable sens concret, et n’est d’ailleurs définie qu’à un facteur d’échelle près. En revanche \( a X_k \) et \( \sigma X_k \) ont une interprétation physique invariante, car ce sont des probabilités. Dans tout ce qui suit il faudra toujours garder en tête que les résultats devront être invariants si les \( X_i \) sont multipliés par un facteur arbitraire à condition que a et \( \sigma \) soient divisés par le même facteur. En particulier \( \text{Var} (a) \) n’a pas de sens “concret”. Seule \( \text{Var} (a X) \) en a un.

Comme précédemment nous allons étudier la variance anticipée, espérance sous modèle de la somme de (11 - 1) et (11 - 2).

Pour le premier terme, l’espérance des produits croisés est nulle, comme de bien entendu. L’espérance sous modèle de ce terme est donc :
\[
X^{-2} \sigma^2 \sum_k X_k^2 / \pi_k
\]

Pour le second terme on trouve : (vu la définition du 11 a)
\[
X^{-2} \sum_k \frac{\sum_i X_i^2}{\pi_k} \cdot \frac{1}{X_k^2} \sum_i \frac{a_k X_i - a_k^2 X_i^2}{\pi_i \mid k} = X^{-2} \sum_k \frac{1}{\pi_k} \sum_i \frac{aX_i - (a^2 + \sigma^2) X_i}{\pi_i \mid k}
\]
Globalement donc :
\[
E \text{Var}(\hat{a} X) = \sigma^2 \sum_k X_k^2 / \pi_k + \sum_{k \in U} \frac{1}{\pi_k} \sum_{i \in k} \frac{aX_i - (a^2 + \sigma^2) X_i}{\pi_i \mid k}
\]
Ici pas de miracle algébrique. Pour simplifier nous admettons que \( (a^2 + \sigma^2) X_i^2 \) est négligeable devant \( aX_i \). Numériquement on peut attendre \( aX_i = 2 \times 5 \times 10^{-2} \) et \( (a^2 + \sigma^2) X_i^2 = 3 \times 30 \times 10^{-4} \).
D'où notre approximation :

\[ E \varnothing \text{Var}(\hat{\theta} | X) = \sigma^2 \sum_{k \in U} X_k^2 / \pi_k + a \sum_{k \in U} \frac{1}{\pi_k} \sum_{i \in k} \frac{X_i}{\pi_i} \]

\[ = \sigma^2 \sum_{k \in U} \frac{X_k^2}{\pi_k} + a \sum_{i} \frac{X_i}{\pi_i} \]

**Optimisation**

Nous utiliserons la fonction de coût suivante :

\[ C = \sum_{i \in k} C_o + C_1 n_k \]

Ici, \( n_k = \sum_{i \in k} \pi_i \) est la taille de l'échantillon tiré dans le district \( k \) (supposé de taille fixe à \( s_1 \) fixé). Son espérance vaut :

\[ C_T = \sum_{k \in U} \pi_k (C_o + C_1 n_k) \]

Posons \( \pi_i | k = n_k P_i \) (avec \( \sum_{i \in k} P_i = 1 \)) et \( Q_k = \pi_k n_k \)

Le problème d'optimisation est maintenant :

\[ \text{Min} : C_o \sum_{k} \pi_k + C_1 \sum_{k} Q_k \]

sous : \[ \sigma^2 \sum_{k} \frac{X_k^2}{\pi_k} + a \sum_{k} \frac{1}{Q_k \sum_{i \in k} \frac{X_i}{P_i}} \leq \gamma_0 \]

Sous cette forme, on constate avec plaisir qu'on peut minimiser les termes en \( \sum X_i / P_i \) indépendamment du reste. Autrement dit, \( n_k \) n'a pas d'incidence sur ce terme.

Laissons l'optimisation du second degré de tirage pour plus tard et notons seulement \( S_k^* \) la valeur optimisée. Avec un multiplicateur de Lagrange \( \lambda \) on obtient par dérivation par rapport aux \( \pi_k \) puis au \( Q_k \) :

\[ *C_o = \lambda \sigma^2 \frac{X_k^2}{n_k^2} \] soit \( \pi_k \) proportionnel à \( X_k \)  

(13 - 1)
\[ C_1 = \lambda \frac{2}{Q_k} \]
d'où \[ n_k = \left( \frac{C_0}{C_1} \right) \frac{1}{\sigma} \frac{\sqrt{a}}{\sqrt{2}} \frac{S_k^2}{X} \] \hspace{1cm} (13 - 2)

En particulier on tirera les districts avec des probabilités proportionnelles à leur difficulté totale.

Passons maintenant au tirage infra-district (deuxième degré de sondage).

Commençons par un cas simple et naïf : on tire les bulletins individuellement. La minimisation conduit à \( P_i \) proportionnelle à \( \sqrt{X_i} = S_i \)

Un calcul simple nous montre qu'alors \( S_k^* = S_k = \sum_{i \in k} S_i \). Ceci nous permet de calculer \( n_k \) grâce à (13 - 2) et notre problème est entièrement résolu.

En fait les choses sont plus compliquées. Pour des raisons assez naturelles, on ne sélectionnera les BI que par ménages entiers. Autrement dit le sondage au second degré est un sondage en grappes. Les valeurs de \( P_i \) seront les mêmes, soit \( P_m \), pour tous les membres d’un même ménage \( m \).

Notons par \( X_m \) la somme des \( X_i \) des individus \( i \) du ménage \( m \).

Le problème est donc de : Minimiser \( \sum X_m 

P_m = \sqrt{\frac{X_m}{n_m}} / \sum n_m \sqrt{X_m} \)

Avec \( X_m = X_m / n_m \), difficulté moyenne du BI du ménage \( m \).

Par suite, on trouve :

\[ S_k^* = \sum n_m \sqrt{X_m} \]

Cette solution nous permet de déterminer le nombre \( n_k \) de BI à tirer grâce à (13 - 2). Le nombre de ménages, en revanche n’est pas déterminé. Cette difficulté était prévisible. La fonction de coût, en effet, ne fait pas intervenir cette contrainte. Pour obtenir le nombre \( m_k \) de ménages à tirer, on s’arrangera de façon à ce que l’espérance du nombre de BI soit égale à \( n_k \). Elle vaut :

\[ m_k \left( \sum n_m \sqrt{X_m} \right) / \sum \sqrt{X_m} \]

232

INSEE Méthodes n°52-53
d'où \[ m_k = n_k \frac{\sum \sqrt{X_m}}{\sum n_m \sqrt{X_m}} \]

Compte tenu de (13 - 2) on a aussi :

\[ m_k = \left( \frac{C_0}{C_1} \right)^{\frac{1}{2}} \frac{a_{\frac{1}{2}}}{\sigma} \frac{\sum \sqrt{X_m}}{X_k} \]

et la probabilité de tirer un ménage vaut alors :

\[ m_k \frac{\sqrt{X_m}}{\sum \sqrt{X_m}} \]

Nous avons obtenu une solution complète du problème.

**Remarque 1** : Dans les deux cas qui ont été traités, \( S_k^* \) est multiplié par \( C^{1/2} \) si les \( X_i \) sont multipliés par \( C \). La formule qui donne \( n_k \) est donc bien invariante à l'échelle de mesure.

**Remarque 2** : La solution du cas 2 privilégie le tirage de petits ménages compliqués.

**Remarque 3** : Ici comme dans les parties précédentes nous déterminons des probabilités d’inclusion simple mais pas des probabilités d’inclusion double. L’algorithme de tirage, qui fixe ces dernières, est donc sans influence. Ceci est relativement naturel si nous nous disons que l’information auxiliaire utilisée pour optimiser le tirage déterminera les \( \pi_k \) et \( \pi_{i|k} \) mais ne peut pas avoir d’influence sur les probabilités doubles.

**Remarque 4** : Ce problème fait apparaître des résultats un peu surprenants sur lesquels il est utile de réfléchir un peu.

Dans un premier cas, nous avons supposé qu’on pouvait isoler chaque bulletin. On tirait alors ceux-ci avec des probabilités proportionnelles à leur difficulté individuelle. On supposait, dans une certaine mesure, que le coût d’utilisation de l’information individuelle était nul.

Dans le second cas, la réalité du contrôle, ce coût était considéré comme infini et la seule information ayant un coût négligeable était celle relative à l’ensemble du ménage. La solution fait alors apparaître des probabilités de tirage des individus (BI) qui est fonction de la difficulté moyenne de codification des bulletins de l’ensemble du ménage auquel appartient cet individu.

Il en irait de même en ce qui concerne le tirage des districts. Si on sait y distinguer les BI, on les tire avec des probabilités proportionnelles à la difficulté totale ; à l’intérieur des districts, on tirera des BI difficiles avec une plus grosse probabilité. Supposons, au contraire, qu’on ne sache pas distinguer les BI à l’intérieur des districts. Ce serait le
cas, par exemple, si la désignation des BI à contrôler ne pouvait pas se faire en temps réel par suite d’une organisation du traitement inadéquate. On tirerait alors les districts proportionnellement à leur difficulté moyenne : à l’intérieur des districts, on serait obligé de réaliser des sondages aléatoires simples.

Dans le premier cas le sondage privilégiera les gros districts à l’intérieur desquels on tirera plutôt les BI difficiles. Dans le second cas, on privilégiera les petits districts difficiles à l’intérieur desquels on tirera des bulletins à probabilités égales. Dans les deux cas, on cherchera à augmenter la probabilité de sonder des BI difficiles. La différence réside simplement dans la possibilité (c’est-à-dire le coût) de mobiliser l’information au moment où on en a besoin.

**Variance anticipée optimale et application**

Après quelques manipulations algébriques, on trouve la valeur de la variance optimale :

\[ E[\var (X)]_{OPT} = \frac{(\sigma X)^2}{m} \left[ 1 + \frac{a}{\sigma} \frac{a^{-1/2} S^*}{X} \left( \frac{C_1}{C_0} \right)^{1/2} \right] \]

Cette forme respecte le caractère homogène des différents facteurs. On a, en particulier

\[ \frac{a^{-1/2} S^*}{X} = \frac{a^{1/2} S^*}{aX} \]

le dénominateur est interprétable comme un nombre total d’erreurs dans un lot, tandis que le numérateur est homogène à une taille.

L’application pratique et numérique de cette théorie repose sur des hypothèses concernant les ordres de grandeurs des différents paramètres (ce qui demande qu’on puisse les raccrocher à une interprétation physique simple). Dans la phase de préparation du recensement, sans mesures préalables très précises, on a utilisé les valeurs \( \sigma/a = 0.5 \) et \( C_1 / C_0 = 0.1 \)

À la suite de diverses hypothèses sur les autres paramètres et de discussion entre experts, il a été décidé un contrôle portant sur 50 districts chacun d’eux étant contrôlé pour environ 20 BI (par région et par semaine). Cet ordre de grandeur initial pouvait, évidemment, être modulé dans la suite du contrôle, les paramètres du modèle pouvant être réestimés après chacun d’eux.
Vue d’ensemble

Cette étude est basée sur l’utilisation de modèles de superpopulation pour anticiper la variance d’une mesure par sondage avant le sondage. On arrive, en utilisant des modèles simples qu’on voudrait néanmoins assez réalistes, à des expressions plus ou moins complexes qu’on parvient à optimiser, parfois rigoureusement, quelquefois de façon approximative.

La solution du dernier de ces problèmes fait apparaître un facteur assez peu étudié dans les problèmes d’optimisation de plan de sondage : le coût lié à la mobilisation d’une information individuelle.


3.3 - LA SAISIE DE L’EXHAUSTIF LÉGER (RP METROPOLE)

Jacques Darmon

Cet article présente pour l’essentiel le bilan de cette opération pour ce qui a trait à la gestion des rapports avec les façonniers sous-traitants. Il ne reprend pas certains développements concernant la gestion informatique interne à l’INSEE.

Le bilan de l’opération sur le plan de la qualité de la saisie est présenté dans l’article sur le contrôle de qualité exhaustif léger. Cet article est découpé en deux parties.

Dans la première partie, nous dressons un historique rapide de l’opération et nous réunissons des éléments chiffrés qui pourront servir dans l’avenir. Dans la deuxième partie, nous résumons et faisons la synthèse des rapports que les directions régionales nous ont adressés.

1 - Historique

1.1. Une opération plus dense

Pour la première fois, l’exploitation exhaustive a été exécutée avant le traitement des échantillons "lourds".

Par là-même, elle a concerné la totalité des bulletins au lieu des 3/4 habituellement.

D’autre part, la saisie exhaustive devait servir à la vérification de la population légale pour les unités de traitement (UT) traitées à l’exhaustif suffisamment tôt pour entrer dans le cadre des délais réglementaires d’établissement de la population légale, ce qui induisait une synergie forte entre les deux opérations. Le délai de réalisation de l’exhaustif se trouvait ainsi réduit à 8 mois environ, alors que la saisie légère des 3/4 du RP82 s’était étalée sur quelques 18 mois.

C’est donc une opération environ 3 fois plus "dense" qu’en 1982 qu’il a fallu réaliser.
1.2. *Un poste de travail commun*

La vérification de la population légale devait bénéficier des comptages issus de la saisie exhaustive et en échange la préparation des unités de traitement (UT) avant expédition chez le façonnier profitait de la saisie des bordereaux de district et de commune prévue dans le cadre de la "Population Légale".

La synergie entre les 2 opérations devait se concrétiser par la mise en place d’un poste de travail "écran-clavier" commun. Des problèmes informatiques non résolus ont conduit à mettre en place deux applications séparées.

1.3. *Le choix des façonniers*

A la suite de l’appel à candidature, 22 sociétés de saisie à façon se sont manifestées. Leurs offres de services (incluant le prix proposé), ainsi que le jeu d’essai qu’elles avaient à saisir, sont parvenus en juillet 1989.

La répétition du printemps 1988 avait montré que ce jeu d’essai n’apporte guère de certitude sur l’aptitude des candidats à faire face à une opération aussi lourde que le RP. C’est donc surtout sur des critères de prix et de références disponibles que le choix final a été arrêté.

Nous savions aussi qu’il fallait éviter les deux écueils :

- trop de sous-traitants, dont certains trop "petits", ce qui rend le suivi plus difficile ;
- peu de sous-traitants, ce qui transforme en catastrophe la défaillance toujours possible de l’un d’entre eux.

Dès septembre 1989, le choix était arrêté et la procédure de passation des marchés engagée.

Les 84 millions d’enregistrements, répartis en 28 lots de 3 millions chacun, ont été attribués aux 8 façonniers retenus selon la grille suivante :

238
<table>
<thead>
<tr>
<th></th>
<th>Prévus (lots)</th>
<th>Réalisés (millions d’enregistrements)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>15,16</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>16,19</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>12,68</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>12,89</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>11,77</td>
</tr>
<tr>
<td>F</td>
<td>3</td>
<td>8,66</td>
</tr>
<tr>
<td>G</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>2</td>
<td>6,20</td>
</tr>
<tr>
<td>TOTAL</td>
<td>28</td>
<td>83,55</td>
</tr>
</tbody>
</table>

On trouve en annexe 1, la répartition par façonnier des bulletins des différentes Directions régionales (DR)\(^1\)

### 1.4. Une nouvelle technique d’échantillonnage

Une cause importante de difficultés à l’exhaustif du RP82 fut le nombre élevé de divergences entre l’échantillon de contrôle et le fichier du façonnier, directement lié à la taille de cet échantillon (5 200 bulletins individuels en moyenne).

En effet ce volume important de divergences à examiner introduisait des délais importants dans la procédure d’acceptation ou de refus du travail des façonniers. Aussi la Division "Méthodes Statistiques et Sondages" a-t-elle mené une étude spécifique pour améliorer l’efficacité de l’échantillonnage qui a aboutit à la règle pratique suivante.

Dans chaque unité de traitement (UT), avec des probabilités proportionnelles au nombre d’habitants recensés, on tire 100 districts et on saisit 15 BI dans chacun d’eux ainsi que les DIC et FL qui les contiennent. En théorie ces 15 bulletins auraient dû être tirés aléatoirement dans le district concerné, mais le coût de recherche aurait été prohibitif. A titre de compromis, la saisie commence alternativement par le début, le milieu, ou la fin du district. La taille de l’échantillon résultant d’un tel tirage est de l’ordre de 2200 enregistrements (bulletin individuel et feuille de logement) par unité de traitement.

\(^{1}\) Les noms des sociétés concernées ont été masqués car des données sur la qualité du travail sont fournies par la suite.

*L'exploitation*  
239
Les conséquences favorables de la taille plusmodeste de l’échantillon furent les suivantes :

- la constitution des échantillons de contrôle par les ateliers de chiffrement-saisie des DR fut de meilleure qualité pour un coût moindre ;

- le volume des listings de divergences fut réduit de beaucoup (on avait, de plus, éliminé les divergences sans incidence sur la qualité) ;

- en conséquence, la charge de travail pour examiner les divergences fut plus légère et le délai pour prévenir le façonnier de ses erreurs plus court.

1.5. *Le calendrier a été tenu*

Dès la mi-avril, tous les façonniers disposaient d’une UT de volume très réduit pour chacun de leurs ateliers afin de tester les procédures. Comme prévu, la saisie réelle pouvait commencer au début du mois de mai. Malgré la défaillance d’un des façonniers les délais annoncés ont pu être tenus, comme en témoigne le tableau d’avancement mensuel suivant.

<table>
<thead>
<tr>
<th>Mois</th>
<th>Millions d’enregistrements</th>
<th>Cumul</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mai</td>
<td>3,3</td>
<td>3,3</td>
<td>4,0</td>
</tr>
<tr>
<td>Juin</td>
<td>7,3</td>
<td>10,6</td>
<td>12,7</td>
</tr>
<tr>
<td>Juillet</td>
<td>13,2</td>
<td>23,8</td>
<td>28,5</td>
</tr>
<tr>
<td>Août</td>
<td>9,4</td>
<td>33,2</td>
<td>39,8</td>
</tr>
<tr>
<td>Septembre</td>
<td>10,4</td>
<td>43,6</td>
<td>52,2</td>
</tr>
<tr>
<td>Octobre</td>
<td>14,3</td>
<td>57,9</td>
<td>69,3</td>
</tr>
<tr>
<td>Novembre</td>
<td>11,0</td>
<td>68,9</td>
<td>82,5</td>
</tr>
<tr>
<td>Décembre</td>
<td>10,6</td>
<td>79,5</td>
<td>95,2</td>
</tr>
<tr>
<td>Janvier</td>
<td>3,8</td>
<td>83,3</td>
<td>99,8</td>
</tr>
</tbody>
</table>
1.6. La qualité obtenue

Sur les 242 UT de la Métropole, 4 seulement ont fait l'objet d'un rejet officiel pour dépassement des seuils admis, suivi de réfection.

Tous les façonniers ne disposant pas d'un "validateur de lot" efficace, de plus nombreuses UT comportaient, en nombre limité, des erreurs de structure, de report d'identifiant ou de doubles. Les erreurs de ce type étaient signalées à l'atelier responsable qui les corrigeait et renvoyait très rapidement une nouvelle bande ; cette procédure a permis d'obtenir des fichiers conformes sans peser sur les délais.

Le taux contractuel d'enregistrements erronés ne devait pas dépasser 3 % pour les BI et 1 % pour les DIC-FL (le taux retenu pour la décision d'acceptation ou de rejet étant supérieur pour tenir compte des aléas du contrôle par sondage). Le tableau suivant donne les valeurs moyennes par façonnier. On serait tenté d'en déduire que les seuils auraient pu être abaissés, mais alors il y aurait eu plus d'UT refusées, entraînant des tensions sur les délais.

<table>
<thead>
<tr>
<th></th>
<th>Taux d'erreur BI</th>
<th>Taux d'erreur FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0,6 %</td>
<td>0,2 %</td>
</tr>
<tr>
<td>D2 *</td>
<td>1,3 %</td>
<td>0,2 %</td>
</tr>
<tr>
<td>F</td>
<td>1,3 %</td>
<td>0,6 %</td>
</tr>
<tr>
<td>E1</td>
<td>1,5 %</td>
<td>0,5 %</td>
</tr>
<tr>
<td>A</td>
<td>1,8 %</td>
<td>0,5 %</td>
</tr>
<tr>
<td>D1</td>
<td>1,9 %</td>
<td>0,3 %</td>
</tr>
<tr>
<td>H</td>
<td>2,0 %</td>
<td>0,5 %</td>
</tr>
<tr>
<td>C</td>
<td>2,2 %</td>
<td>0,7 %</td>
</tr>
<tr>
<td>E2 *</td>
<td>2,3 %</td>
<td>0,9 %</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>1,5 %</td>
<td>0,5 %</td>
</tr>
</tbody>
</table>

* deux des sept façonniers contractants avaient un co-traitant.
2 - Le travail en région

2.1. Préparation des UT

2.1.1 Taille des UT

Si l’on excepte les UT du démarrage à 100 000 BI, le volume moyen des UT s’établit à 255 000 BI.

Une grande majorité des DR a trouvé que ce nombre était un bon compromis entre un volume moindre qui facilite la manutention et le stockage des documents, et un volume plus important qui allège la constitution et la vérification de l’échantillon de contrôle.

De plus, il permet que l’UT tienne sur une seule bande magnétique et il rend supportable la perte du façonnier en cas de rejet.

Certaines directions régionales auraient, quant à elles, souhaité s’en tenir au volume initialement prévu de 300 000 BI.

2.1.2 Tirage de l’échantillon

La méthode proposée par la Division "Méthodes Statistiques et Sondages" a réduit très sensiblement (de 3 à 1 par rapport au RP82) le volume des échantillons de contrôle.

Quelques DR ont signalé que cette opération, qui devait être exécutée très rapidement, a été "lourde et pénible", demandant beaucoup de manutention et nécessitant beaucoup de place. En effet, il fallait retrouver les districts sondés dans les caisses où les bulletins attendaient depuis leur retour des mairies, les extraire, les traiter et les reclasser.

Une direction régionale pense que sa tâche aurait été facilitée si les communes auxquelles appartenaient les districts sondés avaient pu être déterminées dès la constitution de l’UT, pour permettre un rangement facilitant l’extraction ultérieure. Un tirage à deux niveaux aurait forcément diminué la précision de l’échantillon, mais peut-être dans des limites supportables.

Quelles qu’aient été les difficultés rencontrées au tirage, il n’en reste pas moins que le faible volume de cet échantillon à saisir et à vérifier a permis d’alimenter nos sous-traitants au rythme voulu. Sans anticiper sur le point 3, on peut déjà dire ici que le nombre proportionnellement réduit des anomalies à examiner a également permis de mieux tenir les délais qui nous étaient impartis pour accepter ou refuser les UT au retour.
2.1.3 Instruction de chiffrement-saisie

Ce document s'était voulu plus clair et surtout plus léger que son homologue du RP82. Il semble qu'il ait été ainsi reconnu et, dans l'ensemble, apprécié.

La codification du STATUT a paru complexe à certains mais elle n'intervenait que dans les 7% de cas ambigus et elle remplaçait la codification du "Type d'activité" qui était encore plus complexe.

Pour la codification du LIEU DE TRAVAIL, il aurait été souhaitable que l'instruction distingue mieux entre les cas de lieu de travail "variable" et lieu de travail "mobile".

Il manquait, à l'évidence, un paragraphe concernant la saisie des bordereaux de district vides.

Pour ce qui concerne les BI "vierges" rajoutés au cours des opérations de population légale, il est nécessaire de les avoir marqués d'un certaine façon pour qu'ils donnent lieu à la saisie d'un enregistrement à blanc ; en l'absence de marque, ils doivent être ignorés. Sans cette légère exigence, on peut être conduit à saisir n'importe quoi.

Les compléments que nous avons apportés à l'instruction peu après le démarrage, ont été ressentis par certains comme des modifications de consignes. Dans notre esprit, ce devait être des précisions utiles pour la suite du travail, toutes suggérées par le dialogue fructueux que nous avions établi avec les différents ateliers tant de l'INSEE que des façonniers.

2.1.4 Chiffrement-saisie de l'échantillon de contrôle

Les questionnaires de l'échantillon de contrôle du RP82 faisaient l'objet d'une transcription sur un bordereau de saisie avant d'être transmis à l'atelier. Le même processus était prévu pour la répétition générale de 1988, mais il a été dénoncé par la quasi totalité des DR. Pour le RP90, nous avons donc mis en place un chiffrement-saisie direct sur multiclavier effectué par le personnel de l'atelier de dactylocodage.

Dans une majorité des DR, les dactylocodeuses se sont bien adaptées à ce travail nouveau pour elles, y trouvant un enrichissement de leur tâche habituelle. Dans d'autres, cela fut moins bien vécu. Une direction régionale avance pour raison le caractère très intermittent de ce travail pour une région moyenne ou petite.
Les séances de formation préalable des monitrices ont été appréciées ; la seule critique porte sur la date jugée trop précoce de certaines de ces séances. Le support de cette formation a été bien reçu ; il a été jugé souhaitable de disposer d'autant de ces supports que d'opératrices.

Le programme de saisie a fait l'objet de quelques critiques, surtout dans sa première version. Des DR l'ont complété par des programmes de contrôles spécifiques. Dans l'avenir, il faudrait peut-être prévoir à sa place un poste de travail adapté à ce type de tâche, apportant une aide efficace au chiffrement.

Dans un tiers des districts sondés, les 15 BI saisis devaient l'être en commençant par le dernier et en remontant. Cette contrainte a souvent été jugée pénible. Est-on sûr qu'il existe un effet "fin de district" qui impose cette acrobatie ? Il faudrait vérifier s'il n'est pas suffisant d'alterner "début" et "milieu".

Plusieurs DR se sont plaintes du délai souvent trop tendu entre la constitution d'une UT et son départ vers le façonnier.

Dans ce laps de temps, il fallait saisir les pages récapitulatives des bordereaux de district, effectuer un aller-retour au Centre national informatique de Paris (CNIP) pour tirer l'échantillon, extraire les districts sondés, chiffrer-saisir l'échantillon, réaliser un nouvel aller-retour au CNIP pour contrôle, itérer en cas d'erreur, ... ; le tout en craignant de laisser le façonnier en rupture de charge. Dans un tel contexte, il est compréhensible que le temps consacré au chiffrement-saisie ait paru trop court pour permettre un travail soigné et serein.

Ces difficultés ne sont pas provenues seulement du calendrier général très tendu de l'opération mais également des problèmes informatiques mentionnés en 1-2 ci-dessus.

2.1.5 Relations avec la DG et le CNIP

Tous les rapports confirment l'impression que nous avions ressentie d'une excellente collaboration entre toutes les unités ayant participé à cette opération.

2.1.6 Problèmes matériels de constitution des UT

Pour la totalité de l'exhaustif, une évaluation a priori avait donné une masse de 600 tonnes et un volume de 800 m3.

La lecture des différents rapports suggère une image unique : celle d'un "raz-de-marée" de questionnaires, de colis, de palettes, dans un espace toujours trop étroit.
Ces 600 tonnes de papier ont dû être manipulées à plusieurs reprises par un personnel essentiellement féminin. Mais nous n'avions, à ce sujet, aucune solution alternative à proposer. Jusqu'à présent un recensement de population a toujours accouché de cette montagne de papier.

Une initiative a été, en général, bien accueillie. L'imprimerie nationale devait livrer les questionnaires vierges dans des cartons ne dépassant pas 10 kilos et il avait été décidé de réutiliser ces emballages pour les transferts DR- façonnier et retour. Cette fragmentation en colis "à taille humaine" n'a certes pas changé la masse totale mais, de l'avis de certains, a facilité la manutention.

Certaines DR sont déjà confrontées en temps normal à des problèmes liés à la taille et à la disposition des locaux qu'elles occupent. Une opération de cette ampleur ne peut que rendre ces problèmes encore plus sensibles. Une DR signale que les chassés-croisés entre les UT nécessitent, à son sens, une place au sol double de cette utilisée pour l'archivage courant.

C'est pourquoi nous avons vu fleurir le "système D". A titre anecdotique, une DR a installé sur son parking intérieur une tente marabout louée qui a desserré sa contrainte d'espace ; une autre a "réquisitionné" deux salles de réunion et deux garages à voitures.

Notons cependant que certaines des DR ont obtenu l'autorisation de louer pour les opérations liées au RP des locaux plus vastes et mieux adaptés. Cette pratique parait plus logique et devrait s'appliquer à un plus grand nombre de cas.

2.2. Relations avec les façonniers

2.2.1 Désignation du façonnier

La conviction unanime de toutes les DR est que "le façonnier idéal est celui qui traite la totalité des UT de la DR et son atelier doit, sinon être situé dans la ville siège de la DR, du moins ne pas en être trop éloigné."

Toutes les DR qui ont vu ce souhait exaucé expriment leur satisfaction, (à une seule réticence près, quand la puissance insuffisante de l'atelier du façonnier a créé une file d'attente). Par contre chez les autres, on lit le regret de ne pas avoir bénéficié d'une telle facilité, jugée légitime.

Choisir nos sous-traitants n'a pas été une opération facile. Il s'est agi de trouver un petit nombre de sociétés (moins d'une dizaine) dont les références pouvaient justifier notre confiance et capables de réaliser en 7 mois la saisie de 84 millions d'enregistrements à un prix unitaire acceptable.

L'exploitation
Imposer aux ateliers une localisation géographique est d'une part, interdit dans les marchés de l'État, et d'autre part, aurait constitué une contrainte forte qui n'aurait pu que diminuer le poids accordé aux arguments de qualité et de sécurité qui furent les critères fondamentaux du choix.

Une fois connues les sociétés retenues et l'implantation de leurs ateliers, nous avons fait en sorte de minimiser le chemin parcouru par les questionnaires et de créer des couples DR-façonnier aussi stables que possible.

Mais il y avait d'emblée une certaine inadéquation entre le volume attribué à chaque façonnier et le volume disponible dans la ou les DR les plus proches. Puis nous avons dû nous défaire d'un façonnier défaillant. Et surtout, il y a eu le souci constant que jamais aucun façonnier ne se trouve, ne serait-ce qu'une heure, en rupture de charge.

Et ce résultat a pu être atteint sans trop de chassés-croisés. La répartition DR-façonnier se trouve en annexe 1. Les quelques chiffres suivants résument la situation :

- 4 DR n'ont eu affaire qu'à un seul et 11 DR qu'à deux façonniers ;
- parmi les 15 DR qui ont eu plus d'un façonnier, la part du principal a été supérieur à
  - 90 % pour 5 d'entre elles
  - 75 % pour 4 autres.

2.2.2 Prise en charge des documents par le transporteur

Pour la quasi totalité des UT, les transporteurs sont arrivés à l'heure prévue et ont normalement chargé les palettes préparées à leur intention. Mais sans en exagérer l'importance, nous devons citer les quelques bavures relevées par les DR.

