

TABLE DES MATIÈRES

Michel POULAIN - Introduction	1
Didier BLANCHET, Sophie PENNEC – Une application des méthodes de micro-simulation : la gestion du personnel	3
Michele COSTABILE, Gianluca MARZOCCHI - Instruments d'analyse démographique et mesure de la satisfaction du consommateur.....	17
Luc DAL - Contribution à la prévision de la population étudiante universitaire de Belgique francophone.....	33
Giuseppe De BARTOLO, Rosina CHIAPPETTA - Estimation de population locale pour la gestion publique et le marketing.....	47
Thierry EGGERICKX - La démographie comme aide à la gestion locale	55
Carlo M. GALLUCI, Ivana CASABURI et Michele QUINTANO - Influences des familles sur le comportement du consommateur : le cas espagnol.....	71
Jacques MENTHONNEX - De l'utilité et des méthodes pour établir des perspectives démographiques relatives à de petites régions.....	89
Louis G. POL - L'utilisation des méthodes démographiques en démographie appliquée..	97
Michel POULAIN - Quel avenir pour le clergé namurois ? Le point de vue du démographe ...	109
Filomena RACIOPPI - Démographie professionnelle : Opportunités et occasions de formation	117
Djilali SARI - L'approche de localisation optimale	131
Isabelle THOMAS - Analyse spatiale et modèles de localisation optimale. Outils opérationnels d'aide à la décision.....	139
Pierre ARS, Luc DAL et Michel POULAIN - Comment appréhender le problème statistique des petits nombres en démographie ?	155
Christophe BERGOUIGNAN - La démographie post censitaire communale et ses facteurs de qualité.....	179
Jean Louis RALLU - Diversité et dynamiques dans les micro-populations.....	201
Philippe WANNER - La démographie appliquée aux populations de petite taille. Le point de vue du statisticien-démographe	211

Introduction

Michel POULAIN

Université Catholique de Louvain, Belgique

Dans nos sociétés contemporaines, l'analyse démographique trouve une multitude de champs d'application aussi bien dans le domaine de la gestion publique que dans le cadre de la gestion de l'entreprise ou l'étude de la consommation. Le développement de la démographie appliquée s'avère donc indispensable et ceci justifie l'existence d'un groupe de travail spécifique au sein de la *Population Association of America*. Au sein de l'AIDELF, mais également dans le cadre de l'EAPS (European Association for Population Studies), le même souci de collaboration et d'échange dans le champ de la démographie appliquée a vu le jour et s'est concrétisé récemment. En avril 1995, un séminaire de l'AIDELF a été organisé à Cosenza (Calabre, Italie) à l'initiative de Giuseppe DE BARTOLO sur le thème « La démographie appliquée au service de l'entreprise et de la gestion publique ». Par ailleurs, dans le cadre du congrès de Cracovie en juin 1997, une séance spéciale a été consacrée à la démographie appliquée sous les auspices d'un groupe de travail mis sur pied sur ce thème. Ce sont les communications présentées à ces deux occasions qui sont rassemblées dans cet ouvrage publié par l'AIDELF.

Selon l'ordre chronologique de ces deux manifestations scientifiques et par ordre alphabétique des auteurs, nous lirons tout d'abord la contribution de Didier BLANCHET et Sophie PENNEC qui appliquent les méthodes de micro-simulation à la gestion du personnel d'une entreprise. L'âge y est évidemment la variable clé et aux côtés des événements démographiques sont introduits des événements marquant la vie professionnelle pour en déduire des prévisions d'effectifs et favoriser une meilleure gestion du personnel au sein d'une entreprise. Dans leur intervention, Michele COSTABILE et Gianluca MARZOCCHI démontrent de quelle façon les outils classiques de la démographie permettent de mesurer la satisfaction du consommateur en standardisant les données collectées par rapport au sexe et au groupe d'âges, ils en déduisent quelques indicateurs clé de la satisfaction du consommateur. Giuseppe DE BARTOLO et Rosina CHIAPPETTA envisagent quant à eux la problématique de l'estimation de la taille des populations locales dans une perspective de gestion publique et du marketing avec une attention toute particulière à la problématique de la collecte des déchets ménagers.

Le domaine de la gestion publique est le thème central de la présentation de Thierry EGGERICKX qui présentent l'ensemble des sources disponibles pour la démographie appliquée à l'échelle locale en Belgique et s'interroge sur le fait de savoir si la démographie des populations locales est ou non techniquement possible. Avec Carlo GALLUCI, Ivana CASABURI et Michele QUINTANO, on envisage l'influence des variables démographiques décrivant les individus, les familles et les ménages sur le comportement du consommateur dans le cas particulier de l'Espagne, en s'intéressant tout particulièrement au processus de décision dans la famille en matière de consommation. Le recours à une typologie des familles facilitera une analyse plus précise des habitudes d'achat. Jacques MENTHONNEX considère également la problématique des petites régions et rejoint en cela la préoccupation de Thierry Eggerickx et celle de Giuseppe De Bartolo et Rosina Chiappetta. Il envisage ici tout un ensemble de techniques de prévisions démographiques et leur pertinence à l'échelle des petites régions en soumettant les résultats à l'épreuve des faits. Bien souvent dans ce contexte le démographe est forcé au compromis entre méthodes différentes devenant complémentaires et ce faisant l'auteur aborde même les prévisions démographiques à l'échelle des quartiers d'une grande ville, Lausanne, en l'occurrence.