Nous avions pourtant pris soin de faire figurer en annexe au cahier des charges

- les plages horaires pendant lesquelles le transporteur pouvait accéder à chacune de nos directions régionales,
- le gabarit maximum permis au camion, ainsi que
- la nécessité que celui-ci dispose d'un hayon élévateur.

Malgré cette précaution, ces trois règles ont été quelques fois violées.
Par contre, nous n’avions pas imposé que les documents d’une UT soient enlevés d’un seul coup. Il est arrivé que des transporteurs utilisent des véhicules trop petits nécessitant plusieurs voyages ; ce qui a représenté une gêne pour certaines DR, alors que d’autres s’en sont bien accommodées.

Le marché ne prévoyait pas, non plus, à qui incombait le trajet “hall de la DR - camion”, et cette lacune a été, dans un cas, cause de tension.

Il ressort de cela que le cahier des charges doit prendre en compte les moindres détails ; mais sans oublier que le transport est le plus souvent sous-traité par le façonnier, et que, quand tout se passe bien du point de vue de la saisie, nous avons peu de moyens de pression.

2.2.3 Qualité des contacts entre le façonnier et l’Insee

Sur ce point aussi, nous retrouvons une unanimité à reconnaître que ces contacts ont été très bons, et on chercherait en vain une ombre au tableau.

C’est le téléphone qui, presque quotidiennement, s’est révélé le vecteur privilégié de cette collaboration.

Mais il y a eu aussi des visites d’ateliers de façonniers par des agents de certaines DR, et quelques fois, en retour des invitations en DR de monitrices du façonnier pour découvrir l’autre face du travail. De tels contacts n’ont pas manqué d’améliorer les relations de travail entre les différents intervenants.

Il ne faut pas cacher que nous avions craint que ne se développent, à cette occasion, des interprétations “régionales” de l’instruction de chiffrement et de saisie. Mais nous avons l’impression que la répartition des responsabilités entre DG et DR a été bien comprise, et que les questions sont arrivées finalement auprès de qui était habilité à y répondre.

De même, on ne constate pas de distorsion particulière dans le traitement par les différentes DR des divergences constatées sur l’échantillon de contrôle (cf article sur le contrôle de qualité de l’exhaustif léger).

2.2.4 Restitution des documents et des fichiers

Comme pour la prise en charge, les opérations se sont, en général, déroulées dans de bonnes conditions, émaillées de quelques incidents rarissimes.
Un reproche se trouve plus souvent cité : le déclassement des colis sur les palettes du retour, qui compliquait la manutention. Il semble que les façonniers aient replacé les cartons dans l’ordre où ils finissaient d’être saisis plutôt que dans l’ordre initial dont ils ne conservaient pas la mémoire. Nous reconnaissons que nous n’avions pas fourni une consigne stricte à ce sujet mais, l’aurions-nous fourni, il n’est pas sûr que nous ayons pu obtenir son application ; aurions-nous oser demander un aller-retour des palettes chez le façonnier pour qu’il retrie les cartons ?

Notons ici que nos emballages ont vaillamment supporté le périple :

Imprimerie Nationale ==> DR ==> Mairie ==> DR ==> Façonnier ==> DR

2.3 Procédure de refus ou d’acceptation des UT

Peu de directions régionales ont émis un avis sur ce point et, quand elles se sont exprimées, ce fut pour regretter que la décision finale leur échappe au profit de la Direction générale (DG). Nous avons relevé : "... les négociations ont échappé à la DR", "... seule la DG avait l’initiative", "... ce n’était pas très responsabilisant pour les cadres de la DR".

Quand le taux d’erreur du façonnier est soit très faible, soit très élevé et que la décision d’accepter ou de refuser ne fait aucun doute, l’attente d’une décision centralisée peut paraître d’une lourdeur excessive. Dans de nombreux cas intermédiaires, où l’on est en droit de s’interroger, il est bon qu’un centre de décision unique assure à la fois, un traitement équitable de tous les sous-traitants et une même qualité moyenne des fichiers régionaux.

2.4. Les services du minitel

2.4.1 La messagerie

Pendant la saisie exhaustive de l’essai de 1988, qui ne concernait pourtant que 10 DR et 5 façonniers, la coordination des transferts et les précisions techniques à communiquer aux participants avaient tenu le responsable "vissé" à son téléphone tout au long de l’opération. Cette façon toute orale de procéder n’avait permis ni d’éviter de délivrer trop rapidement, dans le feu de l’action, quelques réponses regrettables, ni de conserver de ces échanges la moindre trace écrite.

C’est pour remédier à ces inconvénients que nous nous sommes dotés d’une messagerie, qui devait initialement être accessible aux DR et aux façonniers. Pour des raisons techniques, nous avons dû renoncé à l’ouvrir aux façonniers, ce qui est regrettable.
Cet outil a été jugé, dans l'ensemble, fort utile, quoique ayant beaucoup souffert de la comparaison avec la messagerie de collecte qui était d'un usage plus agréable mais tellement plus onéreuse. De notre point de vue, elle a parfaitement rempli son rôle.

2.4.2 La base de données

Le suivi des 140 UT du RP82 dans les allers-retours entre les DR et les 5 façonniers (y compris les nombreuses réfections) a donné naissance à une immense fresque murale dont l'effet esthétique sinon l'efficacité est resté dans les mémoires. La perspective de 250 UT et 8 façonniers laissait présager une œuvre encore plus colossale. C'est pour y échapper que nous nous sommes dotés d'une base de données implantée sur un micro-ordinateur.

Sur son utilité, les directions régionales se partagent en deux groupes d'égale importance. Nous trouvons, d'un côté, celles qui ont intégré cet outil au processus de suivi de leurs UT et l'estiment, par conséquent, très utile.

À l'opposé, il y a celles pour qui ce ne fut qu'un "pensum" supplémentaire imposé par la DG, attendu qu'"une simple fiche cartonnée" permettait de suivre l'avancement de l'UT.

Nous n'arbitrerons pas ce débat, nous contentant de rappeler que, sans cette base de données, pas plus la Division "Recensements de la population" que la Division PCST n'auraient pu suivre le cheminement des 242 UT (plus les réfections).

Une majorité de DR, de l'un ou l'autre camp, regrette le verrouillage des données saisies dans la base qui, en cas d'erreur, imposait de faire appel à la DG pour corriger. Nous leur donnons acte que cette sécurité s'est révélée plus pénible qu'utile.
CONCLUSION

Les deux critères pour juger une opération comme la saisie exhaustive du RP sont le respect des délais et la qualité du fichier obtenu. Le calendrier du point 6.1.6 pour les délais et le tableau des taux d’erreurs du point 6.1.7 montrent que les engagements ont été tenus.

Nous tenons à remercier ici, pour ce succès qui est avant tout le leur, toutes les équipes régionales qui ont su se mobiliser pour mener à bien cette opération.

Il y a tout au long de ce rapport des éléments qui, nous l’espérons, pourront servir pour le prochain recensement. Qu’il nous suffise pour conclure, de rappeler deux points qui nous paraissent particulièrement importants.

Le poste de travail "Population Légale-Exhaustif"

Dans l’avenir, on ne reviendra sans doute pas sur la place de l’exhaustif par rapport aux échantillons lourds. Il est indispensable dans ce cas de rechercher une synergie maximum entre les opérations de détermination de la population légale et celles de la préparation et du suivi des UT de l’exhaustif. Une analyse du poste de travail commun aux deux applications devra être réalisée.

Le choix des façonniers

Il serait vain de se cacher que nous n’aurions pas tenu nos délais si en chemin un second façonnier s’était révélé défaillant. C’est dire l’importance que revêt le choix des intervenants.

Pour une opération de ce type, la qualité principale du sous-traitant nous semble être sa capacité d’organisation, pour laquelle hélas nous ne connaissons pas de test efficace.

Une connaissance de la profession et une vérification des références des candidats sera indispensable.

Il faudra aussi penser à faire figurer au marché une clause permettant d’éliminer facilement un façonnier pour incompétence pendant le premier mois de la prestation.
### ANNEXE 1 : Répartition des bulletins des DR par façonniers

<table>
<thead>
<tr>
<th>Région</th>
<th>Bulletin</th>
<th>Pourcentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSACE</td>
<td>A</td>
<td>100 %</td>
</tr>
<tr>
<td>AQUITAINE</td>
<td>D</td>
<td>93,2 %</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>6,8 %</td>
</tr>
<tr>
<td>AUVERGNE</td>
<td>D</td>
<td>88,5 %</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>11,5 %</td>
</tr>
<tr>
<td>BOURGOGNE</td>
<td>A</td>
<td>52,1 %</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>47,9 %</td>
</tr>
<tr>
<td>BRETAGNE</td>
<td>E</td>
<td>63,3 %</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>31,5 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5,2 %</td>
</tr>
<tr>
<td>CENTRE</td>
<td>B</td>
<td>100 %</td>
</tr>
<tr>
<td>CHAMPAGNE-ARDENNE</td>
<td>E</td>
<td>94,5 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5,5 %</td>
</tr>
<tr>
<td>HAUTE-NORMANDIE</td>
<td>B</td>
<td>90,2 %</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>9,8 %</td>
</tr>
<tr>
<td>ILE-DE-FRANCE</td>
<td>A</td>
<td>58,3 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>20,3 %</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>19,2 %</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2,2 %</td>
</tr>
<tr>
<td>LANGUEDOC-ROUSSILLON</td>
<td>C</td>
<td>100 %</td>
</tr>
<tr>
<td>LIMOUSIN</td>
<td>D</td>
<td>82,8 %</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>11,9 %</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>5,3 %</td>
</tr>
<tr>
<td>LORRAINE</td>
<td>A</td>
<td>92,4 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>7,6 %</td>
</tr>
<tr>
<td>MIDI-PYRENEES</td>
<td>C</td>
<td>83,4 %</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>16,6 %</td>
</tr>
<tr>
<td>NORD-PAS-DE-CALAIS</td>
<td>B</td>
<td>74,4 %</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>25,6 %</td>
</tr>
<tr>
<td>PAYS DE LA LOIRE</td>
<td>H</td>
<td>100 %</td>
</tr>
<tr>
<td>POITOU-CHARENTE</td>
<td>C</td>
<td>68,0 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>32,0 %</td>
</tr>
<tr>
<td>P.A.C.A</td>
<td>C</td>
<td>64,6 %</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>15,3 %</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>8,1 %</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>7,3 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>4,7 %</td>
</tr>
<tr>
<td>RHONE-ALPES</td>
<td>F</td>
<td>93,7 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>6,3 %</td>
</tr>
<tr>
<td>C.N.E.N</td>
<td>H</td>
<td>59,2 %</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>40,8 %</td>
</tr>
</tbody>
</table>

L'exploitation 251
Bibliographie


Note 1437/RP90 du 21 mars 1990 : compléments à l'annexe technique pour les façonniers.

Note 1538/RP90 du 4 mai 1990 : compléments à l'annexe technique pour les façonniers.

Note 1713/RP90 du 24 juillet 1990 : compléments à l'annexe technique pour les façonniers.
Ces notes présentent les instructions de saisie et de chiffrement.

3.4 - CONTRÔLE DE QUALITÉ DE L’EXHAUSTIF LÉGER: MODALITÉS ET RÉSULTATS

Gérard Badéyan et Serge Fraboul

- Cet article présente la problématique, les modalités retenues et les résultats du contrôle de qualité de l’exhaustif léger. La méthodologie de l’échantillonnage est développée dans l’article "échantillonnages pour le contrôle de qualité des phases de saisie-chiffrement du recensement français de 1990 "(J.C.Deville). On se reportera par ailleurs à l’article "la saisie de l’exhaustif" (J.Darmon) pour un bilan de l’exhaustif léger au niveau de sa gestion et en particulier des rapports avec les façonniers sous-traitants.

- Les parties 1. et 3. (problématique et modalités, mesures de qualité sur les différents codes) ont été traitées par G.Badéyan, la partie 2. (résultats du contrôle de réception) est due à S.Fraboul.

1 - Problématique et modalités.

1.1. Finalités et problématique générale

Le système de contrôle de qualité mis en place pour l’exhaustif léger a poursuivi les deux finalités générales suivantes:

- assurer un niveau de qualité du travail de production correspondant à des normes choisies a priori.

- permettre une mesure de l’erreur issue de cette phase du processus de production détaillée par variable.
Ce système s'est inspiré de l'expérience conduite à l'occasion du recensement de 1982, en innovant sur le deuxième point.

La saisie-chiffrement de l'exhaustif léger a été sous-traitée à des façonniers (neuf au total en comptant deux co-traitants)\(^1\). Les documents (84 millions de BI et de FL) ont été préparés en 242 lots (dénommés "unités de traitements") de volume variable par les établissements régionaux de l'INSEE: aux alentours de 150 000 questionnaires (BI et FL) pour les premiers puis compris entre 350 000 et 500 000 par la suite. Commencer par des lots de taille moindre permettait d'approvisionner très tôt les façonniers et de minimiser le coût d'un refus du travail, refus dont le risque était plus fort au démarrage.

Le contrat de sous-traitance spécifiait un niveau de qualité minimal à respecter pour que le travail soit accepté, niveau fixé en fonction des considérations suivantes:

- obtenir un niveau de qualité au moins égal à celui du recensement précédent,
- être conforme aux normes de qualité de saisie habituelles dans les entreprises de saisie informatique.

Se contenter de prendre comme norme le niveau de qualité observé au recensement précédent peut paraître timide comme objectif. Ce n'était pas le cas car la priorité donnée au traitement de l'exhaustif léger qui permettait une publication beaucoup plus rapide des résultats constituait une nouveauté par rapport aux recensements précédents. Or cette priorité conduisait à un défi redoutable en terme de rythme, défi qui fut tenu: la saisie-chiffrement des 84 millions de documents fut achevée mi-janvier 1992, soit en huit mois et demi. D'un point de vue pluridimensionnel (délais et qualité de saisie-chiffrement) on assurait ainsi une amélioration globale importante de qualité.

Le respect des délais imposait une gestion très rigoureuse des UT et en particulier de leur cheminement entre la DR et le façonnier. En particulier il fallait éviter toute rupture de charge et réaliser très rapidement les opérations de réception des fichiers de saisie retournés par les façonniers. La procédure du contrôle de qualité étant la même qu'en 1982 (saisie-chiffrement d'un échantillon de contrôle, appariement avec le fichier du façonnier et analyse des divergences, voir ci-dessous 1.3.: mode opératoire), la recherche d'amélioration a porté sur la diminution du volume de travail des DR à la fois par la diminution de la taille de l'échantillon de contrôle, ce qui réduit en amont le travail de saisie-chiffrement et en aval le nombre de divergences à analyser pour un taux de divergence donné, et par la recherche des conditions pour abaisser le taux de divergences.

\(^1\) Voir sur ce point l'article "la saisie de l'exhaustif".
1.2 Normes et contrat

Le contrat spécifiait la clause de qualité de la façon suivante:

Dans le fichier livré, chaque document (BI, FL) correspondait à un enregistrement. Un enregistrement était réputé erroné dès lors qu’il contenait au moins un caractère erroné (avec comme cas particulier la commune de travail, voir ci-après). Le fichier livré ne devait pas contenir plus de 3% d’enregistrements de type BI ni plus de 1% d’enregistrements du type FL erronés (conditions à remplir simultanément), les pourcentages d’enregistrements erronés étant estimés par sondage.

L’établissement de taux d’erreurs par enregistrements entiers et non code par code représentait une perte de précision acceptée pour simplifier la règle de décision pour les sous-traitants. En effet on ne tenait pas compte de ce fait de la plus ou moins grande difficulté des enregistrements, en particulier de leur longueur, ni d’éventuelles erreurs multiples sur un même enregistrement.

On a contrepéché en partie ce fait (déjà présent au recensement précédent) en fondant la règle de décision sur deux sous-échantillons (sous-échantillon de FL, sous-échantillon de BI): le sous-échantillon FL permettait de tester la qualité de la saisie pure, le sous-échantillon BI la qualité du chiffrement (si la qualité de saisie était bonne, ce qui était attesté par le résultat sur les FL, les erreurs sur les BI allaient alors provenir pour l’essentiel du chiffrement).

Pour les FL, le taux contractuel retenu (1%) visait à imposer une double saisie aux façonniers. En effet en fonction du taux considéré comme habituel d’erreurs de saisie dans la profession (2 caractères erronés sur mille) et du nombre de caractères à saisir en moyenne par enregistrement, une saisie simple conduisait à des taux prévisibles d’enregistrements erronés qui ne pouvaient qu’induire des refus nombreux du travail, même compte-tenu de la marge prise pour limiter le risque de rejet à tort.

Pour les BI, le taux retenu se calait sur celui de 1982 (4%), diminué à 3% pour tenir compte de l’absence en 1990 d’un chiffrement qui avait causé beaucoup d’erreurs en 1982 (celui du type d’activité.)

Le risque de rejet à tort fut pris faible (1/40), ce qui conduisait à des seuils de décision effectifs relativement élevés (2% et 4,5%). Ces seuils n’étant pas contractuels, on se réservait cependant la possibilité de décider au cas par cas pour tenir compte empiriquement de l’information apportée par les résultats constatés sur les unités de traitement précédentes du même façonnier.
1.3 Mode opératoire

1.3.1 Cheminement d’une unité de traitement

Une fois une unité de traitement constituée et avant son envoi chez le sous-traitant, l’établissement régional saisissait un échantillon de contrôle, en laissant les documents vierges de toute annotation, de sorte que le sous-traitant ne puisse pas repérer les documents qui en faisaient partie. Au retour du fichier réalisé par le sous-traitant, l’échantillon de contrôle préparé par l’établissement régional était apparié à la partie correspondante du fichier sous-traitant. Les enregistrements en divergence étaient alors édités et, par examen des documents correspondants, l’établissement régional déterminait si la divergence constatée était due ou non à une erreur du sous-traitant. On déterminait ainsi les pourcentages d’enregistrements erronés sur les sous-échantillons constitués par les enregistrements de type BI et de type FL.

Simple dans le principe un tel système pose dans le concret la question de l’homogénéité de traitement des divergences par les différents établissements. Dans de nombreux cas, en effet, on peut hésiter entre imputer une erreur au façonnier ou estimer que l’information contenue sur le bulletin est trop floue pour appeler un traitement univoque. On verra dans la partie 2 (résultats du contrôle de réception) qu’on est en droit d’estimer que cette difficulté a été résolue de façon satisfaisante.

Pour les BI, en plus du comptage des enregistrements erronés destiné à la prise de décision concernant l’acceptation ou le rejet du fichier concerné, les établissements régionaux comptabilisaient les erreurs par variable afin d’obtenir des mesures de qualité sur celles-ci prises isolément.

1.3.2. Volume des divergences

Comme indiqué plus haut, on a cherché à raccourcir les délais d’envoi et de réception des unités de traitement en diminuant la taille de l’échantillon de contrôle et en recherchant les conditions pour abaisser le taux de divergences.

La réduction de la taille de l’échantillon diminue en amont le travail de saisie-chiffrement et en aval le nombre de divergences à analyser pour un taux de divergence donné. Cette réduction ne se fait pas au détriment de la précision du sondage mais est permise par une amélioration du plan d’échantillonnage dont la méthodologie est présentée dans l’article "échantillonnages pour le contrôle de qualité des phases de saisie-chiffrement du recensement français de 1990" (J.C.Deville).
Les erreurs ne sont pas réparties uniformément sur l’unité de traitement mais sont plus ou moins concentrées sur certains districts; en effet les districts sont plus ou moins difficiles à saisir, chaque district est traité en règle générale par un même opérateur et, comme les opérateurs sont plus ou moins habiles, un district contiendra plus ou moins d’erreurs selon l’opérateur qui l’aura traité.

Il y a donc intérêt à répartir l’échantillon sur le plus grand nombre possible de districts. Cet élargissement trouve ses limites dans le "coût d’ouverture" des districts, c’est-à-dire le temps de manipulation des documents, et, en fin de compte, le nombre de districts concernés (100) résulte d’un compromis entre ces deux contraintes.

Le système fournit, une fois l’unité de traitement constituée, la liste des districts faisant partie de l’échantillon, chacun de ces districts comptant pour 15 BI, avec les FL correspondantes.

En théorie ces 15 BI devraient être choisis aléatoirement dans les districts concernés. Pour des raisons pratiques on les saisit en fait en séquence. Pour éviter des effets systématiques (par exemple meilleure qualité de saisie au début des districts qu’à la fin) ces 15 BI et leurs FL sont saisis alternativement au début, au milieu et à la fin des districts successifs de l’échantillon.

Trois actions concourent à diminuer le volume de divergences à taille d’échantillon fixée.

La plus notable car ayant comme finalité première la qualité finale des fichiers transmis consiste en une participation à la formation préalable des personnels chez les façonniers. Cette action de formation est nouvelle par rapport à 1982. L’examen des résultats du contrôle de réception (cf 2. ci-dessous) prouve de façon indiscutable et très encourageante l’efficacité d’une formation préalable consistante pour une opération particulière de ce type.

Une bonne formation du personnel des façonniers et des directions régionales diminue le nombre d’erreurs et donc celui des divergences à examiner. Mais l’expérience de 1982 et des tests préalables montrait qu’un nombre important de divergences pouvait ne pas être dû à des erreurs de l’un ou de l’autre mais à des différences légitimes d’interprétation du bulletin. En règle générale en effet il est demandé aux façonniers de saisir le maximum de données brutes, même en cas d’incohérences, pour éviter l’hétérogénéité de la saisie d’un atelier à l’autre, le traitement de ces incohérences étant réalisé par programme par la suite. Ce mode de travail contredit la culture professionnelle des agents de l’INSEE basée au contraire sur l’interprétation du document pour en tirer l’information pertinente. On s’est donc attaché à informer les agents des conditions de saisie chez les façonniers et à les inciter à s’y conformer au maximum, en insistant sur le caractère contingent de la consigne de sorte à ne pas créer de confusion par rapport aux habitudes de travail justifiées par ailleurs.

L’exploitation
La non-prise en compte du code de la commune de lieu de travail procède de la même idée: la connaissance du terrain par les agents des DR leur permet dans de nombreux cas de redresser des orthographes erronées ou de substituer le nom véritable d'une commune à une appellation courante, contrairement au façonnier, d'où l'apparition de divergences qui ne seraient pas, à l'analyse, comptées en erreurs pour le façonnier.
2 - Résultats du contrôle de réception

Les résultats du contrôle de qualité de la saisie de l’exhaustif sont bons. 84 millions d’enregistrements (57 millions de BI et 27 millions de FL) répartis en 242 unités de traitement (UT) ont été saisis avec des taux moyens d’enregistrements erronés de 1,5 % pour les BI et de 0,5 % pour les FL, à comparer aux normes fixées à 3 % pour les BI et 1 % pour les FL. Après avoir donné quelques résultats chiffrés, on tentera d’expliquer les différences observées. Les résultats concernent neuf façonniers (deux des sept façonniers contractants avaient un co-traitant) qui ont travaillé avec dix-neuf établissements régionaux de l’INSEE.

2.1. Les résultats du contrôle qualité

Sur les 242 unités de traitement contrôlées, seules 4 ont fait l’objet d’un rejet, 3 pour dépassement du seuil sur les FL (norme à 1 % et seuil à 1,5 %) et une pour dépassement du seuil pour les BI (norme à 3 % et seuil à 4,5 %)(en fait, pour les FL, les taux des UT rejetées étaient légèrement inférieurs à 2 % mais survenaient après plusieurs UT ayant présenté un taux nettement supérieur à 1 %, cf. ci-dessus 1-2). Ces rejets concernent deux façonniers qui en moyenne auront eu les moins bons résultats. Le contrat avec un façonnier a été dénoncé pour qualité insuffisante.

Les résultats qui suivent, concernent les 242 UT qui ont été acceptées, pour les 4 UT rejetées ce sont les résultats du Mme contrôle (après réfection) qui sont pris en compte. Par façonnier on obtient les taux moyens : (en annexe on donne d’autres statistiques) :

<table>
<thead>
<tr>
<th>Façonnier</th>
<th>T BI</th>
<th>T FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0,6  (0,3)</td>
<td>0,2  (0,2)</td>
</tr>
<tr>
<td>B</td>
<td>1,3  (0,7)</td>
<td>0,2  (0,2)</td>
</tr>
<tr>
<td>C</td>
<td>1,3  (0,6)</td>
<td>0,5  (0,4)</td>
</tr>
<tr>
<td>D</td>
<td>1,5  (0,5)</td>
<td>0,5  (0,2)</td>
</tr>
<tr>
<td>E</td>
<td>1,8  (0,6)</td>
<td>0,5  (0,3)</td>
</tr>
<tr>
<td>F</td>
<td>1,9  (0,6)</td>
<td>0,4  (0,3)</td>
</tr>
<tr>
<td>G</td>
<td>2,0  (0,7)</td>
<td>0,5  (0,3)</td>
</tr>
<tr>
<td>H</td>
<td>2,2  (0,7)</td>
<td>0,7  (0,4)</td>
</tr>
<tr>
<td>I</td>
<td>2,4  (0,9)</td>
<td>1,0  (0,6)</td>
</tr>
<tr>
<td>Ensemble</td>
<td>1,5  (0,8)</td>
<td>0,5  (0,4)</td>
</tr>
</tbody>
</table>

Entre parenthèses : l’écart type
14 UT ont un taux de BI erronés supérieur à 3 % et 20 UT ont un taux de FL erronées supérieur à 1 %, 4 UT dépassant les deux normes à la fois. Le tableau suivant donne la répartition par tranche :

<table>
<thead>
<tr>
<th>Taux d'erreur</th>
<th>BI</th>
<th>FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 à 0,5 %</td>
<td>62</td>
<td>9</td>
</tr>
<tr>
<td>0,5 à 1 %</td>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>1 à 2 %</td>
<td>8</td>
<td>46</td>
</tr>
<tr>
<td>2 à 3 %</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>3 à 4 %</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Sur les 21 UT ayant un taux de BI erronés inférieur à 0,5 %, 19 ont été saisis par le façonnier ayant les meilleurs résultats.

2.2. Décomposition des effets

Les taux d'erreur constatés peuvent être le résultat de plusieurs effets : la qualité proprement dite du travail du façonnier, la "qualité" du document (lisibilité, bon remplissage ...), la difficulté intrinsèque du chiffrement, le traitement des divergences par les DR plus ou moins "sévère". Ces effets peuvent être a priori réduits en deux : un effet façonnier et un effet DR. L'effet façonnier concerne la qualité de la codification et de la saisie des documents par le façonnier. L'effet DR regroupe : le traitement des divergences et les aléas de la collecte : qualité de document, plus ou moins grande complexité de chiffrement.

L'homogénéité du traitement des divergences entre les DR était en particulier une question importante que l'on pouvait se poser. En effet les visites de l'équipe centrale en DR ne pouvait pas assurer cette homogénéité à elles seules. Aussi on a cherché si on pouvait porter une appréciation sur cette question a posteriori.

Un grand nombre de DR ont eu des relations presqu'exclusives avec un seul façonnier ; il n'est donc pas possible de tester globalement s'il existe un effet DR sur les taux d'erreurs. Par contre 7 des façonniers ont vu leurs UT analysées par des DR différentes ; si on réalise une analyse de la variance pour chacun d'entre eux, on conclut à chaque fois à l'absence de significativité de l'effet DR (au seuil habituel de 5 %).

Une autre approche consiste à voir que pour une DR donnée, le classement relatif entre façonniers est respecté.
En fait on répartit les façonniers par groupes qui hiérarchisent significativement les taux moyens d’erreur, ainsi les façonniers du groupe 1 ont, de façon significative, des taux d’erreur plus faibles que ceux du groupe 2. On utilise le test de Waller-Duncan sur la comparaison des moyennes, on définit ainsi 4 groupes pour les BI :

- Groupe 1 : A
- Groupe 2 : B,C,D
- Groupe 3 : E,F,G
- Groupe 4 : H,I

pour les FL on a de même 3 groupes :

- Groupe 1 : A,B,F
- Groupe 2 : C,D,E,G,H
- Groupe 3 : I

Le classement entre les 4 groupes est respecté dans toutes les DR sauf une (une interversion entre les groupes 3 et 4). Notons en particulier que le façonnier du groupe n1 (pour les BI) a travaillé avec onze DR et qu’à chaque fois il réalise la meilleure qualité.

Malgré les limites de l’analyse dues au fait que certains façonniers ont eu des relations presque exclusives avec un seul établissement il apparaît que globalement les DR ont eu un comportement homogène dans le traitement des divergences, ni trop sévère, ni trop laxiste. En fait nos visites dans les ateliers donneraient plutôt une impression de sévérité que de laxisme.

En définitive les facteurs explicatifs des différences observées dans le taux d’erreur sont la qualité du travail du façonnier (qualité de la saisie et de la codification) et dans une nettement moindre mesure la complexité non homogène selon les régions et les zones, du chiffrement de certaines variables. Les erreurs de codification s’expliquent par la complexité du BI et par le travail du façonnier, mais un BI complexe sera plus ou moins bien chiffré selon le façonnier et, en définitive, c’est le travail du façonnier qui sera le plus souvent responsable de l’erreur.
2.3. Eléments d’explication des différences de qualité du travail des façonniers

2.3.1 Le type d’organisation

Pour huit façonniers le schéma retenu consistait à séparer la codification de la saisie. Les BI étaient chiffrés par une équipe spécialisée puis étaient saisis sur du matériel classique de saisie de masse. Cette saisie étaient en principe vérifiée intégralement par double saisie.

Un façonnier avait opté pour un autre schéma : codification/saisie assistée sur micro ordinateur avec une vérification partielle de la saisie. Ce façonnier ayant eu 3 UT rejetées (1 pour les BI et 2 pour les FL) et figurant dans le groupe 4 (pour les BI) il apparaît que le premier schéma d’organisation assurait une meilleure qualité.

2.3.2 La formation initiale et la stabilité du personnel

Le recrutement des personnels a été plus ou moins précoce, plus ou moins sévère selon les façonniers et les périodes de formation ont été plus ou moins longues. Les façonniers des groupes (pour les BI) 1 et 2 ont tous assuré une formation avant le début de l’opération. Le façonnier qui a obtenu les meilleurs résultats a formé l’ensemble de son personnel (chiffreurs et dactylocodeurs) pendant un mois, ce qui lui a permis d’obtenir d’emblée de bons résultats (0,1 % sur les FL et 0,8 % sur les BI en mai 1990). En outre, les dactylocodeurs ayant été formés au chiffrement pouvaient lors de la saisie vérifier la codification et demander à l’encadrement de trancher en cas de désaccord. Cette double vérification explique l’excellence de leurs résultats.

La stabilité du personnel a également favorisé une bonne qualité. En effet outre la formation initiale une période d’apprentissage est nécessaire et la lecture des taux mensuels montre bien une amélioration continue de la qualité entre mai et juillet (les taux d’erreur sur les BI passent de 1,8 % à 1,5 %). Remplacer le personnel implique une nouvelle période d’apprentissage : en été certains façonniers y ont été contraints (départs en vacances, départs définitifs). Le taux moyen est passé à 1,7 % mais chez certains le taux a crû fortement (de 1,4 à 1,9 %, de 1,6 à 2,5 %). Le recrutement supplémentaire chez certains façonniers, dû à la défection d’un de leurs confrères, s’est aussi traduit par une dégradation temporaire.
2.3.3 La vérification totale ou partielle de la saisie

La norme fixée à 1 % sur les FL imposait quasiment une double saisie. Un façonnier avait retenu un schéma d’organisation sans vérification totale de la saisie, deux UT ont été refusées et ses taux FL sont nettement moins bons (20 % d’UT ont un taux supérieur à 1 % contre 8 % pour l’ensemble, 29 % ont un taux inférieur à 0,5 % contre 62 %) De même le façonnier qui a le plus mauvais résultat sur les FL faisait bien de la double saisie mais sa deuxième saisie écrasait systématiquement la première ! Après avoir essuyé un rejet et entendu nos remarques, son taux s’est amélioré.

Plus généralement toute tentative de ne vérifier que partiellement la saisie des FL se traduisait par une augmentation sensible du taux d’erreur et par un avertissement avant sanction. Ainsi un façonnier a vu passer son taux de 0,5 % à 0,9 % puis à 1,1 % en fonction de la diminution de la vérification. Pour les BI cette suppression totale ou partielle de la vérification était moins sensible mais elle accroissait sérieusement le risque de rejet pour peu qu’il y ait des erreurs de codification.

Pour la hiérarchisation des façonniers la qualité de la saisie pure n’est pas indifférente. On peut faire l’hypothèse que le taux d’erreur FL mesure la saisie pure et que la différence entre le taux BI et le taux FL mesure l’erreur sur le chiffrement (en première approximation). Quand on fait le test de comparaisons des moyennes, on obtient trois groupes :

- Groupe 1 : A
- Groupe 2 : B, C, D
- Groupe 3 : E, F, G, H, I

en rapprochant avec les 4 groupes sur les BI, on constate que la différence entre le groupe 4 et le groupe 3 est imputable à la vérification totale ou partielle de la saisie.

L’exploitation 263
Pour conclure sur les facteurs explicatifs des différences de qualité du travail des façonniers on peut établir le tableau suivant :

- Groupe 1 : double saisie totale, codification vérifiée, bonne formation initiale, stabilité du personnel
- Groupe 2 : double saisie totale, bonne formation initiale, assez grande stabilité du personnel
- Groupe 3 : double saisie quasi totale, faible formation initiale
- Groupe 4 : vérification partielle de la saisie.

2.4. Autres éléments d’explication des disparités entre DR

On a vu que les DR ont traité de façon homogène les divergences, l’élément essentiel de discrimination est donc la complexité du BI. Chiffrer l’indicateur de nationalité, le statut ou l’indicateur du lieu de travail dans la Creuse ou dans le Val de Marne ne représente pas la même difficulté. Cela se traduit par des disparités entre DR et au sein des DR : les UT concernant les grandes villes ayant en général des taux d’erreur plus élevés.

Le fait que des façonniers ont travaillé pour un nombre restreint de DR a également réduit les taux d’erreurs grâce à une meilleure connaissance de la région et aux échanges DR/façonnier sur le traitement des problèmes particuliers de codification.

La taille réduite des échantillons (800 FL et 1500 BI en moyenne) a permis aussi une baisse des taux grâce à la rapidité de traitement des divergences et donc de la rapidité de l’information fournie en retour au façonnier sur ses difficultés; le façonnier pouvait agir ainsi immédiatement sur l’UT en cours. Notons enfin que la faible taille des échantillons et donc la faible quantité des divergences a sans doute permis une meilleure homogénéité de traitements en diminuant la taille des équipes.
3 - Mesures de qualité sur les différents codes

Le tableau suivant fournit les taux d’erreurs comptabilisés par variables. On constate que ces taux s’échelonnent de 0,06% pour les variables de saisie les plus simples (sexe) à 0,5% pour certaines variables nécessitant un chiffrement (ilt: indicateur du lieu de travail) ou un quasi-chiffrement (cas du statut pour les bulletins renseignés de façon floue). La faiblesse intéressante, de ces taux globaux ne doit pas masquer le fait que les erreurs peuvent être concentrées sur certaines modalités peu fréquentes et donc, pour ces modalités, représenter un taux relativement élevé. Ces taux n’ont pu être établis en raison du traitement manuel des divergences qui a conduit à limiter à la variable dans sa globalité les comptabilisations demandées aux agents chargés du contrôle.

On peut néanmoins conclure que pour les variables de l’exhaustif léger, les erreurs de saisie-chiffrement apparaissent secondaires par rapport aux erreurs de collecte.

Taux d’erreur en pourcentage

<table>
<thead>
<tr>
<th>Varialble</th>
<th>Taux</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexe</td>
<td>0,06</td>
</tr>
<tr>
<td>Etat matrimonial</td>
<td>0,09</td>
</tr>
<tr>
<td>Jour de naissance</td>
<td>0,12</td>
</tr>
<tr>
<td>Mois de naissance</td>
<td>0,12</td>
</tr>
<tr>
<td>Année de naissance</td>
<td>0,11</td>
</tr>
<tr>
<td>Indicateur du lieu de naissance</td>
<td>0,21</td>
</tr>
<tr>
<td>Indicateur de nationalité</td>
<td>0,17</td>
</tr>
<tr>
<td>Indicateur de résidence antérieure</td>
<td>0,16</td>
</tr>
<tr>
<td>Situation des inactifs</td>
<td>0,27</td>
</tr>
<tr>
<td>Recherche d’emploi</td>
<td>0,26</td>
</tr>
<tr>
<td>Aide familial</td>
<td>0,15</td>
</tr>
<tr>
<td>Indicateur de lieu de travail</td>
<td>0,55</td>
</tr>
<tr>
<td>Département du lieu de travail</td>
<td>0,54</td>
</tr>
<tr>
<td>Commune du lieu de travail</td>
<td>0,25</td>
</tr>
<tr>
<td>Statut</td>
<td>0,59</td>
</tr>
<tr>
<td>Situation particulière</td>
<td>0,11</td>
</tr>
<tr>
<td>Position professionnelle</td>
<td>0,27</td>
</tr>
</tbody>
</table>
Annexe: notes concernant la préparation du contrôle de qualité de l’exhaustif léger.

ANNEXE 1 : taux d'erreur BI

<table>
<thead>
<tr>
<th></th>
<th>Moyenne</th>
<th>Ecart-type</th>
<th>Médiane</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0,6</td>
<td>0,3</td>
<td>0,6</td>
<td>0,2</td>
<td>1,4</td>
</tr>
<tr>
<td>B</td>
<td>1,3</td>
<td>0,7</td>
<td>1,2</td>
<td>0,6</td>
<td>2,5</td>
</tr>
<tr>
<td>C</td>
<td>1,3</td>
<td>0,6</td>
<td>1,2</td>
<td>0,5</td>
<td>3,5</td>
</tr>
<tr>
<td>D</td>
<td>1,5</td>
<td>0,5</td>
<td>1,3</td>
<td>0,7</td>
<td>2,9</td>
</tr>
<tr>
<td>E</td>
<td>1,8</td>
<td>0,6</td>
<td>1,7</td>
<td>0,7</td>
<td>3,6</td>
</tr>
<tr>
<td>F</td>
<td>1,9</td>
<td>0,6</td>
<td>1,7</td>
<td>1,1</td>
<td>3,3</td>
</tr>
<tr>
<td>G</td>
<td>2,0</td>
<td>0,7</td>
<td>1,9</td>
<td>0,9</td>
<td>3,8</td>
</tr>
<tr>
<td>H</td>
<td>2,2</td>
<td>0,7</td>
<td>2,0</td>
<td>1,2</td>
<td>3,8</td>
</tr>
<tr>
<td>I</td>
<td>2,4</td>
<td>0,9</td>
<td>2,5</td>
<td>1,3</td>
<td>3,6</td>
</tr>
<tr>
<td>Ensemble</td>
<td>1,5</td>
<td>0,8</td>
<td>1,4</td>
<td>0,2</td>
<td>3,8</td>
</tr>
</tbody>
</table>
ANNEXE 2 : taux d'erreur FL

<table>
<thead>
<tr>
<th></th>
<th>Moyenne</th>
<th>Ecart-type</th>
<th>Médiane</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0,2</td>
<td>0,2</td>
<td>0,2</td>
<td>0,0</td>
<td>0,9</td>
</tr>
<tr>
<td>B</td>
<td>0,2</td>
<td>0,2</td>
<td>0,1</td>
<td>0,0</td>
<td>0,8</td>
</tr>
<tr>
<td>C</td>
<td>0,5</td>
<td>0,4</td>
<td>0,5</td>
<td>0,0</td>
<td>1,4</td>
</tr>
<tr>
<td>D</td>
<td>0,5</td>
<td>0,2</td>
<td>0,4</td>
<td>0,0</td>
<td>1,0</td>
</tr>
<tr>
<td>E</td>
<td>0,5</td>
<td>0,3</td>
<td>0,5</td>
<td>0,0</td>
<td>1,2</td>
</tr>
<tr>
<td>F</td>
<td>0,4</td>
<td>0,3</td>
<td>0,3</td>
<td>0,0</td>
<td>1,1</td>
</tr>
<tr>
<td>G</td>
<td>0,5</td>
<td>0,3</td>
<td>0,5</td>
<td>0,0</td>
<td>1,0</td>
</tr>
<tr>
<td>H</td>
<td>0,7</td>
<td>0,4</td>
<td>0,7</td>
<td>0,0</td>
<td>1,5</td>
</tr>
<tr>
<td>I</td>
<td>1,0</td>
<td>0,6</td>
<td>0,9</td>
<td>0,1</td>
<td>1,9</td>
</tr>
<tr>
<td>Ensemble</td>
<td>0,5</td>
<td>0,4</td>
<td>0,4</td>
<td>0,1</td>
<td>1,9</td>
</tr>
</tbody>
</table>
3.5 - ÉLÉMENTS SUR LE CHIFFREMENT AUTOMATIQUE POUR LE RP DOM

Chantal Madinier et Lionel Viglino

Le recensement de la population de 1990 s’est déroulé dans les départements d’outre-mer de la même manière qu’en métropole. Les questionnaires étaient très proches (même bulletin individuel, aménagement de la feuille de logement). Par contre, l’exploitation du recensement a été effectuée de manière différente : ce sont des méthodes de chiffrement automatique qui ont été utilisées. Cet article présente brièvement le recensement dans les DOM et ses particularités, il décrit la méthode utilisée, en particularité le logiciel Quid, et présente les principaux résultats.

1 - Le recensement de la population de 1990 dans les départements d’outre-mer (DOM)

Pour la première fois, le recensement dans les DOM a eu lieu en même temps que celui de métropole et a utilisé les mêmes concepts. Le bulletin individuel est le même qu’en métropole. La feuille de logement a été adaptée aux spécificités de l’outre-mer : ajout de questions sur l’eau, l’électricité, l’équipement du ménage, suppression des questions sur le chauffage. La collecte a été effectuée par des enquêteurs qui ont donc rempli eux-mêmes les bulletins.

L’organisation de l’exploitation prévue pour la métropole ne pouvait pas s’adapter directement dans les DOM. En effet, ces derniers sont des territoires relativement peu peuplés, en particulier la Guyane avec 115 000 habitants ; une exploitation du 1/4 des bulletins n’aurait pas donné d’informations suffisamment précises. Aussi le principe d’un dépouillement de tous les bulletins a-t-il été retenu. Dans ces conditions, la solution retenue pour la métropole - chiffrement assisté par ordinateur (Colibri) pour le sondage au 1/4 - était difficile à mettre en œuvre :

- son coût en matériel était très élevé : il fallait une machine supplémentaire et environ 50 écrans ;

- son coût en personnel était également très élevé et il ne pouvait pas être supporté par les services de l’INSEE dans les DOM dont la taille est assez réduite.
C’est pourquoi la solution du chiffrement automatique a été retenue malgré son caractère un peu novateur.

Le principe de l’exploitation a été le suivant :

- saisie de toute l’information recueillie : codes et libellés ainsi que quelques chiffrements simples avec, en supplément, le chiffrement des liens familiaux.

- extraction d’un fichier au format de l’exhaustif "léger" de métropole.

Ceci a constitué l’exploitation légère disponible en juin 1991 :

- chiffrement automatique des libellés saisis et donc obtention du code correspondant ;

- en cas d’échec du chiffrement automatique, reprise des bulletins non chiffrés et chiffrement manuel des codes correspondants ;

- constitution d’un fichier complet qui constitue l’exploitation lourde et exhaustive.

Cette dernière a été disponible en juin 1992 ;

Les différents codes concernés par le chiffrement automatique sont les suivants :

- catégorie socioprofessionnelle (CS) ;

- activité économique de l’établissement employeur de la personne (APE) ;

- communes de lieu de travail, de résidence antérieure ;

- nationalité ;

- pays de naissance ;

- pays de résidence antérieure.

Le chiffrement de la catégorie socioprofessionnelle, de la nationalité et du pays de naissance a utilisé le logiciel Quid. Celui des communes a été réalisé avec un logiciel spécifique à la géographie écrit depuis 20 ans et fonctionnant sur d’autres applications à l’INSEE. Cette codification n’est pas présentée dans cet article. Pour l’activité économique, c’est une méthode mixte qui a été utilisée.
2 - Le système quid

2.1. Introduction

LISTE DES ABRÉVIATIONS

QUID abréviation de QUestionnaire d’IDentification désigne aussi bien la méthode générale que l’arborescence de questionnement.

FA : Fichier d’Apprentissage.

CS : Catégorie socioprofessionnelle.
Nomenclature à 42 postes.

RPDOM : Recensement dans les départements d’outre-mer.

BI : Bulletin individuel (du recensement).


Rappel du problème

Le problème consiste à classer automatiquement un individu enquêté dans un poste défini d’une nomenclature existante (par exemple, la nomenclature des Professions). Pour cela, le système utilise principalement la réponse en clair à la question posée directement (par exemple "Quelle profession ou quel métier exercez-vous actuellement?") et accessoirement d’autres informations figurant dans le formulaire d’enquête et supposées préalablement codifiées (par exemple, le code activité économique de l’entreprise employant l’individu).

Dans notre terminologie, la réponse directe en clair est appelée "intitulé". Les informations codifiées complémentaires sont désignées sous le terme générique de "variables annexes".

L’exploitation
2.2. Le principe de la méthode

2.2.1 L’approche de base

L’approche de base du système QUID consiste à élaborer une base de données très importante constituée d’intitulés typiques des répondants, accompagnés du code correspondant attribué par un expert.

Dans notre terminologie, la base de données s’appelle "base d’apprentissage" ou fichier d’apprentissage (FA).

À partir du fichier d’apprentissage, le système constitue une arborescence de questionnement en interrogeant les libellés pour identifier un code.

Les libellés sont tout d’abord normalisés dans des zones de caractères fixes partitionnées en zones fixes de deux caractères appelées "bigrammes".

Les bigrammes sont les questions posées par l’arborescence QUID pour identifier un code.

QUID utilise la théorie de l’information pour construire une arborescence de longueur moyenne minimale afin de réduire au maximum le temps d’identification d’un code.

Par la suite, nous utiliserons comme vocabulaire indifféremment questionnaire d’identification ou arborescence (de questionnement).

2.2.2 Constitution d’un fichier d’apprentissage

Pour construire le fichier d’apprentissage (FA), nous partons le plus souvent d’une enquête d’une année antérieure, déjà chiffrée manuellement ou par une méthode interactive. Chaque intitulé de la base est accompagné de son code (supposé a priori exact), et de sa "fréquence d’occurrence" observée dans le FA, c’est-à-dire du nombre d’individus ayant répondu par cet intitulé.

La constitution d’un fichier d’apprentissage suffisamment complet et juste est une opération longue et coûteuse, mais ce travail est un investissement durable.
2.2.3 La normalisation des libellés

Les intitulés subissent un traitement automatique de normalisation commandé par un jeu de paramètres externes choisis par l'utilisateur.

Les mots sont séparés et cadrés dans des zones fixes dont la longueur (unique pour tous les mots) et le nombre maximum (unique pour tous les intitulés) sont paramétrés.

Les "mots vides", considérés comme sans importance par l'utilisateur, sont éliminés. Ces mots éliminés sont souvent des articles ou des prépositions mais dépendent du langage traité.

L'utilisateur paramètre les séparateurs des mots, la liste des mots vides, le nombre des mots gardés et le nombre de caractères conservés par mot.

2.2.4 Construction de l'arborescence optimisée

À chaque sommet du questionnaire, le système détermine le bigramme (i.e. la question) qui maximise la réduction d'incertitude sur le code. L'algorithme mis en œuvre par QUID est fondé sur la théorie de l'information et utilise l'entropie de Shannon.

Notons : T = (T_1, T_2, ..., T_j, ..., T_n) le code et l'ensemble des modalités du code.

Q = (q_1, q_2, ..., q_i, ..., q_m) l'ensemble des bigrammes résultant de la normalisation des intitulés

X = l'arborescence à construire

À un sommet quelconque X_o, posons :

N (X_o) la fréquence d'occurrence totale des intitulés associés au sommet X_o.

N (X_o, j) la fréquence d'occurrence du code T_j parmi les intitulés associés au sommet X_o

Hypothèse :

En supposant que la population d'apprentissage soit statistiquement représentative de la population à chiffrer, on peut estimer la probabilité de trouver le code T_j dans la population à chiffrer par :

\[ P (T_j/X_0) = \frac{N (X_o, j)}{N (X_o)} \]
Choix du bigramme :

Au sommet $X_0$, l'incertitude sur le code est mesurée par l'entropie de Shannon $H(T/X_0)$

$$H(T/X_0) = \sum_j P(T_j/X_0) \times \log(1/P(T_j/X_0))$$

Supposons qu'un bigramme (quelconque) $qi$ soit affecté au sommet $X_0$. A chacune des modalités qu'il prend, nous associons la sous-base constituée des intitulés possédant cette modalité.

Notons $(a_{i1}, a_{i2},...,a_{ik},...) le s les modalités prises par le bigramme $qi$. Pour chacune de ces modalités, donc pour chacune des sous-bases engendrées, nous créons un sommet $Y$, successeur immédiat de $X_0$.

L'information apportée par le bigramme $qi$ est mesurée par la réduction moyenne de l'incertitude sur $T$ en passant de $X_0$ à l'un des sommets $Y$.

Soit $I(T,qi,X_0) = H(T/X_0) - \sum_Y P(Y) \times H(T/Y) * H(T/X_0) * H(T/Y)$ avec $Y \in \Gamma(X_0)$

(Entropie conditionnelle au sommet $X_0$)

avec $\Gamma(X_0)$ l'ensemble des sommets $Y$, successeurs du sommet $X_0$, définis par le choix de $qi$.

$$P(Y) = N(X_0,a_{ik}) / N(X_0)$$

$N(X_0,a_{ik})$ est la fréquence d'occurrence de la modalité $a_{ik}$ du bigramme $qi$ parmi les intitulés associés au sommet $X_0$.

L'algorithme effectue ce calcul d'information pour tous les bigrammes $qi$ affectés au sommet $X_0$.

L'algorithme choisit le bigramme qui maximise $I(X_0,T,qi)$ et partage alors effectivement la base en autant de sous-bases que de modalités du bigramme $qi$ rencontrées dans la base. Les sommets $Y$, successeurs de $X_0$, sont alors créés.

Pour chaque sous-base obtenue (donc pour chaque sommet $Y$) l'algorithme recommence exactement le même traitement que celui que nous venons de décrire à propos du sommet $X_0$. 

274

INSEE Méthodes n° 52 - 53
Un sommet est terminal quand l'entropie conditionnelle de tous les bigrammes des libellés associés à ce sommet est nulle.

Il existe deux types de sommets terminaux :

- les sommets de décision

Un sommet de décision est un sommet terminal pour lequel les libellés associés possèdent tous le même code.

- Les sommets d'indécision

Un sommet d'indécision est un sommet terminal associé à des libellés possédant des codes différents qu'aucun nouveau bigramme ne permet de discriminer.

2.2.5 Chiffrement, contrôle de redondance

Le questionnaire d'identification est utilisé pour déterminer le code associé à un libellé à chiffrer.

Le libellé à chiffrer est normalisé avec les mêmes paramètres que ceux utilisés pour construire le questionnaire d'identification.

Ce libellé est découpé en bigrammes et l'arborescence interroge les bigrammes du libellé pour déterminer un code.

Le contrôle de redondance est effectué lors du chiffrement au niveau d'un sommet terminal ; ce contrôle consiste à vérifier la concordance du contenu de quelques bigrammes qui n'ont pas été interrogés par le questionnaire avec celui des libellés associés au sommet terminal pendant la phase de construction de l'arborescence.

La liste des bigrammes à vérifier est un paramètre à la disposition de l'utilisateur de QUID.

Au moment du chiffrement quatre issues sont possibles :

- un écho unique non douteux.
  L'interrogation des bigrammes du libellé à chiffrer a conduit à un sommet de décision du questionnaire (écho unique) et le contrôle de redondance est satisfaisant.
- un écho unique douteux
  Le parcours de l’arborescence est arrivé à un sommet de décision mais le contrôle de redondance signale au moins une différence parmi les bigrammes contrôlés entre le libellé à chiffrer et tous les libellés associé au sommet de décision.

- un écho multiple
  Le libellé à chiffrer est associé à un sommet d’indécision.

- une réponse inconnue
  Le contenu d’un bigramme interrogé par l’arborescence n’est pas connu.

2.2.6 Exemple

Afin d’illustrer le mécanisme de QUID considérons un cas d’école de chiffrage de la catégorie professionnelle :
Exemple pédagogique QUID

Afin d’illustrer le mécanisme de QUID, considérons un cas d’école de chiffrement de la catégorie professionnelle

1ère phase : apprentissage

FICHIER D’APPRENTISSAGE BRUT

64 CHAUFFEUR DE TAXI
64 CHAUFFEUR DE TAXI
31 DOCTEUR
31 DOCTEUR EN MéDECINE
52 FACTEUR
54 EMPLOYÉ
52 EMPLOYÉ
34 CHIRURGIEN
31 CHIRURGIEN DENTISTE

PARAMÈTRES DE NORMALISATION

SÉPARATEURS DES MOTS : " "
MOTS VIDES : "EN", "DE"
NOMBRE DE MOTS CONSERVÉS : 2
NOMBRE DE CARACTÈRES CONSERVÉS PAR MOT : 10
## Fichier d'apprentissage normalisé:

<table>
<thead>
<tr>
<th>CODE</th>
<th>LIBELLÉ</th>
<th>FRÉQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>CHAUFFEUR TAXI</td>
<td>2</td>
</tr>
<tr>
<td>31</td>
<td>DOCTEUR</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>DOCTEUR MÉDECINE</td>
<td>1</td>
</tr>
<tr>
<td>52</td>
<td>FACTEUR</td>
<td>1</td>
</tr>
<tr>
<td>54</td>
<td>EMPLOYÉ</td>
<td>1</td>
</tr>
<tr>
<td>52</td>
<td>EMPLOYÉ</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>CHIRURGIEN</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>IDENTISTE</td>
<td>1</td>
</tr>
</tbody>
</table>

Bigramme 12: 1 2 3 4 5 6 7 8 9 10

## Questionnaire d'identification:

**Question:**

Bigramme 2

- 2 = "AU"
- 2 = "CT"
- 2 = "PL"
- 2 = "IR"

**Code:**

- CODE 64: chauffeur taxi
- CODE 31: docteur médecine
- CODE 52: facteur
- CODE 34: chirurgien

**Codes:**

- CODES 52, 54: employé
- CODE 31: chirurgien dentiste

**Bigrammes:**

- Bigramme 1: 1 = "DO"
- Bigramme 6: 6 = "DE"
PENDANT LA CONSTRUCTION DE L'ARBORESCENCE

2 = "CT" EST UN SOMMET DE QUESTIONNEMENT
LE PROCESSUS S'ARRÊTE QUAND L'ENTROPIE CONDITIONNELLE EST NULLE.
2 = "CT" ; 1 = "DO" EST UN SOMMET DE DÉCISION
2 = "PL" EST UN SOMMET D'INDÉCISION

2e phase : le chiffrement

AU MOMENT DU CHIFFREMENT

"CHAUFFEUR DE TAXI" DONNE UN ÉCHO UNIQUE NON DOUTEUX  Le caractère douteux dépend du contrôle de redondance
"CHAUFFEUR ROUTIER" DONNE UN ÉCHO UNIQUE DOUTEUX
"EMPLOYÉ" CONDUIT À UN ÉCHO MULTIPLE 2 codes sont proposés ; 52 et 54
"AGRICULTEUR" CONDUIT À UN INCONNU La réponse 2=R1 n'est pas connue du questionnaire

LES SUCCÈS DU CHIFFREMENT AUTOMATIQUE SONT LES ÉCHOS UNIQUES NON DOUTEUX.
2.2.7 Mise en œuvre du système QUID

Le logiciel QUID est une bibliothèque de programmes généraux. La mise en œuvre du système sépare deux phases : la phase d’APPRENTISSAGE de celle du CHIFFREMENT proprement dit.


2.2.7.1 Apprentissage expertise

2.2.7.2 Chiffrement
QUID est une méthode générale qui peut être utilisée sur des nomenclatures différentes avec des fichiers d’apprentissage et des paramètres de normalisation appropriés.

QUID est un système expert car :

- d’une part, il reproduit un chiffrement d’expertise

- d’autre part, les connaissances (FA) sont séparées des programmes de mise en œuvre (logiciel QUID).

2.2.8 La justesse du chiffrement

2.2.8.1 L’expertise du fichier d’apprentissage

À sa création, le fichier d’apprentissage n’a que la justesse de la méthode utilisée pour coder les exemples qu’il contient. Mais à chaque apurement, les améliorations apportées au FA restent pérennes et la justesse de la base d’exemples augmente progressivement jusqu’à une qualité d’expertise.

La justesse du chiffrement risque d’être dégradée par la faculté de QUID d’induire une codification à la seule vue des bigrammes interrogés par l’arborescence. Cette faculté est bridée par le contrôle de redondance. Cependant l’élargissement du contrôle de redondance diminue la proportion d’échos uniques non douteux en augmentant le taux de justesse de ceux-ci.

Le paramétrage du contrôle de redondance doit donc tenir compte de l’objectif visé : soit accepter une qualité moindre en diminuant les rejets et en conséquence les coûts de traitement, ou au contraire augmenter la précision, au risque d’élèver le taux de rejets.

2.2.8.2 Les mesures de justesse

Pour le moment, peu de mesures du taux de justesse ont été faites et l’on en reste à l’impression très favorable des responsables d’enquête sur la qualité obtenue par ce mode de chiffrement.

Deux tests effectués, l’un en 1985, l’autre en 1990, ont donné des taux d’erreur de 3 à 4 fois moindres que ceux obtenus avec un chiffrement manuel de qualité courante.

L’exploitation
2.2.9 L’utilisation de variables annexes

Pour un certain nombre de nomenclatures, le chiffrement de libellés peut être précisé par des variables codées appelées variables annexes, obtenues à partir de questions fermées dans le questionnaire d’enquête.

Par exemple, pour chiffrer la catégorie socioprofessionnelle (CS) à partir du bulletin du recensement, une dizaine de variables annexes doivent être considérées (activité, statut salarié ou non, taille de l’entreprise, qualification, fonction, secteur public ou privé, spécialité de l’exploitation agricole, etc.). Le croisement des modalités des variables présentées ci-dessus produit environ 91 000 possibilités.

Pour traiter ces variables annexes, la solution qui consiste à les concaténer à la fin du libellé et à les structurer en bigrammes n’est pas satisfaisante dès que les modalités de ces variables sont nombreuses.

Dans le cas (réaliste) où plusieurs variables annexes apportent la même quantité maximale d’information sur le code, la convention adoptée par QUID est d’interroger la première dans l’ordre de déclaration. Le problème qui se pose, quand les variables annexes n’ont pas toutes la même utilité sur tout le champ de la nomenclature est qu’il n’est pas possible de trouver a priori un bon ordre de déclaration.

QUID risque alors d’identifier un code avec une variable annexe sans signification pour la nomenclature.

Ce problème est lié à l’incomplétude du fichier d’apprentissage vis-à-vis des variables annexes ; il pourrait être résolu en multipliant le fichier d’apprentissage pour chaque appellation de profession par l’ensemble des combinaisons des variables annexes. Mais le nombre trop grand de modalités de l’espace croisé des variables annexes rend cette solution physiquement impossible.

Pour ces raisons, dès que le chiffrement d’un code doit tenir compte de plusieurs variables annexes, une solution à deux niveaux a été retenue :

- le premier niveau est une identification par QUID d’un code qui est soit le code définitif, soit un pointeur appelant une table pour lever une ambiguïté sur la codification.

- le deuxième niveau est un ensemble de tables de décision pour traiter les variables annexes et achever le chiffrement.

Cette solution à deux niveaux s’est imposée pour le traitement de la CS du RPDOM qui utilise 90 tables indépendantes.
2.2.10 Des économies d’échelle

Des économies d’échelle sont déjà réalisées grâce à l’utilisation sur plusieurs enquêtes de fichiers d’apprentissage peu différents.

Il est rentable d’investir dans la construction d’un fichier d’apprentissage, pour exploiter une enquête périodique, ou pour chiffrer un code utilisé par plusieurs enquêtes qui respectent une normalisation minimale de leur questionnaire.

Il est aussi possible de réutiliser un fichier d’apprentissage en sachant a priori que le langage utilisé par les répondants à l’enquête sera un peu différent de celui des libellés du fichier d’apprentissage.

Par exemple, le chiffrement de la catégorie socioprofessionnelle du recensement de la population dans les départements d’outre-mer utilise un fichier d’apprentissage provenant d’une enquête en métropole dont les questionnaires ont été remplis par des enquêteurs de l’INSEE alors que les bulletins individuels du recensement dans les DOM sont remplis par les agents recenseurs, personnels recrutés uniquement pour cette opération et donc moins qualifiés.

Dans ce cas, il est utile d’enrichir le fichier d’apprentissage juste après le début de l’exploitation par les libellés les plus fréquemment non codés (cf. 7.2.1)
3 - Bilan général des procédures de chiffrement pour le RP DOM

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Renseignements traités</th>
<th>Taux de codification</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>Libellé de profession</td>
<td>77,9 %</td>
</tr>
<tr>
<td></td>
<td>Libellé de grade</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variables annexes</td>
<td></td>
</tr>
<tr>
<td>CSANTÉRIEURE</td>
<td>Libellé de profession</td>
<td>89,4 %</td>
</tr>
<tr>
<td></td>
<td>Libellé de commune</td>
<td></td>
</tr>
<tr>
<td>CODE COMMUNE</td>
<td>Code postal</td>
<td>98,3 %</td>
</tr>
<tr>
<td>CODE APE</td>
<td>Libellé d’activité</td>
<td>68,7 %</td>
</tr>
<tr>
<td>Activité de l’établissement</td>
<td>Libellé de raison sociale</td>
<td></td>
</tr>
<tr>
<td>CODE NATIONALITÉ</td>
<td>Libellé de nationalité ou</td>
<td>98,3 %</td>
</tr>
<tr>
<td>CODE PAYS</td>
<td>Libellé de pays</td>
<td></td>
</tr>
</tbody>
</table>

Le recensement de la population dans les départements d’outre-mer traite 1 460 000 bulletins individuels. Les bulletins sont saisis par des façonniers, ensuite 85 % des bulletins sont entièrement traités automatiquement par les différents algorithmes mis en œuvre.

Dans le tableau, les pourcentages sont établis par rapport au nombre de bulletins concernés rubrique par rubrique, ce qui explique que certains taux puissent être inférieurs au pourcentage total (85 %) de bulletins traités entièrement.

Le chiffrement de la CS met en œuvre une arborescence de 28 600 sommets dont 22 630 sommets de décision.

Le chiffrement d’un code CS nécessite en moyenne l’interrogation de 2,5 bigrammes et un temps de traitement de 0,37 ms CPU sur IBM 390-200-E (31 MIPS).

4 - Le traitement de la CS du RPDOM

La prise en compte, en plus du libellé de profession, du libellé de grade et des variables annexes présentes dans le bulletin individuel (BI) a nécessité de compléter la méthode générale de chiffrement par des programmes aidant QUID à trouver et à traiter les informations nécessaires pour aboutir à un chiffrement.

Le libellé découpé en bigrammes et traité par QUID résulte d’un traitement préalable qui concatène au libellé de profession le libellé de grade et une variable de synthèse.
Cette variable de synthèse, à 10 modalités, est calculée à partir des variables annexes du bulletin et représente les grandes divisions en espaces sociaux de la nomenclature des CS.

Parfois la pauvreté des renseignements collectés sur le BI ne permet pas d’établir cette variable de synthèse. Deux QUID sont donc utilisés :

- Le premier appelé QUID-A traite les libellés pour lesquels une variable de synthèse a pu être calculée.

- Le deuxième QUID-B est chargé des libellés sans variable de synthèse.

4.1. **Quid-A**

QUID-A donne une réponse qui est soit le code CS définitif, soit un pointeur appelant une table pour achever la codification. En effet, dans certains cas, le libellé n’est pas assez complet pour déterminer précisément un code, il faut alors interroger des variables annexes du BI pour lever l’ambiguïté sur le code.


Les libellés de professions sont normalisés à 5 mots de 10 caractères, soit 25 bigrammes ; la variable de synthèse est ajoutée en 26ème bigramme. Une étude menée en 1989 a permis de réduire, sans perte de qualité du chiffrage, le contrôle de redondance à la variable de synthèse et aux 3 premiers bigrammes des deux premiers mots. En raison de cette réduction du contrôle de redondance, les échos douteux qui restent sont faux dans 2 cas sur 3. Les échos uniques douteux ne sont donc pas retenus et seuls les échos uniques non douteux sont automatiquement acceptés pour le chiffrage.

**Les tables de QUID-A**

Les tables associées à QUID-A ont pour rôle de lever une ambiguïté possible sur le code définitif. Ces tables sont mises au point au moment de l’expertise du fichier d’apprentissage avant la phase de chiffrage. Les tables n’ont pas été construites systématiquement à partir de la nomenclature mais ont été créées en examinant les sommets d’indécision de l’arborescence et sont rattachées à ces sommets. Ces sommets d’indécision provenaient de libellés pauvres en information conduisant à deux codes possibles lors de la constitution du fichier d’apprentissage.
Cette méthode de construction a l’avantage de faire construire en premier les tables les plus fréquemment utilisées. Elle permet d’obtenir le meilleur rapport qualité/coût pour la constitution des tables.

Il a par ailleurs été décidé de construire des tables indépendantes les unes des autres. D’une façon plus imagée, les tables ne s’appellent pas entre elles au moment de leur exécution. Cette décision a pour but de faciliter la maintenance du fichier. En effet, une modification dans une table reste limitée à cette table et n’a pas de répercussions en cascade sur les spécifications de traitement des autres libellés (les modalités des variables annexes, en particulier les modalités croisées, peuvent en effet conduire à des choix différents dans la nomenclature, selon les professions) ; le coût de maintenance du FA en est ainsi réduit.

Au total, 90 tables ont ainsi été implantées. Elles constituent le prolongement logique de l’arborescence de questionnement du QUID ; mais si QUID-A choisit les bigrammes à interroger dans le libellé pour optimiser la recherche, la solution retenue ici laisse aux experts humains le choix de déterminer le questionnement des variables annexes.

4.2. **Quid-B**

Le second QUID, appelé QUID-B s’occupe des cas où aucune modalité de la variable de synthèse n’a pu être déterminée. Le chiffrage de QUID-B tient compte de l’absence d’information dans le BI qui permettrait de calculer une variable de synthèse. Le code attribué au libellé correspond à celui prévu par la nomenclature dans ce cas. Nous avons vérifié cette option par un test qui a montré que l’application de règles par défaut apporte des résultats de meilleure qualité que le chiffrement manuel à partir du document papier (9 % contre 17 % d’erreur sur cette population).

Le fichier d’apprentissage de QUID-B a été déduit de celui de QUID-A en supprimant la variable de synthèse et en faisant lever les indécisions ainsi créées par des spécialistes du chiffrement de la CS.

Les mêmes paramètres de normalisation du libellé et de contrôle de redondance ont été adoptés.

Là encore, seuls les échos uniques non douteux sont acceptés pour le chiffrement automatique.

Aucune table n’a été implantée pour QUID-B ; les règles par défaut de la nomenclature qui affectent le code le plus probable ont été systématiquement appliquées.
4.3. L'enrichissement des fichiers d’apprentissage en cours d’exploitation

Les possibilités d’apprentissage du système QUID ont été utilisées après le début de l’exploitation du recensement pour enrichir les fichiers d’apprentissage de QUID-A et de QUID-B.

4.4. La reprise des échecs des codifications automatiques

Les rejets de la codification automatique par QUID-A et QUID-B sont traités manuellement avec les rejets des autres procédures automatiques, sur micro-ordinateur, dans chaque département d’origine.

Le traitement de ces rejets est effectué grâce à une application spécialement conçue pour le RPDOM.

Les rejets concernent 22 % des 403 000 actifs pour la CS, et au total 15 % des 1 460 000 BI collectés pour toutes les codifications automatiques.

5 - Le chiffrement de l’activité économique

On chiffre l’activité principale exercée (APE) à partir des informations suivantes :

- raison sociale de l’établissement employeur
- adresse de cet établissement
- activité déclarée sur le bulletin individuel

Le principe est le suivant : à partir des libellés raison sociale et adresse, on recherche l’établissement dans le fichier général des établissements (SIRENE). En cas de succès, on obtient tous les codes utilisés par la suite et relatifs à cet établissement (APE, code public/privé appelé PUB, taille de l’établissement...). En cas d’échec de la recherche dans SIRENE, on utilise le libellé d’activité déclaré.

En pratique, il a été fait appel à trois techniques différentes dont deux très similaires (QUID).
5.1. Mise en Concordance Automatique (MCA)

La mise en concordance automatique est une technique développée par le Centre national d'exploitation de l'INSEE de Nantes qui permet de mettre en relation un enregistrement en entrée avec le fichier SIRENE. Cette technique sera utilisée de façon prioritaire pour déterminer le code APE et le code PUB des établissements des actifs.

Pour la MCA, les informations nécessaires en entrée sont la raison sociale et l'adresse la plus précise possible de l'établissement. La recherche pourra porter indifféremment sur la zone "nom" ou "raison sociale" ainsi que les zones "sigle" ou "enseigne". Pour qu'une recherche aboutisse, il est donc impératif que le libellé figurant sur le bulletin ait une forme approchante d'une des trois zones figurant dans SIRENE. De plus, la zone adresse a une importance considérable, car les enregistrements ayant des échos avec une adresse très divergente de celle figurant sur le bulletin seront rejetés (n'oublions pas que pour SIRENE les quatre DOM forment un même département ayant pour code 97).

De la MCA seront extraits les échos uniques ayant reçu un "code qualité" convenable, et les échos multiples à condition qu’ils aient la même APE, le même code PUB et la même commune.

5.2. Quid Raison sociale

Les échos non traités par SIRENE sont passés dans un QUID portant sur la zone "Raison sociale", ce QUID étant rigoureusement le même que celui portant sur l’APE. Cette décision a été prise pour deux raisons :

- de nombreux enregistrements présentent en zone "Raison sociale" des indications devant sans conteste figurer sur la zone "Activité principale exercée" ou dans la zone "Profession". C’est le cas fréquent d’entrepreneurs individuels ne reportant pas leur nom en zone raison sociale, mais reportant leur activité (artisan, maçon, restaurant, etc.).

- de nombreux enregistrements ne trouvent pas d’échos dans SIRENE alors qu’il est possible de les coder sans difficultés. On retrouve souvent des employeurs pouvant avoir plusieurs noms, et SIRENE n’en connaît qu’un. C’est ainsi que les libellés "commune de...", "ville de...", "municipalité de...", "mairie de...", "école de...", "collège de...", "lycée de...", "perception de...", "poste de..." ont été systématiquement ajoutés dans le QUID pour les 113 communes des DOM. De même, les libellés "conseil général...", "conseil régional...", "région...", "département...", "préfec-
...ture..." ont été eux aussi inclus dans le QUID. D'autres libellés ont été ajoutés, relatifs à des noms usuels mais non officiels de grosses entreprises privées ou publiques ne figurant pas ci-dessus avec les codes APE correspondant.

5.3. Quid APE

Le QUID APE, c'est-à-dire le passage en interrogant la zone "activité principale exercée" est le QUID "de la dernière chance". En effet, après son passage, il n'y a plus aucun moyen automatique de coder l'APE. Il s'effectue évidemment après la MCA et après le passage du QUID raison sociale. Si le QUID raison sociale est privilégié par rapport au QUID APE, c'est parce que l'on sait, par expérience, que la zone "activité principale exercée" n'est pas toujours convenablement renseignée - l'enquêté ayant tendance à donner son activité propre au lieu de celle de l'établissement. Cependant, c'est néanmoins ce quid APE qui, dans l'état actuel des choses, a permis de coder le plus d'enregistrements. Il a été constitué principalement à partir de COLIBRI (utilisation des libellés standards de COLIBRI, enrichis par les synonymes reconnus du système), ainsi que d'un fichier préexistant comprenant environ 30 000 actifs avec les codes APE et les libellés correspondants.

5.4. Codification du Code PUB

Le code PUB (code à 5 modalités : agents de l'État, agents des collectivités locales, agents de la sécurité sociale, agents des grandes entreprises publiques et agents du secteur privé) se détermine de trois façons :

- dans le cas de la codification par MCA, le code PUB découle de la forme juridique détaillée fournie par SIRENE

- dans le cas de la codification du code APE par QUID, la forme juridique détaillée de l'établissement est évidemment inconnue. Le code PUB se calcule alors par une table à partir du code APE. Ceci explique en partie pourquoi, alors que la codification de l'APE ne s'effectuera qu'en NAP 100, il est utile de conserver des positions spécifiques de la NAP 600 car c'est à ce niveau qu'il existe une liaison fonctionnelle entre la NAP et le code PUB.

- Si l'APE n'a pas été calculée, il est possible d'affecter la valeur 5 au code PUB dans le cas où l'individu s'est déclaré, à la question 16, comme étant à son compte.
6 - Le chiffrement des nationalités des pays de naissance ou de résidence antérieure

Les 2 codifications de pays (pays de résidence antérieure et pays de naissance) et celle de la nationalité sont traitées par un QUID unique. Ce QUID est relativement simple, du fait que le nombre de pays est restreint (environ 360 pays reconnus comme tels par l'ONU). Les taux de codification, pour les enregistrements convenablement renseignés, dépassent 95 %.

Pour les Antilles, un problème a été difficilement réglé : il concerne les nationalités, entre la République Dominicaine, dont les ressortissants sont des Dominicains, et la Dominique, dont les ressortissants sont des Dominiquains. Il a été décidé de demander à la collecte de noter les ressortissants de la Dominique, Dominiquais. Cette consigne n'a été que partiellement suivie et il a fallu en dernier ressort utiliser le pays de naissance et celui de résidence antérieure pour affecter correctement Dominicains et Dominiquains.

7 - Bilan de la codification pour l'APE, les nationalités, pays et la CS

7.1. La première codification par QUID

Elle a été réalisée sur environ 15 % des bulletins. Les taux de bulletins non codés étaient les suivants

<table>
<thead>
<tr>
<th>APE</th>
<th>26,7 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS</td>
<td>32,4 %</td>
</tr>
<tr>
<td>Pays de résidence antérieure</td>
<td>9,6 %</td>
</tr>
<tr>
<td>Nationalités</td>
<td>40,1 %</td>
</tr>
<tr>
<td>Pays de naissance</td>
<td>15,4 %</td>
</tr>
</tbody>
</table>

A ce stade, les QUID ont été améliorés en particulier celui des nationalités où ITALIEN avait été omis.
7.2. La campagne d’amélioration

7.2.1 PCS

Les rejets des premiers lots exploités ont été triés par fréquence décroissante ; les 3 % des libellés les plus fréquents représentant 28 % des cas de rejet ont été expertisés puis inclus dans les fichiers d’apprentissage.

Cet enrichissement a permis d’augmenter le taux de chiffrement de 68 % à 78 % sans interrompre significativement l’exploitation.

Environ 400 libellés supplémentaires ont ainsi été chiffrés. Le libellé non chiffré le plus fréquent était "gens de maison".

Les libellés non fréquents qui n’ont pas été codés sont de 3 types :

- métier peu fréquent
- métier fréquent avec appellation non fréquente
- libellés illisibles, sans signification, trop vagues (ce qui pose problème car on rappelle que pour la PCS à 2 positions il n’existe pas de code rebut).

Les 2 premiers types de cas peuvent être traités (cela dépend du temps dont on dispose), le 3e est à voir avec les nomenclaturistes.

7.2.2 APE

Le travail a été long et difficile car le fichier d’apprentissage initial contenait environ 1/4 d’APE fausses ; en effet, le fichier utilisé initialement a été présenté à une MCA SIRENE : on avait ainsi obtenu une APE juste pour l’enregistrement mais qui pouvait être fausse pour le libellé de l’activité. Le fichier d’apprentissage contient environ 40 000 libellés. Tous les libellés apparaissant 3 fois et plus et pouvant être codés l’ont été.

Le problème du chiffrement est très difficile car les personnes interrogées n’ont pas l’habitude de répondre à cette question qui est souvent mal comprise : confusion entre l’activité personelle et celle de l’établissement.

Cette amélioration a été insuffisante et il a fallu reprendre certaines codifications ultérieurement (mais toujours de manière automatique).
Le taux de succès final (70 %) est très satisfaisant.

7.2.3 Nationalité

Les résultats sont excellents. Néanmoins un enseignement peut être tiré de ce chiffrement. Le premier fichier d’apprentissage a été constitué à partir de la nomenclature. Les premiers passages donnaient des résultats désastreux ! L’enrichissement a été simple à réaliser.

7.3. Bilan chiffré pour la Martinique

7.3.1 Bilan global sur la Martinique (département complet)

352 989 BI étaient présents dans le fichier. Le taux de BI rejetés est de 15,2 %

Variable par variable, le taux de rejets est de :

<table>
<thead>
<tr>
<th>Variable</th>
<th>Définition</th>
<th>Taux de rejet</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRA</td>
<td>commune de résidence antérieure</td>
<td>2,2 %</td>
</tr>
<tr>
<td>CLT</td>
<td>commune de lieu de travail</td>
<td>0,5 %</td>
</tr>
<tr>
<td>APE</td>
<td>activité de l’établissement (y compris PUB)</td>
<td>29,3 %</td>
</tr>
<tr>
<td>PCS</td>
<td>profession actuelle</td>
<td>20,7 %</td>
</tr>
<tr>
<td>EMPR</td>
<td>emploi précédent</td>
<td>10,3 %</td>
</tr>
<tr>
<td>PRA</td>
<td>pays de résidence antérieure</td>
<td>2,4 %</td>
</tr>
<tr>
<td>N</td>
<td>nationalité</td>
<td>0,7 %</td>
</tr>
<tr>
<td>PN</td>
<td>pays de naissance</td>
<td>0,8 %</td>
</tr>
</tbody>
</table>

(le taux est calculé sur le nombre de BI à coder pour la variable considérée)

7.3.2 Bilan pour l’APE en Martinique

Pour l’ensemble de la Martinique, 104 630 bulletins étaient à coder pour l’APE ; 74 009 ont eu l’APE chiffrée avec un code PUB et 2443 l’APE chiffrée mais sans code PUB, soit 27 % de rejets pour l’APE et 29 % pour le code PUB. La 1ère opération était le passage à Nantes par la MCA (méthode de correspondance automatique), c’est-à-dire la recherche de l’établissement dans SIRENE. Pour les BI non trouvés dans SIRENE, on recherche l’APE par 2 passages QUID : le premier QUID RS à partir de la raison sociale saisie, le deuxième QUID APE à partir de l’activité déclarée.
La Martinique a été découpée en 4 Unités de Traitement (UT).

Le bilan est le suivant :

<table>
<thead>
<tr>
<th></th>
<th>UT1</th>
<th>UT2</th>
<th>UT3</th>
<th>UT4</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI à coder :</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>104630</td>
</tr>
<tr>
<td>Codés par MCA SIRENE :</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22846</td>
</tr>
<tr>
<td>Non codés par MCA mais non passés dans QUID :</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1994</td>
</tr>
<tr>
<td>A passer dans QUID :</td>
<td>10627</td>
<td>21776</td>
<td>26718</td>
<td>20669</td>
<td>79790</td>
</tr>
<tr>
<td>Codés par QUID RS :</td>
<td>2282</td>
<td>3848</td>
<td>5851</td>
<td>4885</td>
<td>16866</td>
</tr>
<tr>
<td>Codés par QUID APE :</td>
<td>4929</td>
<td>11324</td>
<td>11815</td>
<td>8672</td>
<td>36740</td>
</tr>
<tr>
<td>Rejet QUID :</td>
<td>3416</td>
<td>6604</td>
<td>9052</td>
<td>7112</td>
<td>26184</td>
</tr>
<tr>
<td>TOTAL des rejets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28178</td>
</tr>
</tbody>
</table>

Les BI rejetés par la MCA et ne passant pas dans QUID sont ceux n'ayant qu'une adresse. Certains BI n'ayant que l'activité déclarée ne sont pas présentés à la MCA. Ceci explique la différence entre les BI non codés par la MCA et ceux présentés à QUID.

Le taux de succès de la MCA est faible : il s'explique par :

- la non adaptation de cette méthode aux bulletins du RP. En effet, les données collectées au RP n'ont pas la présentation qui est recommandée pour obtenir le meilleur résultat de la MCA. Il est rare que le salarié connaisse bien la raison sociale ou même le sigle ou l'enseigne déclarée à SIRENE. La méthode n'est pas bien adaptée mais elle avait le mérite d'exister et nous n'avions pas les moyens d'en concevoir une autre.

- la moins bonne qualité de SIRENE- Réunion. Les résultats sont en effet moins bons à la Réunion. En effet, une étude plus poussée est à faire mais les résultats sont nettement meilleurs aux Antilles-Guyane où des gros travaux d'amélioration du fichier SIRENE ont été réalisés (enrichissement par d'autres sources, enquêtes de balayage...).
7.3.3 Bilan chiffré pour les autres UT

Le taux de rejets est le suivant :

<table>
<thead>
<tr>
<th></th>
<th>REUNION UT1 à 5</th>
<th>GUYANE UT1</th>
<th>GUADELOUPE UT1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>10,9</td>
<td>18,5</td>
<td>12,2</td>
</tr>
<tr>
<td>CRA</td>
<td>2,3</td>
<td>14,4</td>
<td>3,5</td>
</tr>
<tr>
<td>CLT</td>
<td>0,0</td>
<td>0,2</td>
<td>0,4</td>
</tr>
<tr>
<td>APE</td>
<td>21,3</td>
<td>33,5</td>
<td>22,1</td>
</tr>
<tr>
<td>PCS</td>
<td>23,2</td>
<td>27,3</td>
<td>18,7</td>
</tr>
<tr>
<td>EMPR</td>
<td>11,0</td>
<td>14,8</td>
<td>8,1</td>
</tr>
<tr>
<td>PRA</td>
<td>1,3</td>
<td>1,7</td>
<td>0,5</td>
</tr>
<tr>
<td>PN</td>
<td>7,0</td>
<td>4,8</td>
<td>0,2</td>
</tr>
</tbody>
</table>

7.4. Bilan général

L’opération s’est révélée très positive. Seulement 15 % des bulletins ont été traités manuellement pour au moins un code. Le coût en personnel pour le chiffrage a donc été divisé par 6 par rapport à une opération de chiffrage classique. En contre partie, la saisie a été nettement plus importante (environ multipliée par 3). Aucune mesure très précise de la qualité n’a été effectuée. Les premières études faites sur les fichiers n’ont pas permis de déceler d’incohérences ou d’erreurs. Néanmoins, le travail d’amélioration des fichiers d’apprentissage a été long et la procédure de recherche dans SIRENE peu performante ce qui a induit un travail amont plus important.


Partie 4

L’ÉCHANTILLONNAGE
L’ÉCHANTILLONNAGE ÉQUILIBRÉ DE DISTRICTS ET L’EXPLOITATION AU 1/20 DU RECENSEMENT

Les quatre études qui suivent ont été écrites entre 1986 et 1988 et sont relatives à l’échantillonnage au 1/20 du RP.


Les trois autres sont des notes d’étapes de ce travail où on trouvera, de manière peut-être moins définie mais plus complète, l’ensemble de la méthodologie et des preuves par simulation des méthodes qui ont été utilisées.

Pour lier ces quatre études -qui parlent de la même chose- il peut s’avérer intéressant de raconter un peu l’histoire de ce travail. Il y a sans doute là un exemple majeur où la théorie rejoint la pratique, voire l’industrie. Un cas exemplaire où des idées extrêmement abstraites au départ ont permis de gagner ou d’économiser beaucoup d’argent.


L’inconvénient de cette technique était relativement bien connu. Comme les unités d’un même district se ressemblent assez, le sondage perd en précision. Une étude de F. CHARTIER en 1978 avait d’ailleurs chiffré cette perte : la variance de la plupart des variables était multipliée par 4 ou 5 ce qui rendait la méthode inutilisable.

L’équipe du recensement demanda cependant à ce que l’étude soit renouvelée. Or la connaissance fait des progrès. L’équipe des sondeurs de l’INSEE était alors fort impressionnée par les travaux de HAJEK (1981) et de ROYALL et HERSON (1973)
(voir la bibliographie de l'article qui suit). Ceux-ci montraient, par des approches radicalement différentes, l'intérêt que pouvaient avoir des échantillons respectant les moyennes de variables de contrôle connues. Ils déploraient la difficulté de construire concrètement de tels échantillons dans la pratique des sondages. J'ai découvert, plus tard, que le grand Pierre THIONNET avait eu, 20 ans auparavant, les mêmes intuitions.

Il se trouve aussi que nous avions quelques idées sur la façon d'arriver à l'équilibrage d'un échantillon qui avaient déjà servi pour de petits problèmes. Il était tentant d'essayer grandeur nature ces "techniques nouvelles". Les facilités de SAS pour enchaîner des opérations statistiques complexes alliées aux talents réunis dans l'équipe sondage permirent la mise au point d'une méthode de tirage équilibrée. Le prototype du programme de tirage permit d'avoir la certitude que, conformément à la théorie de HAJEK en particulier, on arrivait à des gains de précision substantiels qui contrepouvaient complètement l'effet de grappe pour les variables les plus liées à la démographie des personnes et des ménages. Les résultats étaient nettement moins favorables pour les variables liées à la nationalité ou au logement.

Ces dernières toutefois, devaient être traitées grâce à l'exploitation exhaustive légère, de sorte que la technique d'échantillonnage équilibré convenait parfaitement pour le dépouillement au 1/20.

Comme dans un processus industriel, il fallut passer du prototype à un programme opérationnel pour le recensement, ce que l'équipe du recensement su parfaitement réaliser.
1 - Définitions et considérations générales

Soit $X = (X_1, X_2, \ldots, X_p)^T$ un vecteur de variables numériques disponibles pour chaque unité d'une population de taille $N$. La valeur de $X_k$ pour l'unité $i$ est notée $X_{k,i}$ et $X (i) = (X_{1,i}, X_{2,i}, \ldots, X_{p,i})^T$. Un échantillon $s$ de taille $n$ tiré dans la population est dit équilibré pour l'ensemble des variables $X_k$ si on a :

$$
\bar{x} = \frac{1}{n} \sum_{s} X (i) = \frac{1}{N} \sum_{u} X (i) \quad \text{où} \quad \sum_{s} \text{et} \sum_{u} \text{désignent respectivement la sommation sur l'échantillon et sur la population.}
$$

L'obtention d'échantillons équilibrés est une vieille préoccupation des sondeurs. C'est une façon de fabriquer un "échantillon représentatif", car, en un sens naïf, l'échantillon est conçu comme un modèle réduit de la population. Dans une optique de sondage probabiliste, l'équilibrage n'est pas un but en soi et n'est réalisé qu'en espérance mathématique seulement. Dans un contexte d'enquête analytique (ROYALL et HERSON - 1973), l'équilibrage apparaît comme une condition pour obtenir des estimations robustes. HAJEK (1971) insiste sur les performances des échantillons équilibrés pour améliorer la précision des estimateurs les plus classiques.

L'idée d'équilibrer l'échantillonnage n'est pas sans parenté avec celle d'estimation par régression. Supposons que les moyennes empiriques $(\bar{y}, \bar{x})$ de la variables d'intérêt $Y$ et d'un vecteur de variables auxiliaires $X$ suivent approximativement une loi normale multidimensionnelle. L'estimateur par régression $\hat{Y}_r = \bar{y} + b' (\bar{X} - \bar{x})$ est généralement meilleur que la moyenne simple sur l'échantillon comme estimateur de $\bar{Y}$. Si par hasard, il se trouve qu'on ait $\bar{X} = \bar{x}$, les deux estimateurs coïncident. Une idée destinée à conserver la simplicité et l'efficacité de l'estimateur consiste à ne prendre en compte

1 Extrait de Efficient sampling algorithms and balanced samples. Communication invitée au congrès. COMPSTAT (Computational Statistics 1988)
que des échantillons pour lesquels on a l'égalité $\overline{X} = \bar{x}$. Si $X$ est de dimension 1, on peut y arriver de diverses façons (voir par exemple YATES-1949 ou THIONNET-1953). Nous proposons ici un nouvel algorithme pour équilibrer un échantillon quand $X$ est multidimensionnel.

2 - Description de l'algorithme

L'idée informelle est de retirer de l'échantillon courant quelques unités qui sont dans la direction de $\bar{x} - \overline{X}$ pour les remplacer par des unités qui sont dans la direction opposée. Le nombre d'unités à échanger est calculé pour obtenir en moyenne l'équilibre. Le procédé est ensuite itéré jusqu'à ce qu'il ne soit plus utile de procéder à des changements.

Décrivons cette méthode en des termes plus précis. Par convention nous supposerons que $\overline{X} = 0$. On commence par partager la base de sondage en $G$ groupes ($g = 1$ à $G \geq p$) de façon à ce que les moyennes $\overline{X^g}$ dans les différents groupes soient très différentes et à ce que les effectifs $N_g$ de chaque groupe soient à peu près égaux. En pratique, pour $p = 2$ à $5$, on peut construire $2^p$ groupes selon les modalités possibles du vecteur des signes de $(X_1, \ldots, X_p)$. On peut transformer linéairement les variables avant cette opération. Il se révèle spécialement utile d'utiliser les composantes principales de $X$, comme cela a été réalisé dans l'application décrite plus loin.

L'algorithme comporte les étapes suivantes :

- Étape 1 : Stratification avec allocations proportionnelles. 
  Dans chaque groupe $g$, tirer un échantillon aléatoire simple de taille $n_g = n \cdot N_g / N$. Les nombres $n_g$ sont arrondis à des entiers en utilisant la méthode décrite à l'étape 3.
  Posons $A = n$ et $t = 1$.
  Soit $\overline{x^g}$ la moyenne de l'échantillon tiré dans le groupe $g$ et $\overline{x^{(t)}} = \frac{1}{n} \sum_{g} n_g \overline{x^g}$

- Étape 2 : modification de l'échantillon.
  Chercher un ensemble d'entiers $v_g$.
  Si $v_g > 0$, on ajoute $v_g$ nouveaux éléments tirés du groupe $g$ par sondage aléatoire simple.
  Si $v_g < 0$, on élimine ($v_g$) éléments de l'échantillon dans le groupe $g$ par sondage aléatoire simple.
  Pour maintenir à $n$ la taille de l'échantillon et à 0 l'espérance de $\bar{x}$ après modification, les nombres $v_g$ doivent satisfaire à :
\[ \sum_g u_g = 0 \]
\[ \sum_g u_g \bar{X}_g^g = -n \bar{x}(t) \]

- Pour déterminer de façon unique les \( u_g \)'s, on minimise la quantité \( \sum_g u_g^2 / N_g \), la variation de l'allocation entre les groupes.

Soit \( P \) la matrice diagonale dont les éléments diagonaux sont les \( N_g^{-1} \), \( v \) le vecteur des \( u_g \) et \( Z \) la matrice \( (p + 1) \times G \) dont les p premières lignes sont les \( \bar{X}_g^g \) et avec \( Z_{p+1, g} = 1 \).

On cherche à trouver \( v \) minimisant \( v'P v \) sous les contraintes \( Zv = -n \bar{x} \) où \( \bar{x} = (\bar{x}_, (t), 0) \). La solution est :

\[ u = nR\bar{x} \]

avec \( R = -P^{-1}Z' \left( \begin{array}{c} Z \ P^{-1} \ Z' \end{array} \right)^{-1} \)

- Étape 3 : Arrondissage

Arrondir les nombres obtenus à l'étape 2 de la façon suivante :

soit \( v_g^- \) plus grand entier vérifiant \( v_g^- \leq v_g \) et \( r_g = v_g - v_g^- \).

La somme des \( r_g \) est un entier \( d < p \).

On tire \( d \) groupes au hasard avec des probabilités égales aux \( r_g \) en utilisant un algorithme quelconque (voir par exemple HANIF et BROWER (1980) ou SONTNER 1986 ou DEVILLE-GROSBRAS (1987)).

Pour les groupes ainsi tirés on prend \( v_g = v_g^- + 1 \) et pour les autres \( v_g = v_g^- \)

- Étape 4

Si \( \Sigma \left| v_g \right| \geq A \), garder l'échantillon et fin.

Sinon poser \( t = t + 1 \) et \( A = \Sigma \left| v_g \right| \).

Pour chaque groupe \( g \) ayant \( v_g < 0 \) tirer \( \left| v_g \right| \) unités dans l'échantillon par sondage aléatoire simple et les retirer de l'échantillon.

Pour chaque groupe \( g \) ayant \( v_g > 0 \) tirer \( \left| v_g \right| \) nouvelles unités dans le groupe \( g \) par sondage aléatoire simple et les ajouter à l'échantillon. Calculer \( \bar{x}(t) \) et aller à l'étape 2.

L'échantillonnage 303
3 - Quelques aspects théoriques et pratiques

Il est clair que, pour de petites populations et de petits échantillons, l'équilibrage n'est pas toujours possible. Pire, si nous voulons équilibrer au mieux l'échantillon, il peut nous arriver de devoir modifier très fortement les probabilités d'inclusion (réfléchissez par exemple au cas où n = 1 ou n = 2). Les bonnes propriétés probabilistes de l'algorithme décrit ci-dessus sont, en un certain sens, asymptotiques et se basent sur des approximations normales des lois suivies par des quantités comme $n^{1/2} \frac{x}{(t)}$.

Si n est suffisamment grand on peut prouver les propriétés suivantes :

- Pour l'échantillon obtenu à l'itération t de l'algorithme on a :

$$E \frac{x}{(t)} = 0 \left( n^{(1-0)/2} \right)$$

A chaque itération l'ordre de grandeur de $\frac{x}{(t)}$ est divisé $n^{1/2}$ de sorte que l'équilibrage est obtenu à la limite.

- Pour l'unité i du groupe g, la probabilité d'appartenir à l'échantillon est :

$$\frac{n}{N} \left( 1 + \frac{K_i}{n} \right) + 0 \left( 1 / n \right)$$

Où $K_i$ est une quantité finie fixe qui dépend de $X(i) - X^g$.

- Pour deux unités distinctes i et j la probabilité pour le couple (i,j) d'appartenir à l'échantillon est approximativement égale à

$$\left( \frac{n}{N} \right) \left( 1 + \frac{K}{n} \right)$$

Des simulations ont montré que l'équilibrage est obtenu avec une très bonne approximation et qu'il n'apparaît pas de distorsion détectable pour les probabilités d'inclusion.

On doit néanmoins prendre en compte quelques aspects pratiques. D'abord on doit utiliser une règle d'arrêt bien précise de la procédure itérative. Théoriquement la procédure d'arrondissage produit à chaque itération la même variance sur $\frac{x}{(t)}$. Il est logique d'arrêter quand on n'espère plus d'amélioration. Ceci arrive, pratiquement, quand nous devrions faire plus de modifications à l'itération courante que lors de la précédente.

Un autre aspect réside dans le volume de calcul nécessaire par la procédure. On remarquera que la matrice R définie à l'étape 2 reste identique pour toutes les itérations et n'est calculée qu'une seule fois. De même, on peut simplifier l'étape 4 si on utilise la méthode suivante :

- la base de sondage est mise dans un ordre aléatoire pour chaque groupe avant l'étape 1 ;
- Pour l'unité de rang r, les sommes des $X_k$ jusqu’au rang r sont stockées comme de nouvelles variables ;
- à chaque itération, un échantillon requérant \( n_g \) unités dans le groupe \( g \) sera constitué des premières unités du groupe. Le calcul de \( x_{(0)} \) devient alors immédiat.

4 - Une étude par simulation

Le problème de l'équilibrage d'un échantillon vis-à-vis de plusieurs variables s'est posé dans le cadre du dépouillement du recensement français de 1990. Traditionnellement, on publie une estimation rapide des principaux résultats à partir d'un échantillon du 1/20ème des ménages. Jusqu'en 1982 l'échantillon était tiré à la main par tirage systématique. Ce travail était peu attrayant, lent et occasionnait de nombreuses erreurs de report des restes.

La solution alternative consiste à tirer des groupements de logements, les districts, dans le but d'obtenir un traitement plus rapide évitant au maximum les traitements manuels. Cependant ce mode de tirage de l'échantillon augmente les erreurs dues à l'échantillonnage. Pour limiter cet effet on peut utiliser des informations liées à la taille des districts. Plus précisément on dispose à ce stade de trois dénombrements différents à savoir : le nombre de personnes (\( X_1 \)), le nombre de logements (\( X_2 \)) et le nombre d'immeubles (\( X_3 \)) ; ces données sont connues au niveau de chaque district. Malgré la forte corrélation entre ces critères, leur réduction à une seule dimension peut induire une perte en précision notable ; l'usage d'outils classiques comme le sondage systématique sur un fichier trié peut se révéler insatisfaisant.

Les simulations rapportés dans la suite testent l'efficacité de la méthode d'équilibrage qui vient d'être décrite. On y utilise deux critères de contrôle qui sont les deux premières composantes principales tirées de \( X_1, X_2, X_3 \).

On a comparé les méthodes suivantes :
- Sondage aléatoire simple de districts (SRS)
- Première étape de l'échantillonnage équilibré (c'est-à-dire stratification)
- Echantillon équilibré
- Estimation par régression (après stratification initiale)

Les variances ou les écarts-types présentés au tableau 1 et au graphique 1 sont basés sur 100 échantillons tirés pour chacune des méthodes.

Remarques :
  a) La troisième composante principale a été négligée à cause de sa très faible contribution à l'inertie totale. En gardant trois dimensions on obtient des résultats très voisins.
b) L'estimateur par régression a été utilisé sur l'échantillon stratifié tiré de la première étape de l'équilibrage de façon à rendre plus évidente la comparaison entre les méthodes.

c) Les échantillons aléatoires simples résultent nécessairement d'un autre ensemble de simulations. Parfois le sondage aléatoire simple peut sembler plus efficace que d'autres méthodes mais, dans ces cas, les différences ne sont pas significatives.

On a utilisé 36 variables, dont les 3 variables de contrôle, pour évaluer l'efficacité des différentes stratégies. Les données du test proviennent du dernier recensement français. Ce sont des données infra-régionales totalisant 2542 districts. Toutes les simulations sont été réalisées sur un IBM 3090/120E en utilisant le macro-langage SAS® et la PROC MATRIX.

On constate que, pour une bonne moitié ces variables étudiées, les écarts types et donc les intervalles de confiance sont réduits de plus de la moitié en comparaison du sondage aléatoire simple de district (figure 1). Comparé à l'estimation par régression, l'échantillonnage équilibré est légèrement moins efficace, spécialement pour les variables qui ont les meilleurs résultats. Cette plus forte variabilité provient du fait que l'équilibrage n'est jamais réalisé qu'approximativement avec la procédure qui est proposée. En tout état de cause celle-ci autorise une tabulation bien plus aisée, ce qui, pour le dépouillement du recensement, s'avère un avantage décisif.

Enfin, si nous revenons au problème initial, on a comparé l'échantillonnage équilibré au tirage systématique de districts. Comme on pouvait s'y attendre, l'effet de grappe dû au fait qu'on échantillonne des districts demeure assez gros pour des populations rares ou fortement regroupées. Certaines caractéristiques socio-économiques comme la nationalité ou des activités économiques particulières n'ont qu'une faible liaison avec la taille des districts. Ces variables ont un mauvais comportement pour toutes les méthodes qui opèrent au niveau des districts, accusant une perte de précision allant de 2 à 4 en terme de variance, par comparaison à un échantillonnage de logements. A l'opposé l'effet de grappe du tirage en district est complètement effacé pour les "bonnes" variables ; ainsi, celles qui sont très liées au nombre de logements restent un petit peu mieux estimées par le sondage systématique. En revanche les variables bien expliquées par le nombre de personnes donnent de meilleurs résultats en utilisant l'échantillonnage équilibré.


HAJEK, J. (1981), Sampling from a finite population. DEKKER.


**ANNEXE**

### Tableau 1

**COMPARÉE EFFICACITÉ DES MÉTHODES DE SAMPLING**

**RELATIVE ESTIMATED VARIANCE (ESTIMATION ON 100 SAMPLES)**

<table>
<thead>
<tr>
<th>TEST VARIABLES</th>
<th>BALANCED /SRS</th>
<th>STRATIFICATION /SRS</th>
<th>STRATIFICATION+ REGRESSION /SRS</th>
<th>BALANCED+ REGRESSION /SRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ACTIFS EMPLOYÉS CHIMIE DE BASE</td>
<td>1.20</td>
<td>1.20</td>
<td>1.18</td>
<td>1.19</td>
</tr>
<tr>
<td>2 ACTIFS EMPLOYÉS PRODUCTION COMBUSTIBLES</td>
<td>1.14</td>
<td>1.14</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td>3 ACTIFS EMPLOYÉS PHARMACIE ET PHARMACIE</td>
<td>1.26</td>
<td>1.25</td>
<td>1.26</td>
<td>1.27</td>
</tr>
<tr>
<td>4 NOMBRE D'ETRANGERS</td>
<td>0.97</td>
<td>1.37</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td>5 NOMBRE D'ETRANGERS ACTIFS SEXE MASculIN</td>
<td>0.86</td>
<td>0.89</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>6 AGRICULTEURS EXPLOITANTS</td>
<td>0.75</td>
<td>0.94</td>
<td>0.86</td>
<td>0.81</td>
</tr>
<tr>
<td>7 CADRES</td>
<td>0.73</td>
<td>0.81</td>
<td>0.79</td>
<td>0.75</td>
</tr>
<tr>
<td>8 NOMBRE DE PERSONNES DE 75 ANS ET PLUS</td>
<td>0.72</td>
<td>0.86</td>
<td>0.78</td>
<td>0.72</td>
</tr>
<tr>
<td>9 ACTIFS EMPLOYÉS AGRICULTURE ET SEL</td>
<td>0.66</td>
<td>0.82</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td>10 ARTISANS,COMMERCANTS,CHefs D'ENTREPRISE</td>
<td>0.65</td>
<td>0.71</td>
<td>0.63</td>
<td>0.67</td>
</tr>
<tr>
<td>11 NOMBRE DE LOGEMENTS ACTIFS</td>
<td>0.50</td>
<td>0.74</td>
<td>0.64</td>
<td>0.57</td>
</tr>
<tr>
<td>12 NOMBRE DE PERSONNES ACTIVE 50-54 ANS</td>
<td>0.35</td>
<td>0.27</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>13 NOMBRE DE LOGEMENTS VACANTS/SECONDAIRES</td>
<td>0.35</td>
<td>0.65</td>
<td>0.52</td>
<td>0.55</td>
</tr>
<tr>
<td>14 NOMBRE D'ETRANGERS ACTIFS 20-24 ANS</td>
<td>0.48</td>
<td>0.45</td>
<td>0.46</td>
<td>0.44</td>
</tr>
<tr>
<td>15 NOMBRE DE PERSONNES 55-59 ANS</td>
<td>0.47</td>
<td>0.77</td>
<td>0.41</td>
<td>0.45</td>
</tr>
<tr>
<td>16 NOMBRE D'HOMMES 20-24 ANS</td>
<td>0.45</td>
<td>0.50</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>17 NOMBRE DE PERSONNES DE 25 A 34 ANS</td>
<td>0.39</td>
<td>0.50</td>
<td>0.35</td>
<td>0.34</td>
</tr>
<tr>
<td>18 ACTIFS EMPLOYÉS COMMERCE DE DETAIL</td>
<td>0.39</td>
<td>0.44</td>
<td>0.37</td>
<td>0.38</td>
</tr>
<tr>
<td>19 RETRAITES</td>
<td>0.39</td>
<td>0.56</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>20 PROFESSIONS INTERMEDIAIRES</td>
<td>0.28</td>
<td>0.57</td>
<td>0.27</td>
<td>0.26</td>
</tr>
<tr>
<td>21 NOMBRE DE PERSONNES DE 0 A 14 ANS</td>
<td>0.26</td>
<td>0.66</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>22 NOMBRE DE PERSONNES DE 15 A 24 ANS</td>
<td>0.25</td>
<td>0.44</td>
<td>0.27</td>
<td>0.20</td>
</tr>
<tr>
<td>23 ACTIFS EMPLOYÉS INDUSTRIE</td>
<td>0.25</td>
<td>0.48</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>24 EMPLOYÉS</td>
<td>0.25</td>
<td>0.45</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>25 CADRES</td>
<td>0.22</td>
<td>0.41</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>26 AUTRES INACTIFS</td>
<td>0.22</td>
<td>0.39</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>27 ACTIFS EMPLOYÉS TERTIAIRE</td>
<td>0.19</td>
<td>0.45</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>28 NOMBRE DE PERSONNES DE 25 A 34 ANS</td>
<td>0.18</td>
<td>0.47</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>29 NOMBRE DE PERSONNES DE 35 A 44 ANS</td>
<td>0.17</td>
<td>0.55</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>30 NOMBRE DE FEMMES ACTIVES</td>
<td>0.16</td>
<td>0.45</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td>31 NOMBRE D'HOMMES ACTIFS</td>
<td>0.10</td>
<td>0.52</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>32 NOMBRE D'HOMMES</td>
<td>0.09</td>
<td>0.43</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>33 NOMBRE DE LOGEMENTS</td>
<td>0.08</td>
<td>0.53</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>34 NOMBRE DE PERSONNES PRINCIPALES</td>
<td>0.08</td>
<td>0.42</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>35 NOMBRE DE FEMMES</td>
<td>0.07</td>
<td>0.41</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>36 NOMBRE DE PERSONNES</td>
<td>0.07</td>
<td>0.41</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Figure 1**

**Figure 1** : Comparée efficacité des méthodes de sampling

**ESTIMATED STANDARD ERROR RELATIVE TO SRS (ESTIMATION ON 100 SAMPLES)**

![Figure 1](image.png)
4.2 - EXPÉRIMENTATION DU TIRAGE AU 1/20 DU RP COMME UN ÉCHANTILLON DE DISTRICTS
Nicole Roth

L’expérimentation porte sur la possibilité de définir le 1/20 comme un échantillon de districts tiré à partir du 1/4. Une étude analogue a été menée en 1978-79 par CHARTIER (note 648/470 du mars 79) : la comparaison a été faite sur la base de simulations d’échantillons par tirages systématiques de logements d’une part et tirages de districts d’autre part. Il en ressort que la variabilité du tirage par districts est nettement supérieure à celle du tirage par logements. Du point de vue de la précision des estimations, l’échantillonnage par logements apparaît donc plus efficace. D’un point de vue pratique, il faut cependant noter que l’on a constaté dans le 1/20 "classique" des incohérences pouvant provenir de la gestion manuelle du tirage. Cette source d’erreurs devrait être moins importante dans le cas d’un échantillonnage par districts (ce gain est cependant difficile à chiffrer). Par ailleurs, le tirage du 1/20 "classique" coûte cher.

Pour améliorer un éventuel échantillonnage par districts, il a été décidé d’examiner deux variantes des calculs menés par CHARTIER (sur 2 régions "représentatives").

1 - On examinera la possibilité de faire varier le taux de sondage des districts : le tirage de logements étant plus onéreux que le tirage de districts, on peut tirer une fraction $\frac{1}{K} > \frac{1}{5}$ pour un même coût d’exploitation.

En pratique $k = 3$ ou $4$ (en deçà, le différentiel entre $\frac{1}{4}$ et $\frac{1}{4k}$ ne justifierait pas une exploitation séparée).

Symétriquement, on examinera la possibilité d’utiliser un taux plus faible (par exemple $k = 10$), cette possibilité pouvant se justifier dans les régions importantes.

2 - Une procédure de tirage "contraint" sera utilisée (et non plus systématique) pour l’échantillonnage de districts : on tirera un échantillon respectant certaines variables de contrôle (nombre de BI, de logements ou d’immeubles).

Par exemple, on tire un échantillon de districts respectant deux moyennes $\overline{BI}$ et $\overline{LG}$ (moyennes du nombre de BI et du nombre de logements par districts) ; à chaque étape,
on compare les valeurs réalisées aux objectifs \((\overline{\text{BI}}, \overline{\text{LG}})\) et on ajuste le tirage suivant de façon à compenser les écarts constatés.

Les simulations :
Elles porteront sur les variables dont la liste est donnée en annexe 1. Il paraît intéressant d'étudier des variables de type total et de type ratio.
La référence : l'échantillonnage par logements au taux 1/5 à partir de la bande 1/4 (le 1/20 "classique").

On souhaite disposer de 10 simulations pour calculer la variabilité des estimateurs. On ne peut donc pas procéder à des tirages systématiques simples (ce qui fournirait au maximum 5 échantillons différents), pour un ordre donné du fichier. Deux autres solutions sont envisageables :
- tirage systématique à entrées multiples,
- tirage stratifié par bloc de taille 5 (cf. annexe 2).

L'échantillonnage par districts :
- On effectue 10 simulations pour trois taux différents \((k = 3, 5\) et 10).

Les opérations à effectuer sont les suivantes, pour chacune des régions choisies NORD - PAS DE CALAIS, AUVERGNE.

1 - Extraction des variables utiles à partir des fichiers 1/4 du RP :
- variables identifiant les districts :
  le district est le véritable district de recensement pour les communes de TCD\(\geq1000\),
  ce "district" est la commune prise en totalité pour les communes de TCD<1000.
On retiendra les variables : D, C, CIL, IL, FIL, TCD, CLD, ACTR, AGEP, DU, TU, U, TZPIU
- variables identifiant les logements : IMM, LOG
- variables "test" :
  au niveau immeuble : AAC1
  au niveau logement : CL
  au niveau individu : S, TA, AD, AG, CS8, CS, AE40, IN

2 - Extraction d'un fichier de districts à partir du précédent, puis appariement de ce fichier sur la population légale fichier DISTR.

Le fichier obtenu comprend, outre l'identifiant des districts, les variables de la population légale (nombre de BI, de logements et d'immeubles). Il serait intéressant de disposer de certaines variables de type "total" :
- nombre d'étrangers,
- nombre de moins de 15 ans,
- nombre de plus de 65 ans,
- nombre d’ouvriers,
- nombre de cadres supérieurs et professions libérales,
- nombre d’agriculteurs exploitants,
- nombre de chômeurs.

3 - Extraction d’un fichier de logements à partir du premier (comprenant les identifiants des logements) que la tabulation du 1/4 qui servira de référence ; on étudiera notamment la variance des estimateurs, mais aussi d’autres indicateurs de dispersion ou d’asymétrie. → fichier LOGMT.

4 - Tirages de districts à partir du fichier DISTR. On crée ainsi 30 = 3 x 10 indicatrices d’appartenance à l’échantillon (10 simulations pour chacun des trois taux).

5 - Tirage "quasi - systématiques" de logements pour obtenir la variabilité de référence des estimateurs. On crée ainsi 10 variables indicatrices d’appartenance à l’échantillon.

6 - "Appariement du fichier initial au 1/4 sur le fichier au niveau district" (comportant 30 indicatrices) et sur le fichier au niveau logement (comportant 10 indicatrices).

7 - Le calcul des estimateurs :
- On estimera les variables de type total, moyenne ou ratio pour la population totale par leurs équivalents sur les différents échantillons obtenus.
- Pour chacune des quatre méthodes, on obtient ainsi dix estimateurs, et ceci pour chacune des 46 variables étudiées (cf. annexe 1).
- Pour estimer de façon précise la variabilité des résultats, la division méthodes de sondages souhaite disposer du fichier donnant pour chaque région et pour chacune des 46 variables, les 40 estimateurs obtenus lors des simulations d’échantillonnage ainsi que la tabulation du 1/4 qui servira de référence ; on étudiera notamment la variance des estimateurs, mais aussi d’autres indicateurs de dispersion ou d’asymétrie.
ANNEXE 1

Liste des variables à étudier

1 - Nombre total de logements
2 - Nombre de résidences principales CL = 1 à 5
3 - Nombre de logements vacants et résidences secondaires CL = 6 ou 7
4 - Nombre de logements récents (1975 ou après) AAC1 = 7
5 - Nombre de personnes
6 - Nombre d'hommes au total S = 1
7 - Nombre d'hommes actifs S = 1 et TA = 1 ou 3
8 - Nombre de femmes au total S = 2
9 - Nombre de femmes actives S = 2 et TA = 1 ou 3
10 - Nombre d'hommes de 20 à 24 ans, au total S = 1 et AQ = 20
11 - Nombre d'hommes de 20 à 24 ans actifs S = 1 et AQ = 20 et TA = 1 ou 3
12 - Nombre de femmes de 30 à 34 ans, au total S = 2 et AS = 30
13 - Nombre de femmes de 30 à 34 ans, actives S = 2 et AQ = 30 et TA = 1 ou 3
14 - Nombre de personnes de moins de 15 ans AQ 15
15 - Nombre de personnes de 15 à 24 ans AQ = 15 ou 20
16 - Nombre de personnes de 25 à 34 ans AQ = 25 ou 30
17 - Nombre de personnes de 35 à 54 ans AQ = 35 ou 40 ou 45 ou 50
18 - Nombre de personnes de 55 à 74 ans AQ = 55 ou 60 ou 65 ou 70
19 - Nombre de personnes de 75 et plus AQ 70
20 - Nombre d'étrangers, au total IN = 20
21 - Nombre d'étrangers, actifs, hommes IN = 20 et TA = 1 ou 3 et S = 1
22 - Age moyen AD
23 - Age moyen des étrangers
24 - Nombre de personnes de CS = 3 CS8 = 3
25 - Nombre d'hommes de CS = 3 CS8 = 3 et S = 1
26 - Nombre de personnes de CS = 5 CS8 = 5
27 - Nombre d'hommes de CS = 5 CS8 = 5 et S = 1
28 - Nombre de personnes de CS = 6 CS8 = 6
29 - Nombre d'hommes de CS = 6 CS8 = 6 et S = 1
30 - Nombre d'inactifs CS = 71,72, 74, 75, 77, 78, 83 à 86
31 - Nombre d'hommes inactifs et S = 1

Nombre de personnes d'activité économique
32 - Production de combustibles minéraux solides AE40 = 040
33 - Chimie de base AE40 = 110
34 - Parachimie et industrie pharmaceutique AE40 = 120
35 - Commerce de détail AE40 = 270 ou 280
Ratios
36 - % de logements achevés après 75 : (4) / (1)
37 - Taux d'activité des hommes : (7) / (6)
38 - Taux d'activité des femmes : (9) / (8)
39 - Taux d'activité des hommes de 20 à 24 ans : (11) / (10)
40 - Taux d'activité des femmes de 30 à 34 ans : (13) / (12)
41 - % de moins de 25 ans ( (14) + (15) ) / (5)
42 - % de plus de 75 ans (19) / (5)
43 - % de cadres supérieurs parmi les actifs (24) / (5) - (30)
44 - % d'ouvriers parmi les actifs (28) / ( (5) - (30) )
45 - % d'étrangers (20) / (5)
46 - % d'étrangers actifs hommes/actifs hommes : (21) / (7)

L'échantillonnage
ANNEXE 2

Un tirage systématique classique au 1/5 ne permettant d’obtenir que 5 échantillons, on risque ainsi de biaiser la comparaison des différentes méthodes. Il est donc souhaitable de définir une méthode permettant d’obtenir 10 simulations différentes d’échantillonnage "quasi-systématique".

Deux techniques "voisines" sont envisageables :
- un tirage stratifié après découpe de la population en blocs de taille 5 (on tire une unité par strate),
- un tirage systématique à entrées multiples : on tire S unités sans remise dans le premier bloc et les suivants par tirage systématique initialisé sur le premier bloc.

W. GAUTSCHI a étudié les deux types de méthodes sous différentes hypothèses de superpopulation : en particulier, dans le cas où il y a corrélation entre individus (corrélation décroissante et fonction uniquement de la différence de rang entre individus), il montre que la variance d’un sondage systématique à entrées multiples est plus forte que celle d’un sondage stratifié avec tirage d’une unité par strate.

Un deuxième argument vient étayer la préférence pour le sondage stratifié : c’est celui qui, en pratique, semble le plus proche du tirage classique du 1/20 ème.
4.3 - RÉSULTATS OBTENUS DANS L’EXPÉRIMENTATION DU TIRAGE DU 1/20ème COMME UN ÉCHANTILLON DE DISTRICTS
(note 1222/470 du 28 avril 1986)

Nicole Roth

Le but de l’opération est de comparer les performances d’un échantillonnage de districts à celles de l’échantillonnage classique de logements :
- pour les régions 31 et 83 : on a simulé 10 échantillons pour chacun des 4 types de tirage : tirage contrôlé de districts au 1/10ème, au 1/5ème et au 1/3, tirage quasi-systématique de logements au taux 1/5.
- cette étude a été reprise sur les régions 52 et 73, comparant les performances de 3 échantillonnages à taux fixe (1/5) : tirage quasi systématique de districts (non contrôlé), tirage contrôlé de districts, tirage quasi-systématique de logements.

1 - Les critères de contrôle

Pour chaque région, les tirages contrôlés de districts se font à l’intérieur de strates définies par un critère de type taille d’unité urbaine (cf. annexe). Le principe du tirage contrôlé consiste à se rapprocher au mieux de k critères de contrôle ; on définit donc $2^k$ groupes. A chaque pas du tirage, la valeur des critères définit le groupe dans lequel on doit effectuer le prochain tirage pour aller dans le sens du rééquilibrage. En pratique, cette procédure peut conduire à une impossibilité (groupes dissymétriques).

Dans le cas de l’échantillonnage de districts, les variables de contrôle sont :
- le nombre de BI (PSDC)
- le nombre de FL (FL)
- le nombre d’immeubles (A)
définissant théoriquement 8 groupes. Cette situation aboutissant souvent à l’échec du tirage (insuffisance d’un groupe) on s’est limité à 2 critères construits à partir des trois précédents : ces 2 critères synthétiques sont obtenus à partir d’une analyse en composantes principales des variables originelles : on retient les deux premiers axes pour chaque croisement strate x région ; la part d’inertie expliquée par ces 2 axes dépasse généralement 96 % (cf. annexe). Cette même étude montre qu’aucune des trois variables n’est vraiment redondante :

- sur le premier axe, les trois variables ont une importance sensiblement égale (effet taille)
- sur le deuxième axe, la variable FL apparaît cependant moins importante que les deux autres (elle explique en moyenne environ 16 % de l’inertie contre 33 % pour PSDC et 51 % pour A)
- l’axe 3 en contrepartie est expliqué à 47 % par la variable FL (en moyenne), la partie résiduelle étant expliquée par la variable A pour les strates TU = 1 à 3 et par la variable PSDC pour les strates TU = 4 et plus.

L’effet de chacune des 3 variables peut s’apprécier de façon plus approfondie à partir de leur signe respectif sur les axes factoriels notamment sur les 2e et 3e axes (l’effet taille jouant sur le 1er axe, les 3 variables apparaissent du même côté).

On crée pour cela 2 variables de type densité \[ \{ \begin{align*}
\text{PSDC/FL (densité de population)} \\
A/FL (densité de logements)
\end{align*} \]

Les variables entourées dans le tableau sont celles correspondant aux contributions les plus fortes ; pour les strates "rurales" (0 à 3), la variable pertinente sur le 2ème axe est la variable densité de population ; sur les strates urbaines (4 et plus), c’est la variable densité de logements qui prédomine ; en contrepartie, on récupère sur le 3ème axe une composante densité de logements pour les strates rurales et une composante densité de population pour les strates urbaines.

En conclusion, il apparaît donc que :

1) On ne peut négliger aucune des variables : nombre de BI, nombre de FL, nombre d’immeubles.

2) Négliger l’axe 3 dans le tirage revient à relâcher le contrôle sur la variable densité de logements (nombre de logements, nombre d’immeubles) dans les strates "rurales" et sur la variable densité de population (nombre de logements, nombre de BI) dans les strates "urbaines".
### Tableau 1

<table>
<thead>
<tr>
<th>Région</th>
<th>Strate</th>
<th>PSDC (2ème axe)</th>
<th>FL</th>
<th>A</th>
<th>PSDC/FL</th>
<th>A/FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>1, 2, 3</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>4, 5</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>31</td>
<td>0, 1</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2, 3</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>4, 5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>76201</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75902</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75901</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>70052</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>70051</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>1, 2, 3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>4, 5</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>67201</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>64901</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>64401</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>74401</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>73</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>1, 2</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
2 - Les résultats obtenus sur les variables tests

Pour juger de la qualité des échantillonnages, on s’intéresse essentiellement au rapport des variances entre 2 types de tirage. On note (cf. tableaux 1 à 4 en annexe).

TAB : effectif exact (tabulé)

MDISn : effectif moyen estimé à partir des échantillons contrôlés de districts au taux $1/n$

MDIS : effectif moyen estimé à partir des échantillons non contrôlés de districts (uniquement pour les régions 52 et 73)

MLOG : effectif moyen estimé à partir des échantillons systématiques de logements au taux 1/5.

On note VDISn, VDIS et VLOG les variances estimées correspondantes.

Dans un deuxième temps, on examine plus en détail les résultats de simulation pour les régions 83 et 31.

2.1. Résultats globaux

Remarques préliminaires :

1. Le faible échantillon de simulation (10 pour chaque tirage) ne permet pas de conclure de façon certaine quant à la précision des échantillonnages. En particulier, certaines "incohérences" peuvent apparaître entre les résultats du tirage au 1/10, au 1/5 et au 1/3.

Pour "lisser" les irrégularités provenant du faible nombre de tirages, on a calculé une efficacité moyenne, notée EFFMOY : par une moyenne pondérée des efficacités des 3 tirages contrôlés (VLOG/VDIS10, VLOG/VDIS5, VLOG/VDIS3) de coefficients (1/4, 1/2, 1/4) (l’objet principal de l’étude étant l’échantillonnage au taux 1/5, on lui accorde un poids plus important qu’aux 2 autres tirages).

2. La variabilité observée sur le nombre de logements dans le tirage systématique de logements au 1/5 résulte de la présence dans le fichier de logements fictifs. La variable intitulée "nombre de logements" dénombre en fait les logements non fictifs (CL = 1 à 7).
Pour évaluer les résultats, on distingue 3 groupes de variables :
(1) les variables de type moyenne ou taux
(2) les variables liées au nombre de logements
(3) les variables liées au nombre de BI.

Le critère utilisé est un critère global (variance des différents échantillonnages calculée à partir de 10 simulations).

(1) **Variables de type moyenne ou taux** :

Ces variables sont bien estimées par les différents types de tirages (différences non significatives entre valeurs estimées et valeur tabulée : cf. tableaux 1 et 2).

(2) **Variables liées au nombre de logements (L)** :

<table>
<thead>
<tr>
<th>(L)</th>
<th>Nombre de logements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nombre de résidences principales</td>
</tr>
<tr>
<td></td>
<td>Nombre de résidences secondaires</td>
</tr>
<tr>
<td></td>
<td>Nombre de logements récents</td>
</tr>
</tbody>
</table>

La variabilité des tirages contrôlés est nettement plus élevée que celle du tirage systématique de logements (cf. tableaux 3 et 4, lignes 1 à 4), et ceci bien que le nombre de FL soit l'une des variables de contrôle. Ceci s'explique pour les variables 1 et 2 (nombre de logements et nombre de résidences principales), d'une part par le fait d'arrêter le tirage au bout d'un nombre fixé de pas (un déséquilibre apparaissant en fin de procédure ne peut être résorbé), et d'autre part, par le fait qu'en négligeant un axe, même mineur, on perd une partie de l'information contenue dans FL. Pour les variables 3 et 4 (résidences secondaires et logements récents) sujettes à un fort effet de grappe, il est clair qu'un tirage de logements est plus approprié.

(3) **Variables liées au nombre de BI** :

On distingue les variables directement liées au nombre de BI (B1) de celles qui sont plus sujettes à un effet de grappe (B2).

<table>
<thead>
<tr>
<th>B1</th>
<th>Nombre de personnes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nombre d'hommes au total</td>
</tr>
<tr>
<td></td>
<td>Nombre d'hommes actifs</td>
</tr>
<tr>
<td></td>
<td>Nombre de femmes au total</td>
</tr>
<tr>
<td></td>
<td>Nombre de femmes actives</td>
</tr>
<tr>
<td></td>
<td>Nombre d'hommes de 20 à 24 ans, au total</td>
</tr>
</tbody>
</table>
Nombre d'hommes de 20 à 24 ans actifs  
Nombre de femmes de 30 à 34 ans, au total  
Nombre de femmes de 30 à 34 ans, actives  
Nombre de personnes de moins de 15 ans  
Nombre de personnes de 15 à 24 ans  
Nombre de personnes de 25 à 34 ans  
Nombre de personnes de 35 à 54 ans  
Nombre de personnes de 55 à 74 ans  
Nombre de personnes de 75 ans et plus

B2  Nombre d'étrangers, au total  
Nombre d'étrangers, actifs, hommes  
Nombre de personnes de CS = 3  
Nombre d'hommes de CS = 3  
Nombre de personnes de CS = 5  
Nombre d'hommes de CS = 5  
Nombre de personnes de CS = 6  
Nombre d'hommes de CS = 6  
Nombre d'inactifs  
Nombre d'hommes inactifs  
Nombre de personnes d'activité économique  
production de combustibles et minéraux solides  
chimie de base  
parachimie et industrie pharmaceutique  
commerce de détail

Les résultats détaillés figurent dans les tableaux 3 et 4. 
On donne ici quelques résultats plus synthétiques : pour les variables de type B1, on résume la distribution des efficacités sur l'ensemble des variables par les statistiques d'ordre X (3), X (8), X (13) ; X (8) s'interprète donc comme l'efficacité médiane et l'intervalle X (3), X (13) comme un indicateur de dispersion (idem pour le type B2 avec X (3), X (8) et X (12) ).

Tableau 2

<table>
<thead>
<tr>
<th>Région</th>
<th>31</th>
<th>83</th>
<th>52</th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>[X(3), X(13)]</td>
<td>[0,24 ; 0,93]</td>
<td>[0,44 ; 2,72]</td>
<td>[0,25 ; 1,01]</td>
</tr>
<tr>
<td></td>
<td>X(8)</td>
<td>0,66</td>
<td>1,00</td>
<td>0,54</td>
</tr>
<tr>
<td>B2</td>
<td>[X(3), X(12)]</td>
<td>[0,27 ; 0,78]</td>
<td>[0,17 ; 0,64]</td>
<td>[0,26 ; 1,06]</td>
</tr>
<tr>
<td></td>
<td>X(8)</td>
<td>0,51</td>
<td>0,42</td>
<td>0,41</td>
</tr>
</tbody>
</table>
Interprétation : Lorsque VLOG/VDIS5 <1, le tirage de logements est meilleur que le tirage systématique de districts et inversement quand ce rapport est supérieur à 1.

Globalement, on constate que :
- les variables de type B1 sont relativement mieux restituées que les variables de type B2, plus sujettes à un effet de grappe.
- les régions où le tirage de districts donne les meilleurs résultats par rapport au tirage de logements sont les régions 83 et 73, notamment pour les variables de type B1. Ceci peut s’expliquer par plusieurs raisons :
  + certaines strates des régions 52 et 31 se caractérisent par un pourcentage d’inertie sur le 3e axe (négligé) relativement important (cf. annexe : strate TU = 4 ou 5 en région 31, strate TUC = 64401 en région 52).
  + lorsque le tirage systématique de logements est lui-même sujet à une forte variabilité, le tirage de districts détériore relativement peu cette situation.
  + les effets de grappe jouent de façon différenciée selon les régions.

En résumé, sur les 4 régions considérées, la perte du tirage systématique de districts par rapport à celui de logements (au même taux : 1/5) varie entre 0,9 et 1,85 en termes de variance (valeurs médianes) pour les variables de type B1 et entre 1,95 et 2,45 pour les variables de type B2. (Une perte dans un rapport égal à 2 signifie qu’il faudrait doubler la taille de l’échantillon de districts pour obtenir une précision égale à un échantillon de logements).

Sur les régions 52 et 73, on a étudié plus spécifiquement ce que peut apporter la procédure de contrôle dans le tirage systématique de districts.

Globalement, l’échantillonnage contrôlé domine le tirage non contrôlé, surtout sur les variables directement liées à la taille de la population.
2.2. **Résultats détaillés des simulations (régions 31 et 83) :**

On examine ici de façon plus précise la distribution des simulations (au nombre de 10) pour chaque variable et chacun des 2 types d’échantillonnage :
- tirage systématique de districts au taux 1/5
- tirage systématique de logements au taux 1/5.

Le critère de type variance utilisé dans la partie A) apparaît en effet trop réducteur par certains aspects. On considère donc ici pour chaque variable de type L, B1 ou B2 la distribution des simulations résumée par les statistiques d’ordre 2, 5 et 8. On raisonne ici en termes de déviation relative par rapport à la valeur vraie (tabulée) :

Soit \( E_i(j) = \frac{|S_{ij} - TAB_i|}{TAB_i} \);

où \( S_{ij} \) désigne le résultat de la simulation \( j \)

\( TAB_i \) la vraie valeur

Les \( E_i(j) \) sont alors ordonnés et on retient comme indicateurs \( E_i^{(2)} \), \( E_i^{(5)} \) et \( E_i^{(8)} \) (cf. tableau 5 en annexe où ces valeurs sont exprimées en %).
### ANNEXES

(Résultats des ACP (par v. dem x strate))

<table>
<thead>
<tr>
<th>Région</th>
<th>Strate</th>
<th>Effectif</th>
<th>Inertie axe 1</th>
<th>Inertie axe 1-2</th>
<th>ACP axe 1</th>
<th>FR1 axe 1</th>
<th>ACP axe 2</th>
<th>FR1 axe 2</th>
<th>ACP axe 3</th>
<th>FR1 axe 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>03</td>
<td>TV = 0</td>
<td>2370</td>
<td>95,3</td>
<td>99,6</td>
<td>52</td>
<td>34</td>
<td>54</td>
<td>67</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>TV = 1, 2, 3</td>
<td>1007</td>
<td>92,9</td>
<td>96,6</td>
<td>55</td>
<td>34</td>
<td>53</td>
<td>67</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>TV = 4, 5</td>
<td>2547</td>
<td>95,7</td>
<td>97,0</td>
<td>56</td>
<td>35</td>
<td>29</td>
<td>14</td>
<td>15</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>TV = 7</td>
<td>1696</td>
<td>90,4</td>
<td>97,5</td>
<td>58</td>
<td>56</td>
<td>26</td>
<td>10</td>
<td>16</td>
<td>74</td>
</tr>
<tr>
<td>31</td>
<td>TV = 0, 1</td>
<td>2647</td>
<td>95,8</td>
<td>99,5</td>
<td>52</td>
<td>73</td>
<td>33</td>
<td>63</td>
<td>73</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TV = 2, 3</td>
<td>1463</td>
<td>95,7</td>
<td>98,4</td>
<td>53</td>
<td>34</td>
<td>33</td>
<td>63</td>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>TV = 4, 5</td>
<td>2803</td>
<td>94,6</td>
<td>94,0</td>
<td>52</td>
<td>35</td>
<td>33</td>
<td>61</td>
<td>2</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>TV = 6</td>
<td>2504</td>
<td>96,4</td>
<td>98,0</td>
<td>56</td>
<td>35</td>
<td>29</td>
<td>12</td>
<td>17</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>TVC = 70 201</td>
<td>3555</td>
<td>97,4</td>
<td>98,7</td>
<td>54</td>
<td>34</td>
<td>32</td>
<td>17</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>TVC = 75 902</td>
<td>5295</td>
<td>79,9</td>
<td>96,9</td>
<td>50</td>
<td>36</td>
<td>26</td>
<td>9</td>
<td>17</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>TVC = 75 901</td>
<td>1070</td>
<td>94,0</td>
<td>99,5</td>
<td>51</td>
<td>34</td>
<td>32</td>
<td>21</td>
<td>12</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>TVC = 70 052</td>
<td>1692</td>
<td>97,1</td>
<td>99,1</td>
<td>55</td>
<td>34</td>
<td>33</td>
<td>20</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>TVC = 70 051</td>
<td>1174</td>
<td>94,6</td>
<td>98,0</td>
<td>51</td>
<td>34</td>
<td>32</td>
<td>15</td>
<td>10</td>
<td>67</td>
</tr>
</tbody>
</table>

**Note** | TVC = (TV, HI, D) |
---|---|
| ISC : population sans double compte |
| FR : nombre de logements |
| A : nombre d'immobiliers |

L'échantillonnage 325
### Résultats région 83

<table>
<thead>
<tr>
<th></th>
<th>TAD</th>
<th>MD1510</th>
<th>MD155</th>
<th>MLOGS</th>
<th>VD1510</th>
<th>VD155</th>
<th>VD15S</th>
<th>VLOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 AGE MOYEN</td>
<td>38.18</td>
<td>58.030</td>
<td>58.149</td>
<td>58.067</td>
<td>38.154</td>
<td>0.130620</td>
<td>0.025099</td>
<td>0.052312</td>
</tr>
<tr>
<td>2 AGE MOYEN ET RANGERS</td>
<td>29.10</td>
<td>29.122</td>
<td>29.244</td>
<td>29.006</td>
<td>29.058</td>
<td>0.790483</td>
<td>0.258649</td>
<td>0.184385</td>
</tr>
<tr>
<td>3 LOGEMENTS POSTS</td>
<td>0.14</td>
<td>0.139</td>
<td>0.140</td>
<td>0.139</td>
<td>0.140</td>
<td>0.000032</td>
<td>0.000000</td>
<td>0.000054</td>
</tr>
<tr>
<td>4 Taux Activité Hommes</td>
<td>0.52</td>
<td>0.525</td>
<td>0.522</td>
<td>0.525</td>
<td>0.525</td>
<td>0.000025</td>
<td>0.000016</td>
<td>0.000025</td>
</tr>
<tr>
<td>5 Taux Activité Femmes</td>
<td>0.54</td>
<td>0.544</td>
<td>0.542</td>
<td>0.542</td>
<td>0.545</td>
<td>0.000027</td>
<td>0.000018</td>
<td>0.000018</td>
</tr>
<tr>
<td>6 Taux Act Hommes 20-24</td>
<td>0.72</td>
<td>0.717</td>
<td>0.722</td>
<td>0.717</td>
<td>0.724</td>
<td>0.001001</td>
<td>0.000129</td>
<td>0.000225</td>
</tr>
<tr>
<td>7 Taux Act Femmes 30-54</td>
<td>0.69</td>
<td>0.697</td>
<td>0.690</td>
<td>0.691</td>
<td>0.691</td>
<td>0.000179</td>
<td>0.000111</td>
<td>0.000054</td>
</tr>
<tr>
<td>8 Moins de 25 ans</td>
<td>0.54</td>
<td>0.544</td>
<td>0.542</td>
<td>0.544</td>
<td>0.542</td>
<td>0.000027</td>
<td>0.000018</td>
<td>0.000027</td>
</tr>
<tr>
<td>9 Plus de 75 ans</td>
<td>0.08</td>
<td>0.077</td>
<td>0.080</td>
<td>0.080</td>
<td>0.080</td>
<td>0.000025</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>10 % Sup/Actifs</td>
<td>0.05</td>
<td>0.035</td>
<td>0.050</td>
<td>0.055</td>
<td>0.050</td>
<td>0.000025</td>
<td>0.000000</td>
<td>0.000025</td>
</tr>
<tr>
<td>11 Ouvriers/Actifs</td>
<td>0.55</td>
<td>0.331</td>
<td>0.331</td>
<td>0.332</td>
<td>0.332</td>
<td>0.000077</td>
<td>0.000010</td>
<td>0.000040</td>
</tr>
<tr>
<td>12 Étrangers</td>
<td>0.05</td>
<td>0.047</td>
<td>0.047</td>
<td>0.049</td>
<td>0.048</td>
<td>0.000068</td>
<td>0.000025</td>
<td>0.000010</td>
</tr>
<tr>
<td>13 % Étran Act Hom/Actifs Hom</td>
<td>0.06</td>
<td>0.056</td>
<td>0.057</td>
<td>0.056</td>
<td>0.056</td>
<td>0.000071</td>
<td>0.000023</td>
<td>0.000018</td>
</tr>
<tr>
<td></td>
<td>RESULTATS REGION 31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------------------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TABLE MDIS10 MDIS5 MDIS5 VLOG5 VDIS10 VDIS5 VDIS5 VLOG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>AGE MOYEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>33.94</td>
<td>33.861</td>
<td>54.059</td>
<td>53.945</td>
<td>0.038255</td>
<td>0.002366</td>
<td>0.0088992</td>
<td>0.008028</td>
</tr>
<tr>
<td>3</td>
<td>AGE MOYEN ÉTRANGERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>33.44</td>
<td>53.240</td>
<td>35.623</td>
<td>33.409</td>
<td>33.323</td>
<td>0.440489</td>
<td>0.133088</td>
<td>0.121934</td>
</tr>
<tr>
<td>5</td>
<td>ÉLÉMENTS POST75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.14</td>
<td>0.144</td>
<td>0.140</td>
<td>0.159</td>
<td>0.140</td>
<td>0.00049</td>
<td>0.000000</td>
<td>0.0000100</td>
</tr>
<tr>
<td>7</td>
<td>TAUX ACTIVITÉ HOMMES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.49</td>
<td>0.490</td>
<td>0.491</td>
<td>0.491</td>
<td>0.491</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.0000100</td>
</tr>
<tr>
<td>9</td>
<td>TAUX ACTIVITÉ FEMMES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.29</td>
<td>0.289</td>
<td>0.290</td>
<td>0.290</td>
<td>0.290</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.0000100</td>
</tr>
<tr>
<td>11</td>
<td>TAUX ACT HOMMES 20-24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.72</td>
<td>0.716</td>
<td>0.717</td>
<td>0.719</td>
<td>0.714</td>
<td>0.000249</td>
<td>0.000025</td>
<td>0.0000544</td>
</tr>
<tr>
<td>13</td>
<td>TAUX ACT FEMMES 30-34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.58</td>
<td>0.573</td>
<td>0.578</td>
<td>0.573</td>
<td>0.579</td>
<td>0.000068</td>
<td>0.000018</td>
<td>0.0000233</td>
</tr>
<tr>
<td>15</td>
<td>MOINS DE 25 ANS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.41</td>
<td>0.410</td>
<td>0.410</td>
<td>0.410</td>
<td>0.410</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>17</td>
<td>PLUS DE 75 ANS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.05</td>
<td>0.051</td>
<td>0.051</td>
<td>0.050</td>
<td>0.050</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>19</td>
<td>CADAUP/ACTIFS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.06</td>
<td>0.060</td>
<td>0.060</td>
<td>0.060</td>
<td>0.060</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>21</td>
<td>OUVRIERS/ACTIFS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.41</td>
<td>0.415</td>
<td>0.411</td>
<td>0.414</td>
<td>0.414</td>
<td>0.000050</td>
<td>0.0000100</td>
<td>0.0000277</td>
</tr>
<tr>
<td>23</td>
<td>ÉTRANGERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.05</td>
<td>0.050</td>
<td>0.050</td>
<td>0.050</td>
<td>0.050</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>25</td>
<td>ÉTRANGERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0.06</td>
<td>0.058</td>
<td>0.060</td>
<td>0.059</td>
<td>0.060</td>
<td>0.000040</td>
<td>0.000000</td>
<td>0.000100</td>
</tr>
<tr>
<td>Nom de l'article</td>
<td>TD1510</td>
<td>TD1515</td>
<td>MO5</td>
<td>VLOG VLOG/VD150</td>
<td>VLOG/VD155</td>
<td>VLOG/VD155 EFG10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>-----------------</td>
<td>-------------</td>
<td>-------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NB ET PLUS</td>
<td>105416</td>
<td>103589</td>
<td>10581</td>
<td>103465</td>
<td>101247</td>
<td>580019</td>
<td>1.09157</td>
<td>0.59209</td>
</tr>
<tr>
<td>JS ET PLUS</td>
<td>105416</td>
<td>103589</td>
<td>10581</td>
<td>103465</td>
<td>101247</td>
<td>580019</td>
<td>1.09157</td>
<td>0.59209</td>
</tr>
<tr>
<td>INACTIFS</td>
<td>510712</td>
<td>512459</td>
<td>51115</td>
<td>512164</td>
<td>508662</td>
<td>650464</td>
<td>1.86684</td>
<td>0.68548</td>
</tr>
<tr>
<td>PARACHIMIE</td>
<td>31356</td>
<td>32368</td>
<td>31571</td>
<td>312451</td>
<td>312451</td>
<td>312451</td>
<td>312451</td>
<td>312451</td>
</tr>
<tr>
<td>COMBUST/MINERAUX</td>
<td>584</td>
<td>440</td>
<td>619</td>
<td>428</td>
<td>584</td>
<td>11004</td>
<td>0.20786</td>
<td>0.19029</td>
</tr>
<tr>
<td>CHIMIE DE BASE</td>
<td>2500</td>
<td>2796</td>
<td>2661</td>
<td>428</td>
<td>584</td>
<td>11004</td>
<td>0.20786</td>
<td>0.19029</td>
</tr>
<tr>
<td>VLOG VLOG/VD150</td>
<td>510712</td>
<td>512459</td>
<td>51115</td>
<td>512164</td>
<td>508662</td>
<td>650464</td>
<td>1.86684</td>
<td>0.68548</td>
</tr>
<tr>
<td>VLOG VLOG/VD155</td>
<td>510712</td>
<td>512459</td>
<td>51115</td>
<td>512164</td>
<td>508662</td>
<td>650464</td>
<td>1.86684</td>
<td>0.68548</td>
</tr>
<tr>
<td>VLOG/VD155 EFG10</td>
<td>510712</td>
<td>512459</td>
<td>51115</td>
<td>512164</td>
<td>508662</td>
<td>650464</td>
<td>1.86684</td>
<td>0.68548</td>
</tr>
</tbody>
</table>
### Tableau 4

#### RÉSULTATS REGION 51

<table>
<thead>
<tr>
<th>TAB</th>
<th>MBLS10</th>
<th>MBLS3</th>
<th>MBLS5</th>
<th>MLOGS</th>
<th>VLOG VLOG/VDLS10 VLOG/VDLS3 VLOG/VDLS EFFROI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1466900</td>
<td>1471258</td>
<td>1471206</td>
<td>1471228</td>
<td>1466822</td>
</tr>
<tr>
<td>2</td>
<td>517072</td>
<td>512045</td>
<td>512186</td>
<td>512206</td>
<td>512182</td>
</tr>
<tr>
<td>3</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
</tr>
<tr>
<td>4</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
</tr>
<tr>
<td>5</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
<td>510570</td>
</tr>
</tbody>
</table>

#### PARACHIMIE PHARMACIE

<table>
<thead>
<tr>
<th>HOMMES</th>
<th>CS°5</th>
</tr>
</thead>
<tbody>
<tr>
<td>81108</td>
<td>118728</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES ACTIFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>166488</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES 20-24 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>111908</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FEMMES ACTIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>152400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FEMMES 30-34 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>88128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MOINS DE 15 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>355470</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15-24 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>672512</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>25-34 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>621528</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>35-44 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>85976</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>45-54 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>628984</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>55-74 ANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>207697</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>75 ET PLUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>192538</td>
</tr>
</tbody>
</table>

#### ÉTRANGERS

<table>
<thead>
<tr>
<th>ÉTRANGERS ACTIFS HOMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>54172</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES CS°5</th>
</tr>
</thead>
<tbody>
<tr>
<td>109050</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES CS°3</th>
</tr>
</thead>
<tbody>
<tr>
<td>197080</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES CS°</th>
</tr>
</thead>
<tbody>
<tr>
<td>556726</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES CS°5</th>
</tr>
</thead>
<tbody>
<tr>
<td>105328</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CS°6</th>
</tr>
</thead>
<tbody>
<tr>
<td>650354</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOMMES CS°6</th>
</tr>
</thead>
<tbody>
<tr>
<td>492626</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INACTIFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2403240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INACTIFS HOMMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>973558</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMBUST/MINERAUX SOLIDES</th>
</tr>
</thead>
<tbody>
<tr>
<td>21884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMMERCE DE BASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>12872</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARACHIMIE PHARMACIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>9240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMMERCE DE DETAIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>118884</td>
</tr>
</tbody>
</table>

#### 17:55 TUESDAY, JUNE 24, 1986
Tableau 5

<table>
<thead>
<tr>
<th>Région 93</th>
<th>Region 34</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>NAME</em></td>
<td>QD1</td>
</tr>
<tr>
<td>W1</td>
<td>0.17</td>
</tr>
<tr>
<td>W2</td>
<td>0.02</td>
</tr>
<tr>
<td>W3</td>
<td>0.48</td>
</tr>
<tr>
<td>W4</td>
<td>0.72</td>
</tr>
<tr>
<td>W5</td>
<td>0.20</td>
</tr>
<tr>
<td>W6</td>
<td>0.61</td>
</tr>
<tr>
<td>W7</td>
<td>0.27</td>
</tr>
<tr>
<td>W8</td>
<td>0.24</td>
</tr>
<tr>
<td>W9</td>
<td>0.55</td>
</tr>
<tr>
<td>W10</td>
<td>1.84</td>
</tr>
<tr>
<td>W11</td>
<td>0.95</td>
</tr>
<tr>
<td>W12</td>
<td>0.86</td>
</tr>
<tr>
<td>W13</td>
<td>0.70</td>
</tr>
<tr>
<td>W14</td>
<td>0.27</td>
</tr>
<tr>
<td>W15</td>
<td>0.56</td>
</tr>
<tr>
<td>W16</td>
<td>0.58</td>
</tr>
<tr>
<td>W17</td>
<td>0.22</td>
</tr>
<tr>
<td>W18</td>
<td>0.51</td>
</tr>
<tr>
<td>W19</td>
<td>0.88</td>
</tr>
<tr>
<td>W20</td>
<td>2.74</td>
</tr>
<tr>
<td>W21</td>
<td>0.63</td>
</tr>
<tr>
<td>W22</td>
<td>1.28</td>
</tr>
<tr>
<td>W23</td>
<td>2.58</td>
</tr>
<tr>
<td>W24</td>
<td>0.56</td>
</tr>
<tr>
<td>W25</td>
<td>1.28</td>
</tr>
<tr>
<td>W26</td>
<td>0.78</td>
</tr>
<tr>
<td>W27</td>
<td>0.58</td>
</tr>
<tr>
<td>W28</td>
<td>0.20</td>
</tr>
<tr>
<td>W29</td>
<td>0.15</td>
</tr>
<tr>
<td>W30</td>
<td>24.66</td>
</tr>
<tr>
<td>W31</td>
<td>4.80</td>
</tr>
<tr>
<td>W32</td>
<td>2.41</td>
</tr>
<tr>
<td>W33</td>
<td>2.89</td>
</tr>
</tbody>
</table>

Quantiles d'ordre 2, 5, 8 de la variable K1 (exprimée en %)
QD : échantillonnage de districts
QL : échantillonnage de logements

INSEE Méthodes n° 52 - 53
Graphique 1:

L'échantillonnage
Graphique 2

Quantiles relatifs aux variables $B$
Graphique 3

*Quantiles relatifs aux variables B2*

L'échantillonnage
4.4- ÉCHANTILLONNAGE CONTRÔLÉ DE DISTRICTS POUR LE TIRAGE DU 1/20 DU RECENSEMENT

Nicole Roth

1- Rappel de la méthode :

On cherche à équilibrer un échantillon de taille \( n \) sur \( k \) critères, c'est-à-dire un échantillon respectant \( x_k = \bar{X}_k \) pour \( k = 1, \ldots, K \), ceci tout en respectant l'équiprobabilité du tirage. On constitue alors \( 2^K \) groupes indiqués par \( \sigma \) selon les positions relatives de \( x_{ik}^{\sigma} \) par rapport à \( X^k \). Ainsi par exemple, pour \( K = 2 \), on a 4 groupes qu'on peut noter \( ++, + -, - + \) et \( -- \) :

\[
\begin{array}{c|cc}
\sigma & \text{++} & \text{+ -} \\
\hline
\bar{x}_1^2 & \text{+ -} & \text{+ +} \\
\end{array}
\]

Un algorithme simple et naturel consiste à calculer à l'étape \( v \) du tirage les sommes partielles \( \sum_{i=1}^{v} x_{ik}^{\sigma} \), et à tirer à l'étape \( v + 1 \) dans le quadrant "opposé" pour aller dans le sens du rééquilibrage (le premier individu étant tiré au hasard). Cette procédure assure des probabilités quasi-égales quand \( K = 1 \), mais pose quelques problèmes dès qu'on a un critère multivarié ; en effet, on n'assure pas en général l'équiprobabilité de tirage des différents groupes \( \sigma \). Pour pallier cet inconvénient, un second algorithme est proposé (c'est celui qui a été utilisé dans les simulations présentées ci-dessous). On procède de la façon suivante :

- 1ère itération : elle consiste à faire un tirage stratifié avec allocation proportionnelle selon la variable appartenance aux groupes \( \sigma \) (on alloue \( n_{\sigma} = np_{\sigma} \), où \( p_{\sigma} \) désigne le
poids de la strate $\sigma$). En négligeant les problèmes d'arrondis, on obtient ainsi un échantillon $E_0$ et on peut calculer le $K$-critère : 

$$U_0 = \sum_\sigma n_0^\sigma \bar{x}_\sigma^{(K,1)}$$

- 2nde itération : on cherche à modifier à la marge l'échantillon initial $E_0$ de façon à rééquilibrer vers la vraie moyenne $\bar{X}$ (qu'on peut supposer nulle pour simplifier). On cherche donc une allocation $\nu = \{\nu^\sigma\}$ assurant en moyenne un rééquilibrage, sous contrainte d'effectif global constant, connaissant la réalisation de l'échantillon initial, soit : on cherche $\nu$ tel que

$$\sum \nu^\sigma \bar{x}^\sigma = -U_0 \quad \text{s.c.} \quad \sum \nu^\sigma = 0$$

Ce problème admet plusieurs solutions ; on retient celle qui résulte d'un certain critère d'optimisation, consistant à s'écarteler le moins possible de l'allocation proportionnelle.

On minimise donc la distance $\left(\chi^2\right)$ entre $p^\sigma$ et $\hat{p}^\sigma = n_0^\sigma + \nu^\sigma \frac{n}{n}$ ce qui conduit au critère quadratique : 

$$\min \sum \frac{(\nu^\sigma)^2}{p^\sigma}$$

Si on note : $P = \text{diag} \left( p^1, p^2, \ldots, p^{2K} \right)$

$$\bar{X}^{(K+1,2K)} = \begin{pmatrix} \bar{X}^1 \\ \vdots \\ \bar{X}^1 \\ \bar{X}^1 \\ \vdots \\ \bar{X}^K \\ 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} U_0 \\ 0 \end{pmatrix} = \begin{pmatrix} U_0^1 \\ U_0^K \end{pmatrix}$$

le problème s'écrit alors sous forme matricielle : 

$$\min \quad \nu \quad \text{s.c.} \\ \bar{X} \nu = -\bar{U}_0$$

d'où : 

$$\nu = -P^{-1} \bar{X}, \quad (\bar{X} P \bar{X})^{-1} \bar{U}_0 = RU\bar{U}_0$$

(aux arrondis près)

- Cette étape peut-être itérée jusqu'à ce que $\nu = 0$ ou, en pratique, jusqu'à vérification d'un critère d'arrêt. On montre, de façon théorique, qu'on respecte ainsi l'égalité
(ou presque) des probabilités d’inclusion (ce qui peut être vérifié de façon empirique) et qu’on converge vers la vraie moyenne $\bar{X}$.

Remarque : L’algorithme utilisé ici permet apparemment de prendre en compte un nombre de critères plus élevé que l’algorithme alterné où le non-respect des $p^\circ$ peut conduire souvent à un échec du tirage (groupe épuisé). Cependant, la multiplication du nombre de groupes dilue le contrôle sur chacun des critères pris séparément, ce qui peut ne pas être acceptable (cf. composantes principales). Par ailleurs, pour des groupes de petite taille, le problème des arrondis devient alors crucial ; la gestion aléatoire des arrondis masque alors l’effet du contrôle.

2 - Comparaison de 2 méthodes d’échantillonnage contrôlé

L’algorithme décrit ci-dessus a été utilisé afin de tester 2 méthodes de tirage contrôlé du 1/20 à partir du fichier 1/4. Ces deux méthodes diffèrent par l’information auxiliaire supposée connue :

- 1ère méthode (notée G) : on connaît pour chaque district les 3 variables de taille
  - $C_1$ : nombre de BI
  - $C_2$ : nombre de FL
  - $C_3$ : nombre d’immeubles

dont on extrait deux composantes principales par une A C P (ces deux composantes expliquent environ 95 % d’inertie).

- 2ème méthode (notée H) : outre ces variables de taille, on connaît 8 autres grandeurs :
  - $C_4$ : nombre de personnes de moins de 20 ans
  - $C_5$ : nombre de personnes de plus de 60 ans
  - $C_6$ : nombre de femmes actives
  - $C_7$ : nombre d’ouvriers et paysans ($CS_8 = 1,6$)
  - $C_8$ : nombre d’artisans, commerçants, cadres et professions intermédiaires ($CS_8 = 2, 3, 4$)
  - $C_9$ : nombre d’étrangers
  - $C_{10}$ : nombre de logements postérieurs à 1975
  - $C_{11}$ : nombre de logements antérieurs à 1948

On extrait 3 composantes principales à partir d’une A C P sur les 11 variables ainsi définies.

Les 4 régions suivantes ont été étudiées :

1 Auvergne (83)
2 Nord-Pas de Calais (31)
3 Pays de Loire (52)
4 Midi-Pyrenées (73)
Dans tous les cas, on procède à une stratification préalable selon la tranche d'unité urbaine ; on rappelle que le fichier utilisé est un fichier de districts issu du 1/4, où on a regroupé les districts des communes de petite taille (TCD<1 000).

L'efficacité du mode de tirage a été testée sur 36 variables dont la liste suit :

- W1) Nombre total de logements
- W2) Nombre de résidences principales
- W3) Nombre de logements vacants et résidences secondaires
- W4) Nombre de logements récents (1975 ou après)
- W5) Nombre de personnes
- W6) Nombre d'hommes au total
- W7) Nombre d'hommes actifs
- W8) Nombre de femmes au total
- W9) Nombre de femmes actives
- W10) Nombre d'hommes de 20 à 24 ans, au total
- W11) Nombre d'hommes de 20 à 24 ans actifs
- W12) Nombre de femmes de 30 à 34 ans, au total
- W13) Nombre de femmes de 30 à 34 ans, actives
- W14) Nombre de personnes de moins de 15 ans
- W15) Nombre de personnes de 15 à 24 ans
- W16) Nombre de personnes de 24 à 34 ans
- W17) Nombre de personnes de 35 à 54 ans
- W18) Nombre de personnes de 55 à 74 ans
- W19) Nombre de personnes de 75 et plus
- W20) Nombre d'étrangers, au total
- W21) Nombre d'étrangers, actifs, hommes

Nombre de personnes d'activité économique :

- W22) Production de combustibles et minéraux solides
- W23) Chimie de base
- W24) Parachimie et industrie pharmaceutique
- W25) Commerce de détail
- W26) Primaire
- W27) Industrie
- W28) Tertiaire
- W29) Agriculteurs exploitants
- W30) Artisans, commerçants, chefs d'entreprise
- W31) Cadres
- W32) Professions intermédiaires
- W33) Employés
- W34) Ouvriers
- W35) Retraités
- W36) Autres sans activité professionnelle

INSEE Méthodes n° 52 - 53
Les résultats sont présentés essentiellement sous deux formes :

1) Les graphes représentant les rapports des écarts-type des 2 méthodes :

$$R_{AHPG} = \frac{\sigma_H}{\sigma_G}$$

chacun des écarts-type étant calculé à partir de 15 simulations différentes, et ceci pour chacune des 36 variables-test (cf. tableaux 1 et 2, graphiques 1 et 2).

Remarque 1 : Sur ces graphiques, les variables ayant servi plus ou moins directement au contrôle H sont marquées d’un 0, à savoir :

\[ W1 = C2 \]
\[ W4 = C10 \]
\[ W5 = C1 \]
\[ W9 = C6 \]
\[ W20 = C9 \]
\[ W29 + W34 = C7 \]
\[ W30 + W31 + W32 = C8 \]

Remarque 2 : Pour les graphes 1 et 2, les ordonnées sont graduées en échelle logarithmique afin de donner la même importance visuelle à un gain ou à une perte de même importance. Un rapport $R_{AHPG} < 1$ correspond au cas où H est plus efficace que G et inversement.

Remarque triviale mais importante : On peut raisonner en gain en écart-type ou en variance ; un raisonnement en écart-type se réfère à une précision en termes d’intervalle de confiance (± 2 fois l’écart-type) ; utiliser le gain en variance revient à raisonner en termes de taille d’échantillon équivalente (pour obtenir la même précision). Ainsi un gain relatif de 0,7 en écart-type, c’est-à-dire un gain de 0,49 en variance, signifie qu’on obtient un résultat aussi précis avec un échantillon réduit de moitié.

On donne également, pour la région 83 seulement, le graphique représentant l’efficacité d’un contrôle sur 4 facteurs issus de C1-C11 (méthode L) par rapport à un contrôle sur 3 facteurs issus de la même ACP (méthode H) : cf. graphique 3.

Résultats globaux : comptage des efficacités supérieur ou inférieure à 1

<table>
<thead>
<tr>
<th></th>
<th>$R_{AHPG} &lt; 1$</th>
<th>$R_{AHPG} \geq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 83$</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>$R = 31$</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>$R = 52$</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>$R = 73$</td>
<td>18</td>
<td>18</td>
</tr>
</tbody>
</table>

Globalement, sauf pour la région Midi-Pyrénées, on observe donc un gain de H par rapport à G, et ceci d’autant plus que H permet de gagner quantitativement plus sur certaines variables qu’il ne fait perdre sur d’autres.

L’échantillonnage 339
Il reste donc à détailler, variable par variable, l’efficacité relative de H par rapport à G ;

pour cela on considère six groupes :

- groupe 1 : W1 à W4 ⟷ logement
- groupe 2 : W5 à W13 ⟷ sexe x activité
- groupe 3 : W14 à W19 ⟷ tranches d’âge
- groupe 4 : W20 à W21 ⟷ étrangers
- groupe 5 : W22 à W28 ⟷ activité économique
- groupe 6 : W29 à W36 ⟷ catégorie socio-professionnelle

1. **groupe 1** : les résultats sont globalement meilleurs pour G que pour H, sauf pour la variable W4, qui est introduite comme variable de contrôle dans l’échantillonnage H (W4 = C10) et dont la contribution au 3ème facteur de l’ACP est relativement forte pour l’ensemble des strates x régions.

2. **groupe 4** : c’est sur ces variables que le critère du coefficient de corrélation multiple est le plus amélioré en passant de G à H. De fait, on observe un gain important d’efficacité pour ces 2 variables.

3. **groupe 5** : ce groupe est constitué de variables sujettes à un fort effet de grappe : W22 à W24 et les résultats sont assez mauvais quelle que soit la méthode employée. L’efficacité H/G varie de façon assez erratique sans que cela puisse être considéré comme significatif.

Pour les autres variables du groupe, H améliore souvent par rapport à G, sauf dans le cas de W27.

4. **groupes 2, 3 et 6** : de même que pour le nombre de logements, la méthode H a tendance à détériorer les résultats pour les variables liées directement à un effet de taille : W5, W6 et W8. Pour les autres variables, on observe en moyenne des résultats meilleurs pour H, mais les résultats sont variables d’une région à l’autre (on ne gagne pas systématiquement sur les mêmes variables).

**Conclusion** : Sur l’ensemble des régions, il n’y a donc pas de résultat général et systématique, sauf pour les variables directement liées à un effet taille, privilégiées par la méthode G, et quelques variables (W4, W7, W19, W20, W26 et W35) qui sont mieux contrôlées par H que par G. Pour les autres, on observe des "sautillements" selon les régions.

Les résultats relatifs aux 4 régions sont récapitulés sur le graphique 4 qui représente la moyenne (simple) des efficacités relatives.
Mini étude sur le nombre de facteurs de contrôle.

Le rajout d’un critère de contrôle (qui revient à multiplier le nombre de groupes) peut avoir un effet pervers dû à une certaine dilution du contrôle sur chacun des critères pris séparément.

Ainsi, quand on passe de 3 à 4 facteurs de contrôle issus de C1-C11 pour la région 83, le gain en corrélation multiple ne se retrouve pas en termes d’efficacité de l’échantillonnage L par rapport à H ; les 4 composantes principales étant traitées lors du contrôle sur un pied d’égalité, le gain sur le 4e facteur peut s’accompagner d’une détérioration importante sur le 1er (qui explique à lui seul de l’ordre de 70 % d’inertie).

La multiplication extrême des facteurs de contrôle ne paraît donc pas souhaitable (cf. propriétés théoriques de l’algorithme : pour obtenir la convergence vers $X$, il faut avoir une taille suffisante de l’échantillon dans chaque groupe).

Un "optimum" empirique paraît être, dans le cas étudié, le contrôle sur 3 facteurs.

3. Comparaison détaillée de divers échantillonnages et estimateurs

Pour la région 83, il a été procédé à une étude détaillée de 9 stratégies d’échantillonnage et d’estimation notées de la façon suivante :

- échantillonnage contrôlé
  - G : tirage contrôlé sur 2 facteurs issus de C1-C3
  - H : tirage contrôlé sur 3 facteurs issus de C1-C11
  - L : tirage contrôlé sur 4 facteurs issus de C1-C11

- échantillonnage "classique"
  - I : tirage stratifié par groupe $\sigma$ (= 1 à 4) avec allocation proportionnelle
  - J : tirage systématique avec tri préalable sur un critère taille (1er facteur issu de C1-C3)
  - K : tirage aléatoire simple

- estimateur par la régression sur fac1 à fac3(issus de C1-C11)
  - R : estimateur par la régression sur le fichier K
  - T : estimateur par la régression sur le fichier I
  - V : estimateur par la régression sur le fichier H

Les méthodes I et J correspondent à des améliorations "classiques" de tirage aléatoire simple (K) ; l’échantillonnage J avec tri préalable sur un critère taille permet de bénéficier d’une stratification implicite. La méthode I permet de bénéficier quant à elle d’une stratification fine qui correspond en fait à la 1ère itération de G. Quant aux trois
dernières méthodes, elles combinent une stratégie d'échantillonnage (K, I, H) et une stratégie d'estimation "complexe" : estimation par la régression (cf. SARNDAL) qu'on rappelle rapidement :

Soit U l'univers 
S l'échantillon.

On suppose qu'on peut poser le modèle :

\[ \begin{align*}
E_m (Y_k) &= x_k \beta \\
V_m (Y_k) &= v_k
\end{align*} \]

où Em, Vm désignent l'espérance et la variance par rapport au modèle.

Le plan d'échantillonnage se caractérise quant à lui par les probabilités d'inclusion au 1er et au 2nd ordre : \{\pi_k, \pi_{ke}\}

Ayant observé \((Y_k, x_k)\) pour \(k \in S\), on effectue un ajustement empirique pour estimer \(\beta\).

Le critère d'optimisation \(\sum_{S} (Y_k - x_k \beta)^2 / v_k \pi_k\) est un estimateur sans biais du critère défini sur U, à savoir : \(\sum_{U} (Y_k - x_k \beta)^2 / v_k\),

d'où : \(\hat{\beta} = \left( \sum x_k x_k' / V_k \pi_k \right)^{-1} \sum x_k y_k / V_k \pi_k\)

On peut alors estimer \(Y_k\) pour \(k \in U \backslash S\) par \(\hat{Y}_k = x_k \hat{\beta}\)

Posant \(r_k = Y_k - \hat{Y}_k\) pour \(K \in S\), SARNDAL définit ainsi l'estimateur par la régression multiple : \(\hat{Y}_r = \sum_{U} \hat{Y}_k + \sum_{S} r_k / \pi_k\)

(Remarque : cet estimateur est du type "design-based" car les résidus estimés sont dilatés selon le plan d'échantillonnage ; d'autres estimateurs par la régression sont envisageables : par exemple l’approche de ROYALL-HERSON consiste à poser

\(\hat{Y}_{RH} = \sum_{U} \hat{Y}_k + \sum_{S} r_k = \sum_{U \backslash S} \hat{Y}_k + \sum_{S} Y_k\)

Dans le cas qui nous intéresse ici, les \(\pi_R\) sont égaux et on fait une régression avec constante ; les deux estimateurs coïncident donc).
Le graphique 5 présente les principaux résultats, à savoir :

1) Pour les méthodes d'échantillonnage contrôlé, la meilleure méthode semble être celle reposant sur les 3 facteurs (H : cf. partie 2).

2) Les méthodes classiques I et J améliorent les résultats par rapport à l'échantillonnage aléatoire simple mais pas de façon décisive. Comparées à l'échantillonnage contrôlé H, elles donnent toujours des résultats nettement moins bons.

3) L'estimateur par la régression permet d'améliorer assez notablement les résultats par rapport à un estimateur simple (valeurs dilatées) sur le même échantillon. Si on compare l'effet du contrôle seul (H) à l'effet de l'estimateur par la régression sur le fichier issu d'un tirage aléatoire simple (R), les résultats sont nettement favorables à H mais cette supériorité peut s'expliquer par la 1ère étape de stratification mise en œuvre dans H.

Remarque : l'estimateur par la régression combiné avec l'échantillonnage contrôlé (sur les mêmes trois facteurs) ne s'avère pas redondant. L'estimation par la régression redresse en effet certaines imperfections du tirage contrôlé tenant en particulier au problème des arrondis ou à une certaine dilution du contrôle sur certains facteurs.

Par contre, si on utilise un estimateur par la régression, le contrôle sur les mêmes variables auxiliaires paraît superflu ; une stratification en groupes suffisamment fine semble suffisante.

Conclusion : Une stratégie combinant échantillonnage et estimateur paraît donc bénéfique, (elle resterait cependant à tester sur une autre région). Il faut cependant souligner l'inconvénient (majeur ?) d'une méthode d’estimation non linéaire, à savoir sa mise en œuvre pratique qui peut s'avérer délicate. La stratégie (échantillonnage contrôlé + estimateur simple) ne souffre évidemment pas de ce défaut.
### Tableau 1

<table>
<thead>
<tr>
<th>CBS</th>
<th>NOM</th>
<th>MHT TABB</th>
<th>MHT G</th>
<th>MHT STG</th>
<th>MHT FH</th>
<th>STDH</th>
<th>FRPHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>123 657</td>
<td>89 012 34</td>
<td>56 78 9</td>
<td>01 234 56</td>
<td>78 90 12</td>
<td>345 67</td>
<td>89 01</td>
</tr>
</tbody>
</table>

### Tableau 2

<table>
<thead>
<tr>
<th>CBS</th>
<th>NOM</th>
<th>MHT TABB</th>
<th>MHT G</th>
<th>MHT STG</th>
<th>MHT FH</th>
<th>STDH</th>
<th>FRPHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>123 657</td>
<td>89 012 34</td>
<td>56 78 9</td>
<td>01 234 56</td>
<td>78 90 12</td>
<td>345 67</td>
<td>89 01</td>
</tr>
</tbody>
</table>

### Comparison G,H

### Région 83

<table>
<thead>
<tr>
<th>CBS</th>
<th>NOM</th>
<th>MHT TABB</th>
<th>MHT G</th>
<th>MHT STG</th>
<th>MHT FH</th>
<th>STDH</th>
<th>FRPHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>123 657</td>
<td>89 012 34</td>
<td>56 78 9</td>
<td>01 234 56</td>
<td>78 90 12</td>
<td>345 67</td>
<td>89 01</td>
</tr>
</tbody>
</table>

### ANNEXES

<table>
<thead>
<tr>
<th>CBS</th>
<th>NOM</th>
<th>MHT TABB</th>
<th>MHT G</th>
<th>MHT STG</th>
<th>MHT FH</th>
<th>STDH</th>
<th>FRPHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>123 657</td>
<td>89 012 34</td>
<td>56 78 9</td>
<td>01 234 56</td>
<td>78 90 12</td>
<td>345 67</td>
<td>89 01</td>
</tr>
</tbody>
</table>

INSEE Méthodes n° 52 - 53
### Tableau 3

**Comparaison G,H**

**Région 52**

<table>
<thead>
<tr>
<th>Tableau 3</th>
<th>G</th>
<th>H</th>
<th>STOG</th>
<th>STOH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>125</td>
<td>126</td>
<td>354</td>
<td>354</td>
</tr>
<tr>
<td>2</td>
<td>127</td>
<td>128</td>
<td>357</td>
<td>357</td>
</tr>
<tr>
<td>3</td>
<td>129</td>
<td>130</td>
<td>360</td>
<td>360</td>
</tr>
<tr>
<td>4</td>
<td>131</td>
<td>132</td>
<td>363</td>
<td>363</td>
</tr>
<tr>
<td>5</td>
<td>133</td>
<td>134</td>
<td>366</td>
<td>366</td>
</tr>
<tr>
<td>6</td>
<td>135</td>
<td>136</td>
<td>369</td>
<td>369</td>
</tr>
<tr>
<td>7</td>
<td>137</td>
<td>138</td>
<td>372</td>
<td>372</td>
</tr>
<tr>
<td>8</td>
<td>139</td>
<td>140</td>
<td>375</td>
<td>375</td>
</tr>
</tbody>
</table>

**Comparaison G,H**

**Région 73**

<table>
<thead>
<tr>
<th>Tableau 4</th>
<th>G</th>
<th>H</th>
<th>STOG</th>
<th>STOH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>141</td>
<td>142</td>
<td>378</td>
<td>378</td>
</tr>
<tr>
<td>2</td>
<td>143</td>
<td>144</td>
<td>381</td>
<td>381</td>
</tr>
<tr>
<td>3</td>
<td>145</td>
<td>146</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>4</td>
<td>147</td>
<td>148</td>
<td>387</td>
<td>387</td>
</tr>
<tr>
<td>5</td>
<td>149</td>
<td>150</td>
<td>390</td>
<td>390</td>
</tr>
<tr>
<td>6</td>
<td>151</td>
<td>152</td>
<td>393</td>
<td>393</td>
</tr>
<tr>
<td>7</td>
<td>153</td>
<td>154</td>
<td>396</td>
<td>396</td>
</tr>
<tr>
<td>8</td>
<td>155</td>
<td>156</td>
<td>399</td>
<td>399</td>
</tr>
</tbody>
</table>

L'échantillonnage

345
Région 83

OBS 908
1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8
9
20
21
22
25
24
25
26
27
28
29
30
31
32
35
54
35
36

346

MOTTA8

0986

sTDD

sTol

sTDJ

5763.2
635348 3868.48 4934.76
9450.2
w 1
7157.4 3274.2
485976 2545.91 3300.84
1/ 2
5443.4
4436.4
151372 2015.72 4788.44
45
4260.8
2626.8
1/
88264 3067.54 1547.93
1./ 5 1332848 7664.52 6951.53 21467.5 14048.4
6644.5
651660 4671.17 3529.27 10684.3
1/ 6
5241.2
6484.8
11 7 340948 2634.48 1999.04
7920.6
681188 3488.71 3768.87 11276.9
11 8
3715.4
4384.5
233104 2328.78 2635.90
1.! 9
50136 1491.27 1707.66 1321.3 1410.7
1/10
956.7
1026.4
968.55
868.37
411
56056
840.13
904.42
1214.4 1381.7
49020
1412
1135.7
752.8
33824
651.83 863.68
413
5566.6 3886.4
255000 4018.96 3687.93
1/14
3544.1 4037.8
201468 3894.47 3141.56
1/15
3303.6 3988.4
416
195524 2260.79 2223.20
4128.3
7252.8
311552 3956.27 3673.69
417
3606.4
5096.0
263908 2101.80 2933.28
1758
3048.5
3165.2
105416 1868.00 1193.43
419
2351.6 4710.6
61944 3254.00 1990.19
5/20
1091.2
628.89
918.2
18900 1171.64
5521
243.1
354.8
289.28
334.40
584
422
373.3
2500
433.95
577.79
650.6
423
339.9
366.03
318.07
361.3
3156
424
1068.1
859.86
1485.8
56308
752.53
L/25
2202.1 3024.7
76260 1929.41 1903.70
5126
188156 5780.49 1885.85 3086.1 3528.4
1/27
4860.7
262528 3801.42 3414.41 8016.3
428
3019.6
70408 1923.13 1982.31 2247.5
5129
2263.2 1581.2
55848 1432.36 1257.84
5/30
932.6
2079.5
431
30312 1295.43 1039.44
2106.4
5132
87336 2022.88 1107.50 3499.0
2927.2 5311.4
5133
131284 1879.19 1581.85
3620.8 3837.8
190012 3588.85 2217.25
5/34
3456.5
4110 .6
219092 2848.96 2634.17
435
9820.6 10620.7
548556 6818.01 6175.25
1/36

STON

STOL

ST08

ST0T

STDV

12889.1 4951.8 6174.05 5282.83 5555.57
5472.1 3571.28 3202.94 2728.35
10622.5
5474.3 5138.63 4034.16 4446.95
4955.9
1855.0 2075.07 1529.46 1605.81
2971.8
27539.7 13593.1 7098.66 4860.96 3867.80
13566.6 7821.3 4083.41 3105.45 2453.22
6534.6 3717.1 1934.91 2198.95 1518.27
6896.7 3523.97 3363.23 2040.26
14064.9
2526.7 2444.33 1637.42 1949.88
5101.1
1846.6 1737.55 1351.68 1754.63
1922.5
755.68
895.38
902.6 772.24
960.9
825.9 921.96 950,80 745.82
1156.8
814.85
719.26 773.55
738.4
839.5
4041.1 3864.56 2553.56 2398.74
6060.6
4164.2 3568.57 2881.85 3110.84
5261.8
3183.4 2376.59 2548.81 1731.55
4587,2
7142.4 3740.8 3284.53 3684.01 3504.88
3934.4 3801.64 2666.05 2942.08
6514.6
2102.1 2234.38 1865.98 1254.27
3255.2
1993.8 1865.65 1575.65 2000.60
3938.4
560.77
453.5
582.69
524.73
1071.4
371.04
335.14
433.5
371.99
286.2
690.84
575.47
423.2
467.10
511.6
321.74
227.7 344.37 314.34
363.7
1493.1 1057.2 1072.84 1115.40 874.53
1692.9 2398.15 1750.99 1911.53
4004.9
2789.8 2778.07 2638.67 1697.49
5312.8
5742.5 3467.37 3567.69 2618.12
6511.4
5886.4 1802.0 2391.80 1809.64 1979.28
2177.8 1654.9 1633.95 1664.21 1143.43
746.3 984.74 1560.93 1035.58
950.4
899.90
1798.9 1688.94 2075.75
2555.4
3316.7 2047.7 1805.01 2301.19 1225.12
2722.2 2331.79 2670.12 2247.32
5520.4
2169.2 3196.68 2520.43 2889.30
6084.4
7959.1 7545.65 5213.86 5295.87
13770.8

INSEE Méthodes n° 52 - 53


Tableau 5

**Corrélations multiples entre les variables - test et tac1, fac2 issus de l'ACP sur C1-C3 (méthodes G) : COR2.**

<table>
<thead>
<tr>
<th>OBS NOM</th>
<th>CORG</th>
<th>OBS NOM</th>
<th>CORG</th>
<th>OBS NOM</th>
<th>CORG</th>
<th>OBS NOM</th>
<th>CORG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 W 1</td>
<td>0.982586</td>
<td>1 W 1</td>
<td>0.979591</td>
<td>1 W 1</td>
<td>0.984869</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 W 2</td>
<td>0.979242</td>
<td>2 W 2</td>
<td>0.976070</td>
<td>2 W 2</td>
<td>0.980744</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 W 3</td>
<td>0.586411</td>
<td>3 W 3</td>
<td>0.745916</td>
<td>3 W 3</td>
<td>0.701014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 W 4</td>
<td>0.575795</td>
<td>4 W 4</td>
<td>0.599221</td>
<td>4 W 4</td>
<td>0.612720</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 W 5</td>
<td>0.985775</td>
<td>5 W 5</td>
<td>0.992504</td>
<td>5 W 5</td>
<td>0.690711</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 W 6</td>
<td>0.972166</td>
<td>6 W 6</td>
<td>0.975527</td>
<td>6 W 6</td>
<td>0.975901</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 W 7</td>
<td>0.952769</td>
<td>7 W 7</td>
<td>0.953057</td>
<td>7 W 7</td>
<td>0.946462</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 W 8</td>
<td>0.985745</td>
<td>8 W 8</td>
<td>0.892323</td>
<td>8 W 8</td>
<td>0.981627</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 W 9</td>
<td>0.910704</td>
<td>9 W 9</td>
<td>0.926996</td>
<td>9 W 9</td>
<td>0.953768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 W 10</td>
<td>0.677984</td>
<td>10 W 10</td>
<td>0.619945</td>
<td>10 W 10</td>
<td>0.648598</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 W 11</td>
<td>0.721777</td>
<td>11 W 11</td>
<td>0.671751</td>
<td>11 W 11</td>
<td>0.686486</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 W 12</td>
<td>0.798627</td>
<td>12 W 12</td>
<td>0.782954</td>
<td>12 W 12</td>
<td>0.813024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 W 13</td>
<td>0.693084</td>
<td>13 W 13</td>
<td>0.707904</td>
<td>13 W 13</td>
<td>0.740541</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 W 14</td>
<td>0.895612</td>
<td>14 W 14</td>
<td>0.885025</td>
<td>14 W 14</td>
<td>0.891280</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 W 15</td>
<td>0.852624</td>
<td>15 W 15</td>
<td>0.846622</td>
<td>15 W 15</td>
<td>0.849704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 W 16</td>
<td>0.894523</td>
<td>16 W 16</td>
<td>0.873981</td>
<td>16 W 16</td>
<td>0.894676</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 W 17</td>
<td>0.891592</td>
<td>17 W 17</td>
<td>0.690131</td>
<td>17 W 17</td>
<td>0.950197</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 W 18</td>
<td>0.790999</td>
<td>18 W 18</td>
<td>0.764538</td>
<td>18 W 18</td>
<td>0.865111</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 W 19</td>
<td>0.515035</td>
<td>19 W 19</td>
<td>0.479220</td>
<td>19 W 19</td>
<td>0.659648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 W 20</td>
<td>0.457478</td>
<td>20 W 20</td>
<td>0.373507</td>
<td>20 W 20</td>
<td>0.596911</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 W 21</td>
<td>0.542285</td>
<td>21 W 21</td>
<td>0.542418</td>
<td>21 W 21</td>
<td>0.504264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 W 22</td>
<td>0.150655</td>
<td>22 W 22</td>
<td>0.088808</td>
<td>22 W 22</td>
<td>0.136585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 W 23</td>
<td>0.295267</td>
<td>23 W 23</td>
<td>0.140011</td>
<td>23 W 23</td>
<td>0.207404</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 W 24</td>
<td>0.255810</td>
<td>24 W 24</td>
<td>0.197772</td>
<td>24 W 24</td>
<td>0.209935</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 W 25</td>
<td>0.657229</td>
<td>25 W 25</td>
<td>0.620890</td>
<td>25 W 25</td>
<td>0.702510</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 W 26</td>
<td>0.269268</td>
<td>26 W 26</td>
<td>0.539837</td>
<td>26 W 26</td>
<td>0.445302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 W 27</td>
<td>0.871418</td>
<td>27 W 27</td>
<td>0.842355</td>
<td>27 W 27</td>
<td>0.870905</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 W 28</td>
<td>0.635964</td>
<td>28 W 28</td>
<td>0.870905</td>
<td>28 W 28</td>
<td>0.879359</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 W 29</td>
<td>0.245561</td>
<td>29 W 29</td>
<td>0.283594</td>
<td>29 W 29</td>
<td>0.559194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 W 30</td>
<td>0.457295</td>
<td>30 W 30</td>
<td>0.546307</td>
<td>30 W 30</td>
<td>0.701936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 W 31</td>
<td>0.402462</td>
<td>31 W 31</td>
<td>0.498085</td>
<td>31 W 31</td>
<td>0.578843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32 W 32</td>
<td>0.745655</td>
<td>32 W 32</td>
<td>0.725850</td>
<td>32 W 32</td>
<td>0.798287</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 W 33</td>
<td>0.840444</td>
<td>33 W 33</td>
<td>0.844522</td>
<td>33 W 33</td>
<td>0.612185</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 W 34</td>
<td>0.862697</td>
<td>34 W 34</td>
<td>0.841111</td>
<td>34 W 34</td>
<td>0.846507</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35 W 35</td>
<td>0.609772</td>
<td>35 W 35</td>
<td>0.678800</td>
<td>35 W 35</td>
<td>0.806154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36 W 36</td>
<td>0.940896</td>
<td>36 W 36</td>
<td>0.950084</td>
<td>36 W 36</td>
<td>0.942106</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Tableau 6**

**Corrélations multiples entre les variables - test et fac1, fac2, fac3 issus de l'ACP sur C1-C11 (méthodes H) : COR3.**

<table>
<thead>
<tr>
<th>R = r3</th>
<th>R = r4</th>
<th>R = r5</th>
<th>R = r7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RBS NOM</strong></td>
<td><strong>COR3</strong></td>
<td><strong>RBS NOM</strong></td>
<td><strong>COR3</strong></td>
</tr>
<tr>
<td>1 W 1 0.952281</td>
<td>2 U 1 0.962436</td>
<td>3 U 1 0.962436</td>
<td>5 W 1 0.949534</td>
</tr>
<tr>
<td>2 U 2 0.977357</td>
<td>2 U 2 0.976807</td>
<td>5 W 5 0.701339</td>
<td>2 U 2 0.970553</td>
</tr>
<tr>
<td>3 U 3 0.952665</td>
<td>3 U 3 0.974566</td>
<td>4 U 4 0.839400</td>
<td>5 W 3 0.615860</td>
</tr>
<tr>
<td>4 U 4 0.910895</td>
<td>4 U 4 0.886325</td>
<td>5 U 5 0.982357</td>
<td>6 U 4 0.929088</td>
</tr>
<tr>
<td>5 U 5 0.982839</td>
<td>5 U 5 0.981284</td>
<td>6 U 5 0.982357</td>
<td>5 W 5 0.985860</td>
</tr>
<tr>
<td>6 U 6 0.974084</td>
<td>6 U 6 0.965122</td>
<td>7 W 6 0.967540</td>
<td>6 U 6 0.967625</td>
</tr>
<tr>
<td>7 W 7 0.972435</td>
<td>7 W 7 0.976973</td>
<td>7 W 7 0.971429</td>
<td>7 U 7 0.973967</td>
</tr>
<tr>
<td>8 W 8 0.975107</td>
<td>8 W 8 0.977550</td>
<td>8 W 8 0.971428</td>
<td>8 W 8 0.979475</td>
</tr>
<tr>
<td>9 W 9 0.958084</td>
<td>9 W 9 0.943716</td>
<td>9 W 9 0.954616</td>
<td>9 U 9 0.952183</td>
</tr>
<tr>
<td>10 W 10 0.656115</td>
<td>10 W 10 0.660821</td>
<td>10 W 10 0.625586</td>
<td>10 W 10 0.619668</td>
</tr>
<tr>
<td>11 W 11 0.712535</td>
<td>11 W 11 0.725344</td>
<td>11 W 11 0.705901</td>
<td>11 W 11 0.705142</td>
</tr>
<tr>
<td>12 W 12 0.843250</td>
<td>12 W 12 0.821692</td>
<td>12 W 12 0.799725</td>
<td>12 W 12 0.837721</td>
</tr>
<tr>
<td>13 W 13 0.802357</td>
<td>13 W 13 0.751824</td>
<td>13 W 13 0.747512</td>
<td>13 W 13 0.778953</td>
</tr>
<tr>
<td>14 W 14 0.935813</td>
<td>14 W 14 0.915089</td>
<td>14 W 14 0.912556</td>
<td>14 W 14 0.950185</td>
</tr>
<tr>
<td>15 W 15 0.846629</td>
<td>15 W 15 0.845634</td>
<td>15 W 15 0.835211</td>
<td>15 W 15 0.856728</td>
</tr>
<tr>
<td>16 W 16 0.911687</td>
<td>16 W 16 0.913824</td>
<td>16 W 16 0.881414</td>
<td>16 W 16 0.900735</td>
</tr>
<tr>
<td>17 W 17 0.923225</td>
<td>17 W 17 0.877232</td>
<td>17 W 17 0.879725</td>
<td>17 W 17 0.925997</td>
</tr>
<tr>
<td>18 W 18 0.916071</td>
<td>18 W 18 0.880085</td>
<td>18 W 18 0.856466</td>
<td>18 W 18 0.928273</td>
</tr>
<tr>
<td>19 W 19 0.763535</td>
<td>19 W 19 0.751930</td>
<td>19 W 19 0.658729</td>
<td>19 W 19 0.800028</td>
</tr>
<tr>
<td>20 W 20 0.795317</td>
<td>20 W 20 0.797280</td>
<td>20 W 20 0.695571</td>
<td>20 W 20 0.955102</td>
</tr>
<tr>
<td>21 W 21 0.835427</td>
<td>21 W 21 0.735039</td>
<td>21 W 21 0.765438</td>
<td>21 W 21 0.809268</td>
</tr>
<tr>
<td>22 W 22 0.079893</td>
<td>22 W 22 0.254632</td>
<td>22 W 22 0.085898</td>
<td>22 W 22 0.140576</td>
</tr>
<tr>
<td>23 W 23 0.158616</td>
<td>23 W 23 0.304976</td>
<td>23 W 23 0.140632</td>
<td>23 W 23 0.220489</td>
</tr>
<tr>
<td>24 W 24 0.565974</td>
<td>24 W 24 0.255097</td>
<td>24 W 24 0.201943</td>
<td>24 W 24 0.527925</td>
</tr>
<tr>
<td>25 W 25 0.691852</td>
<td>25 W 25 0.654980</td>
<td>25 W 25 0.627066</td>
<td>25 W 25 0.714586</td>
</tr>
<tr>
<td>26 W 26 0.580809</td>
<td>26 W 26 0.311010</td>
<td>26 W 26 0.416613</td>
<td>26 W 26 0.493881</td>
</tr>
<tr>
<td>27 W 27 0.889170</td>
<td>27 W 27 0.915031</td>
<td>27 W 27 0.678291</td>
<td>27 W 27 0.865317</td>
</tr>
<tr>
<td>28 W 28 0.928555</td>
<td>28 W 28 0.907003</td>
<td>28 W 28 0.887116</td>
<td>28 W 28 0.911114</td>
</tr>
<tr>
<td>29 W 29 0.562043</td>
<td>29 W 29 0.260228</td>
<td>29 W 29 0.350564</td>
<td>29 W 29 0.455341</td>
</tr>
<tr>
<td>30 W 30 0.404004</td>
<td>30 W 30 0.461372</td>
<td>30 W 30 0.535576</td>
<td>30 W 30 0.686695</td>
</tr>
<tr>
<td>31 W 31 0.660128</td>
<td>31 W 31 0.594025</td>
<td>31 W 31 0.557844</td>
<td>31 W 31 0.649191</td>
</tr>
<tr>
<td>32 W 32 0.370592</td>
<td>32 W 32 0.821988</td>
<td>32 W 32 0.765995</td>
<td>32 W 32 0.858797</td>
</tr>
<tr>
<td>33 W 33 0.907182</td>
<td>33 W 33 0.877882</td>
<td>33 W 33 0.861953</td>
<td>33 W 33 0.858045</td>
</tr>
<tr>
<td>34 W 34 0.926521</td>
<td>34 W 34 0.926521</td>
<td>34 W 34 0.920568</td>
<td>34 W 34 0.927660</td>
</tr>
<tr>
<td>35 W 35 0.877652</td>
<td>35 W 35 0.877652</td>
<td>35 W 35 0.810560</td>
<td>35 W 35 0.988885</td>
</tr>
<tr>
<td>36 W 36 0.917651</td>
<td>36 W 36 0.917651</td>
<td>36 W 36 0.913687</td>
<td>36 W 36 0.926683</td>
</tr>
</tbody>
</table>
### Tableau 7

**Gains d'efficacité (ou perte) H/G : RAPHG**

**Rapport des coefficients de corrélation multiples :** \( \text{CORHG} = \frac{\text{CORH}_3}{\text{CORH}_2} \)

<table>
<thead>
<tr>
<th>Observations</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Region 3</th>
<th>Region 4</th>
<th>Region 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBS NOM</td>
<td>RAPHG</td>
<td>CORHG</td>
<td>OBS NOM</td>
<td>RAPHG</td>
<td>CORHG</td>
</tr>
<tr>
<td>1 W 0.71516</td>
<td>0.97951</td>
<td></td>
<td>1 W</td>
<td>1.28107</td>
<td>0.97329</td>
</tr>
<tr>
<td>2 W 1.36425</td>
<td>0.97951</td>
<td></td>
<td>2 W</td>
<td>1.15956</td>
<td>0.99252</td>
</tr>
<tr>
<td>3 W 0.99359</td>
<td>0.80027</td>
<td></td>
<td>3 W</td>
<td>1.51251</td>
<td>0.91808</td>
</tr>
<tr>
<td>4 W 0.33639</td>
<td>1.55965</td>
<td></td>
<td>4 W</td>
<td>0.75966</td>
<td>1.43522</td>
</tr>
<tr>
<td>5 W 1.52536</td>
<td>0.99242</td>
<td></td>
<td>5 W</td>
<td>1.67547</td>
<td>0.98904</td>
</tr>
<tr>
<td>6 W 0.98661</td>
<td>0.99260</td>
<td></td>
<td>6 W</td>
<td>0.98632</td>
<td>0.99583</td>
</tr>
<tr>
<td>7 W 0.86714</td>
<td>0.10159</td>
<td></td>
<td>7 W</td>
<td>0.51306</td>
<td>0.19130</td>
</tr>
<tr>
<td>8 W 1.17057</td>
<td>0.99370</td>
<td></td>
<td>8 W</td>
<td>1.70832</td>
<td>0.98772</td>
</tr>
<tr>
<td>9 W 0.53441</td>
<td>1.04174</td>
<td></td>
<td>9 W</td>
<td>0.86028</td>
<td>1.03181</td>
</tr>
<tr>
<td>10 W 0.62716</td>
<td>0.97469</td>
<td></td>
<td>10 W</td>
<td>0.76566</td>
<td>0.80878</td>
</tr>
<tr>
<td>11 W 0.64776</td>
<td>0.10190</td>
<td></td>
<td>11 W</td>
<td>0.67704</td>
<td>0.10476</td>
</tr>
<tr>
<td>12 W 0.64334</td>
<td>1.05276</td>
<td></td>
<td>12 W</td>
<td>0.78779</td>
<td>0.10212</td>
</tr>
<tr>
<td>13 W 0.84785</td>
<td>0.80478</td>
<td></td>
<td>13 W</td>
<td>0.97359</td>
<td>0.10523</td>
</tr>
<tr>
<td>14 W 0.81477</td>
<td>1.02403</td>
<td></td>
<td>14 W</td>
<td>1.03461</td>
<td>0.10544</td>
</tr>
<tr>
<td>15 W 0.79815</td>
<td>0.99206</td>
<td></td>
<td>15 W</td>
<td>0.97815</td>
<td>1.01017</td>
</tr>
<tr>
<td>16 W 1.02827</td>
<td>1.02555</td>
<td></td>
<td>16 W</td>
<td>1.18555</td>
<td>0.10620</td>
</tr>
<tr>
<td>17 W 1.15271</td>
<td>0.98889</td>
<td></td>
<td>17 W</td>
<td>0.99591</td>
<td>0.98856</td>
</tr>
<tr>
<td>18 W 0.61975</td>
<td>1.12474</td>
<td></td>
<td>18 W</td>
<td>0.81378</td>
<td>0.10493</td>
</tr>
<tr>
<td>19 W 0.93288</td>
<td>1.42213</td>
<td></td>
<td>19 W</td>
<td>0.60323</td>
<td>1.34442</td>
</tr>
<tr>
<td>20 W 0.70651</td>
<td>2.08677</td>
<td></td>
<td>20 W</td>
<td>0.49294</td>
<td>2.53720</td>
</tr>
<tr>
<td>21 W 0.57348</td>
<td>2.14444</td>
<td></td>
<td>21 W</td>
<td>0.73530</td>
<td>2.29955</td>
</tr>
<tr>
<td>22 W 0.64728</td>
<td>1.60303</td>
<td></td>
<td>22 W</td>
<td>0.99317</td>
<td>0.93359</td>
</tr>
<tr>
<td>23 W 1.09524</td>
<td>1.05128</td>
<td></td>
<td>23 W</td>
<td>0.59993</td>
<td>1.04444</td>
</tr>
<tr>
<td>24 W 1.10558</td>
<td>1.07089</td>
<td></td>
<td>24 W</td>
<td>0.64076</td>
<td>1.07410</td>
</tr>
<tr>
<td>25 W 0.89041</td>
<td>1.07276</td>
<td></td>
<td>25 W</td>
<td>0.98354</td>
<td>0.99994</td>
</tr>
<tr>
<td>26 W 0.88532</td>
<td>1.04828</td>
<td></td>
<td>26 W</td>
<td>0.92860</td>
<td>1.22592</td>
</tr>
<tr>
<td>27 W 0.87130</td>
<td>1.04632</td>
<td></td>
<td>27 W</td>
<td>1.57115</td>
<td>1.04241</td>
</tr>
<tr>
<td>28 W 0.67691</td>
<td>1.04992</td>
<td></td>
<td>28 W</td>
<td>1.17994</td>
<td>0.12861</td>
</tr>
<tr>
<td>29 W 0.78249</td>
<td>1.06059</td>
<td></td>
<td>29 W</td>
<td>0.92168</td>
<td>1.26351</td>
</tr>
<tr>
<td>30 W 1.33584</td>
<td>0.00537</td>
<td></td>
<td>30 W</td>
<td>0.95416</td>
<td>0.98193</td>
</tr>
<tr>
<td>31 W 0.60397</td>
<td>1.03889</td>
<td></td>
<td>31 W</td>
<td>1.36284</td>
<td>1.19162</td>
</tr>
<tr>
<td>32 W 0.54914</td>
<td>1.10354</td>
<td></td>
<td>32 W</td>
<td>0.86544</td>
<td>1.05622</td>
</tr>
<tr>
<td>33 W 0.30715</td>
<td>1.04442</td>
<td></td>
<td>33 W</td>
<td>1.05167</td>
<td>1.01846</td>
</tr>
<tr>
<td>34 W 1.21757</td>
<td>1.07407</td>
<td></td>
<td>34 W</td>
<td>0.87257</td>
<td>1.09359</td>
</tr>
<tr>
<td>35 W 0.53897</td>
<td>1.12241</td>
<td></td>
<td>35 W</td>
<td>0.64849</td>
<td>1.19508</td>
</tr>
<tr>
<td>36 W 1.26354</td>
<td>0.98233</td>
<td></td>
<td>36 W</td>
<td>1.65591</td>
<td>0.98161</td>
</tr>
</tbody>
</table>
NOTE: 1 OBS HAD MISSING VALUES OR WERE OUT OF RANGE

MODIFIED RANK PLOT OF RAPAL BY COMA - NON
L'échantillonnage
<table>
<thead>
<tr>
<th>NUM</th>
<th>ER1</th>
<th>ER2</th>
<th>ER3</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>17</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>19</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>21</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>22</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>23</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>24</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>25</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>26</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>27</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>28</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>29</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>30</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>31</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>32</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>33</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>34</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

L'échantillonnage
<table>
<thead>
<tr>
<th>NOM</th>
<th>ERT</th>
<th>ER1</th>
<th>ER2</th>
<th>ER3</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v2</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v3</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.85</td>
<td>0.85</td>
<td>-0.03</td>
<td>0.85</td>
</tr>
<tr>
<td>v4</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.04</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>v5</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v6</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>v7</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>v8</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>v9</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>v10</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v11</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v12</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v13</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>v14</td>
<td>-0.02</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>v15</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v16</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v19</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>v20</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v21</td>
<td>-0.02</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>v22</td>
<td>-0.02</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>v23</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v24</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v25</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v26</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v28</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v30</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v31</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v34</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>v35</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>v36</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
### Section 52

#### 1/76

<table>
<thead>
<tr>
<th>IDNR</th>
<th>ER1</th>
<th>ER2</th>
<th>ER3</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
</tbody>
</table>

### Section 52

#### 1/77

<table>
<thead>
<tr>
<th>IDNR</th>
<th>ER1</th>
<th>ER2</th>
<th>ER3</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
</tr>
</tbody>
</table>

**L'échantillonnage** 357
<table>
<thead>
<tr>
<th>nom</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>v2</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v3</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v4</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v5</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v6</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v7</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v8</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v9</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v10</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v11</td>
<td>-0.03</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v12</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v13</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v14</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v15</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v16</td>
<td>-0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v17</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v18</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v19</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v20</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v21</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v22</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v23</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v24</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v25</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v26</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v27</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v28</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v29</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v30</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v31</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v32</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v33</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v34</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v35</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v36</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**Méthodes n°52 - 53**
Éfficacité relative moyenne sur l'ensemble des régions
Box - Plots résumant les efficacités relatives H/G pour les 4 régions

**Box-Plots résumant les efficacités relatives H/G pour les 4 régions**

- Largeur
  - Min: 0.353654
  - Max: 2.375543

<table>
<thead>
<tr>
<th>NOM</th>
<th>RST1HG</th>
<th>RST2HG</th>
<th>RST5HG</th>
<th>RST5HG</th>
</tr>
</thead>
<tbody>
<tr>
<td>V 1</td>
<td>0.71</td>
<td>1.28</td>
<td>1.48</td>
<td>1.28</td>
</tr>
<tr>
<td>V 2</td>
<td>1.58</td>
<td>1.16</td>
<td>1.43</td>
<td>1.30</td>
</tr>
<tr>
<td>V 3</td>
<td>1.00</td>
<td>1.51</td>
<td>1.12</td>
<td>2.58</td>
</tr>
<tr>
<td>V 5</td>
<td>1.52</td>
<td>1.48</td>
<td>1.34</td>
<td>0.91</td>
</tr>
<tr>
<td>W 1</td>
<td>1.17</td>
<td>1.74</td>
<td>1.71</td>
<td>1.02</td>
</tr>
<tr>
<td>W 17</td>
<td>1.15</td>
<td>1.98</td>
<td>1.34</td>
<td>0.75</td>
</tr>
<tr>
<td>W 27</td>
<td>1.23</td>
<td>1.57</td>
<td>1.41</td>
<td>0.50</td>
</tr>
<tr>
<td>W 83</td>
<td>1.76</td>
<td>1.44</td>
<td>1.65</td>
<td>0.91</td>
</tr>
</tbody>
</table>

**Box-Plots résumant les efficacités relatives G/H pour les 4 régions**

- Largeur
  - Min: 0.553654
  - Max: 2.375543

<table>
<thead>
<tr>
<th>NOM</th>
<th>RST1HG</th>
<th>RST2HG</th>
<th>RST5HG</th>
<th>RST5HG</th>
</tr>
</thead>
<tbody>
<tr>
<td>V 4</td>
<td>0.55</td>
<td>0.74</td>
<td>1.00</td>
<td>0.51</td>
</tr>
<tr>
<td>V 7</td>
<td>0.86</td>
<td>0.51</td>
<td>0.75</td>
<td>0.76</td>
</tr>
<tr>
<td>V 10</td>
<td>0.66</td>
<td>0.75</td>
<td>0.73</td>
<td>1.15</td>
</tr>
<tr>
<td>V 11</td>
<td>0.63</td>
<td>0.68</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>V 12</td>
<td>0.84</td>
<td>0.79</td>
<td>0.27</td>
<td>1.00</td>
</tr>
<tr>
<td>V 13</td>
<td>0.85</td>
<td>0.97</td>
<td>0.74</td>
<td>1.35</td>
</tr>
<tr>
<td>V 14</td>
<td>0.81</td>
<td>1.03</td>
<td>1.08</td>
<td>0.92</td>
</tr>
<tr>
<td>V 15</td>
<td>0.90</td>
<td>0.95</td>
<td>1.16</td>
<td>0.81</td>
</tr>
<tr>
<td>V 16</td>
<td>0.42</td>
<td>0.82</td>
<td>0.93</td>
<td>1.40</td>
</tr>
<tr>
<td>V 19</td>
<td>0.95</td>
<td>0.60</td>
<td>0.92</td>
<td>0.64</td>
</tr>
<tr>
<td>V 20</td>
<td>0.71</td>
<td>0.49</td>
<td>0.91</td>
<td>0.61</td>
</tr>
<tr>
<td>V 21</td>
<td>0.57</td>
<td>0.75</td>
<td>0.60</td>
<td>0.74</td>
</tr>
<tr>
<td>V 24</td>
<td>0.88</td>
<td>0.95</td>
<td>0.72</td>
<td>0.99</td>
</tr>
<tr>
<td>V 28</td>
<td>0.68</td>
<td>1.13</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>V 29</td>
<td>0.78</td>
<td>0.92</td>
<td>0.57</td>
<td>1.03</td>
</tr>
<tr>
<td>W 2</td>
<td>0.71</td>
<td>0.66</td>
<td>1.27</td>
<td>0.55</td>
</tr>
<tr>
<td>W 5</td>
<td>0.33</td>
<td>0.35</td>
<td>0.84</td>
<td>0.24</td>
</tr>
<tr>
<td>W 28</td>
<td>0.92</td>
<td>0.77</td>
<td>0.62</td>
<td>0.82</td>
</tr>
<tr>
<td>W 83</td>
<td>0.54</td>
<td>0.62</td>
<td>0.84</td>
<td>0.93</td>
</tr>
</tbody>
</table>

**Légende**
- A: Nord-Pas-de-Calais
- B: Pays de Loire
- C: Midi-Pyrénées
- D: Auvergne

INSEE Méthodes n° 52 - 53
Aujourd'hui être femme en France...

... c'est être majoritaire dans la population et vivre plus longtemps que les hommes. C'est être plus souvent diplômée que dans le passé, mais encore cantonnée dans une profession "féminine", plus engagée dans le monde du travail, mais plus souvent chômeuse, avoir une influence dans la vie sociale et politique en progrès, mais toujours loin de l'égalité avec les hommes. La collection "Contours et Caractères" présente un portrait des femmes en 80 thèmes : famille, santé, éducation, emploi, ressources, conciliation de la vie familiale et du travail, loisirs, relations avec la société et la politique sont autant de sujets abordés.

225 pages, format 16 x 24 cm. 80 FF

A NOTER : 36 68 07 60*, l'INSEE vous informe 24h sur 24. indices, informations, adresses utiles. * 0,19 le minute.
**BULLETIN D'ABONNEMENT**

A retourner à : INSEE - CNGP
BP 2718 - 80027 AMIENS Cedex 01

**Nom, Prénom**
**Adresse complète**

**INSTITUT NATIONAL**
**DE LA STATISTIQUE**
**ET DES ÉTUDES ÉCONOMIQUES**

**INFORMATIONS RAPIDES**
Dans les meilleurs délais, la mise à jour de l'information conjoncturelle.
Abonnement 1 an : (350 numéros) France : 1 670 FF Etranger : 2 088 FF
Etranger avion : 2 591 FF

**BULLETIN MENSUEL DE STATISTIQUE**
10 000 séries mensuelles, trimestrielles et annuelles concernant l'ensemble de la vie économique et sociale.
Etranger avion : 562 FF
Etranger avion : 249 FF

**NOTE DE CONJONCTURE**
Une analyse de la situation et des perspectives de l'économie française.
Abonnement 1 an : (3 numéros + 2 points conjoncturels) France : 167 FF Etranger : 209 FF
Etranger avion : 255 FF

**NOTE DE CONJONCTURE INTERNATIONALE**
3 fois par an, un panorama de la conjoncture mondiale dressé par la Direction de la Prévision.
Abonnement : (3 numéros) France : 137 FF Etranger : 171 FF
Etranger avion : 222 FF

**CONJONCTURE IN FRANCE**
Synthèse en anglais à destination de l'étranger.
Abonnement 1 an : (2 numéros) France : 44 FF Etranger : 55 FF
Etranger avion : 67 FF

**L'INFORMATION SUR L'INFORMATION**

**BLOC-NOTES**
de INSEE Info Service
A la fois un répertoire et un guide de l'information économique. Le "thème du mois" fournit des repères sur un sujet d'actualité.
Abonnement 1 an : (11 numéros) France : 163 FF Etranger : 204 FF
Etranger avion : 324 FF

**INSEE INFOS**
Chaque mois, sous forme de brèves, une vitrine des produits de diffusion de l'INSEE.
Abonnement gratuit

**INSEE NEWS, LA CARTA DEL INSEE, LA LETTRE DE L'INSEE**
A destination des lecteurs étrangers, en anglais, en espagnol et en français, une chronique bi-annuelle des principaux titres et travaux de l'INSEE.
Abonnement gratuit

**COURRIER DES STATISTIQUES**
Quatre fois par an, une vue d'ensemble de la vie du système statistique français.
Abonnement 1 an : (4 numéros) France : 135 FF Etranger : 169 FF
Etranger avion : 234 FF

**SCRIBECO**
Une revue bibliographique reflet du fond documentaire de l'INSEE.
Abonnement 1 an : (6 numéros) France : 657 FF Etranger : 621 FF
Etranger avion : 892 FF

Veuillez noter mon abonnement aux publications suivantes :

**LA CONJONCTURE**
**TENDANCES DE LA CONJONCTURE**
**INFORMATIONS RAPIDES**
**BULLETIN MENSUEL DE STATISTIQUE**
**NOTE DE CONJONCTURE**
**NOTE DE CONJONCTURE INTERNATIONALE**
**CONJONCTURE IN FRANCE**

Je souhaite recevoir le catalogue des publications de l'INSEE

Ci-joint un chèque de ............... FF*
*pour l'étranger libellé en FF.

(Total des abonnements) à l'ordre de l'Insee

Date :
Signature:

---

**INSEE INFOS**
Chaque mois, sous forme de brèves, une vitrine des produits de diffusion de l'INSEE.
Abonnement gratuit

**INSEE NEWS, LA CARTA DEL INSEE, LA LETTRE DE L'INSEE**
A destination des lecteurs étrangers, en anglais, en espagnol et en français, une chronique bi-annuelle des principaux titres et travaux de l'INSEE.
Abonnement gratuit

**COURRIER DES STATISTIQUES**
Quatre fois par an, une vue d'ensemble de la vie du système statistique français.
Abonnement 1 an : (4 numéros) France : 135 FF Etranger : 169 FF
Etranger avion : 234 FF

**SCRIBECO**
Une revue bibliographique reflet du fond documentaire de l'INSEE.
Abonnement 1 an : (6 numéros) France : 657 FF Etranger : 621 FF
Etranger avion : 892 FF

Veuillez noter mon abonnement aux publications suivantes :

**LA CONJONCTURE**
**TENDANCES DE LA CONJONCTURE**
**INFORMATIONS RAPIDES**
**BULLETIN MENSUEL DE STATISTIQUE**
**NOTE DE CONJONCTURE**
**NOTE DE CONJONCTURE INTERNATIONALE**
**CONJONCTURE IN FRANCE**

Je souhaite recevoir le catalogue des publications de l'INSEE

Ci-joint un chèque de ............... FF*
*pour l'étranger libellé en FF.

(Total des abonnements) à l'ordre de l'Insee

Date :
Signature :
LES PÉRIODIQUES

INSEE RESULTATS

Les données détaillées des enquêtes et opérations statistiques de l'INSEE.

INSEE RESULTATS

est articulé en 5 thèmes :
- Economie générale
- Démographie-Société
- Consommation-Modes de vie
- Système productif
- Emploi-Revenus.

Il est possible de s'abonner à un, plusieurs ou tous les thèmes.

Abonnement INSEE RESULTATS

Economie générale (20 numéros)
France : 1 454 FF - Etranger : 1 818 FF - Etranger avion : 2 075 FF

Démographie-Société (7 numéros)
France : 909 FF - Etranger : 536 FF - Etranger avion : 726 FF

Consommation-Modes de vie (15 numéros)
France : 1 091 FF - Etranger : 1 364 FF - Etranger avion : 1 557 FF

Système productif (15 numéros)
France : 1 091 FF - Etranger : 1 364 FF - Etranger avion : 1 557 FF

Emploi-revenus (13 numéros)
France : 945 FF - Etranger : 1 181 FF - Etranger avion : 1 344 FF

Ensemble des 5 thèmes (70 numéros)
France : 5 090 FF - Etranger : 6 363 FF - Etranger avion : 7 259 FF

Abonnement sur microfiches INSEE RESULTATS
(ensemble des thèmes)
France : 2 284 FF - Etranger : 2 955 FF - Etranger avion : 3 755 FF

INSEE METHODES

La méthodologie des travaux de l'INSEE et les modèles.

Abonnement 1 an (15 numéros)

ANNÉES D'ÉCONOMIE ET DE STATISTIQUE

Des travaux originaux de recherche théorique ou appliquée dans les domaines de l'économie, de l'économétrie et de la statistique.

Abonnement 1 an (4 numéros)

Pour les particuliers (édition papier) :
Abonnement 1 an (4 numéros)
France : 157 FF - Etranger : 196 FF - Etranger avion : 237 FF

RECUDE D'ETUDES SOCIALES

Sélection d'études déjà publiées par différents organismes publics français d'études et de statistiques.

Abonnement 1 an (3 numéros)

BULLETIN D'ABONNEMENT

A retourner à : INSEE - CNGP
BP 2718 - 80027 AMIENS Cedex 01

Nom, Prénom
Société
Activité
Adresse complète

Date :
Signature :

*Pour l'étranger libellé en FF.
(Total des abonnements) à l'ordre de l'Insee :
L’INSEE DANS VOTRE RÉGION

VOUS Y TROUVEREZ :
- Salle de documentation en libre consultation
- Bureau de vente des publications de l’INSEE
- Adresses des entreprises et établissements (SIRENE).
- Accès au fond documentaire et aux banques de données de l’INSEE
- Travaux à la demande...

EN OUTRE - MER :

ANTILLES-GUYANE
INSEE : Direction Inter-Régionale Tour Secrid. 7ème Etage. Place de la Réunion, BP 863, 97153 POINTE-A-PITRE CEDEX Tél. : 19 590 91 99

GUADELOUPE
INSEE : Service Régional Rue Paul Lacaze. BP 96, 97152 BASSE-TERRA Tél. : 19 590 81 42 50

INSEE - DIRECTION GÉNÉRALE
Unité Communication Externe 18, Bd Adolphe Pinard 75675 Paris Cedex 14 - FRANCE

NORD-PAS-DE-CALAIS
10-12 boulevard Vauban. 59800 LILLY Tél. : 20 12 87 20

BASSE-NORMANDIE
93-95 rue de Gécele. 14052 CAEN CEDEX Tél. : 31 85 35 70

HAUTER-NORMANDIE
4, quai de la Bourse. 76037 ROUEN CEDEX Tél. : 35 52 49 11

PAYS DE LA LOIRE
105, rue des Français Libres. BP 2189, 44304 NANTES CEDEX 92 Tél. : 40 41 75 75

LORRAINE
15, rue du Général Bulot, BP 3846. 54029 NANCY CEDEX 4 Tél. : 83 91 85 85

LE SERVICE INSEE 24H/24
36 68 07 60 *
- indices
- informations
- adresses

et sur minitel 36.15 - 36.16 INSEE
* 2,23 Fonds France entière

EN OUTRE - MER :

ANTILLES-GUYANE
INSEE : Direction Inter-Régionale Tour Secrid. 7ème Etage. Place de la Réunion, BP 863, 97153 POINTE-A-PITRE CEDEX Tél. : 19 590 91 99

GUADELOUPE
INSEE : Service Régional Rue Paul Lacaze. BP 96, 97152 BASSE-TERRA Tél. : 19 590 81 42 50

INSEE - DIRECTION GÉNÉRALE
Unité Communication Externe 18, Bd Adolphe Pinard 75675 Paris Cedex 14 - FRANCE

Tél. renseignements : (1) 41 17 66 11
Tél. administration : (1) 41 17 50 50
Fax : (1) 41 17 51 77
IMPRIMERIE NATIONALE

Impression d'après documents fournis

5 170272 T
Sous une apparence de forte continuité, le recensement de 1990 comme chacun de ses prédécesseurs a en fait largement innové dans ses procédures. Cet "INSEE-Méthodes" décrit quelques unes des innovations en matière de collecte, un contrôle de qualité de la saisie et de la codification en ligne sensiblement plus économe grâce à un recours important aux techniques statistiques, la première utilisation censitaire de procédures automatiques de codification, des procédures renouvelées d'imputation de valeurs manquantes, et surtout, dans un plan d'exploitation assez différent de celui du RP de 1982, l'avancée méthodologique importante qu'a constitué le tirage équilibré de districts pour la constitution de l'échantillon au 1/20e ; bonne illustration de ce qu'apportent les allers-retours entre méthodologues et responsables techniques d'une opération statistique de grande ampleur : la mise au point d'algorithmes efficaces puis leur utilisation a été précédée d'une recherche plus fondamentale pour transformer une idée simple en y introduisant les avantages de procédures probabilistes après en avoir vérifié la qualité.

Ce volume présente également des appréciations sur la qualité du recensement (exhaustivité et qualité des déclarations).

Par contre il néglige les avancées substantielles également réalisées dans la mise à disposition des résultats.