Louis POL dresse dans sa contribution un large aperçu de cette sous-discipline de la démographie étiquetée *démographie appliquée* qui s'est développée aux États-Unis depuis le début des années 80. Par la suite il passe en revue les différents outils propres à la démographie qui sont d'usage courant en démographie appliquée. Quel avenir pour le clergé namurois ? Telle est la problématique envisagée par Michel POULAIN avec les outils prospectifs du démographe et en prenant en considération, tout comme l'ont fait Didier Blanchet et Sophie Pennec, les événements spécifiques à la carrière sacerdotale. Filomena RACIOPPI s'intéresse aux différents types de formation en démographie proposés en Italie par comparaison aux États-Unis et examine les différents débouchés professionnels offerts aux démographes, essentiellement dans le domaine de la démographie appliquée. Si Djalila SARI analyse la localisation optimale d'infrastructure en fonction de l'occurrence de certaines catastrophes comme les inondations, Isabelle THOMAS fait état de divers outils opérationnels d'aide à la décision dans ce même domaine de l'analyse spatiale et de la localisation optimale. Sa recherche d'optimalité est ici développée à partir de l'exemple de la localisation des bureaux de poste dans une phase de réduction de leur nombre.

Dans le cadre de la séance spéciale du Congrès de l'EAPS qui s'est tenu à Cracovie, Pierre ARS, Luc DAL et Michel POULAIN reviennent de façon plus approfondie sur le problème statistique des petits nombres en Démographie et débouchent sur des propositions d'estimation par intervalles de confiance et tests d'hypothèses en se basant sur les théories statistiques adéquates. Christophe BERGOUIGNAN touche également à la problématique des petites populations en proposant des méthodes d'estimation post-censitaire de la population des 542 communes de la Gironde (France). Il dresse un bilan fort à propos proposant une analyse empirique des écarts et dispersions résiduelles ainsi obtenues. Jean-Louis RALLU s'intéresse également aux populations de petite taille dans le cadre de trois petites îles du Pacifique Sud afin de mettre en évidence la diversité de leur population et les dynamiques démographiques qui leur sont propres. Philippe WANNER enfin clôture cet ouvrage en fournissant le point de vue du statisticien-démographe sur cette même problématique de la démographie appliquée aux populations de petite taille. Pour cela il prend l'exemple pratique d'une ville de taille moyenne (100.000 habitants) et propose enfin un ensemble cohérent de recommandations.

À la lecture de ces différentes communications, il ne fait plus aucun doute, la démographie appliquée existe bel et bien dans le paysage de la démographie, qu'elle soit francophone ou européenne. Les méthodes utilisées sont les mêmes que celles développées en démographie en général, mais le plus souvent les populations concernées sont des sous-populations spécifiques ou des populations de territoire restreint. Ainsi, le problème des populations de petite taille est-il récurrent en démographie locale et les contributions à ce recueil démontrent qu'une démographie appliquée aux petites populations est possible en prenant quelques précautions d'ordre statistiques.

Une application des méthodes de micro-simulation : la gestion du personnel

Didier BLANCHET

Institut National d'Études Démographiques, Paris, France

Institut National de la Statistique et des Études Économiques, Paris, France

Sophie PENNEC

Institut National d'Études Démographiques, Paris, France

Les méthodes de micro-simulation, introduites dans le domaine économique par Orcutt (1957,1961) et dans le domaine démographique par Hyrénus et Adolfsson (1964) ou Sheps (1963), sont d'un usage croissant pour l'évaluation des politiques sociales ou fiscales et pour la prévision démographique désagrégée (structures familiales). Mais, à notre connaissance, elles n'ont guère été exploitées dans le domaine de la projection et de la gestion prévisionnelle des populations professionnelles. Pourtant, elles présentent à cet égard deux propriétés intéressantes :

- D'une part, elles sont particulièrement à même de rendre compte de règles de promotion ou de comportements d'entrée-sortie dont le détail est difficile à saisir à l'aide des méthodes matricielles classiques, sauf à envisager des désagréations très poussées qui rendent très complexes la programmation et la gestion des modèles correspondants.
- D'autre part, en mettant l'accent sur l'individu, elles permettent de clarifier les liens entre évolution globale des structures et perspectives d'évolution individuelle, le caractère stochastique de la simulation permettant de surcroît de décrire ces dernières en termes probabilistes, c'est-à-dire en termes de distribution de carrières possibles, et non pas seulement en termes de profil de carrière moyen.

L'objet de cette communication est de présenter le fonctionnement et quelques résultats d'un modèle illustrant ces deux aspects. La première section rappellera quelques principes généraux concernant la construction de modèles de projection de populations professionnelles et la façon dont ils sont mis en œuvre dans le présent programme. La section 2 présentera le détail de la structure du modèle. Les résultats feront l'objet de la section 3.

1. Logique d'ensemble des modèles de projection de personnel

Une population professionnelle est constituée d'un ensemble de sous-populations, en général hiérarchisées, soumises à diverses formes d'échange, soit entre elles (promotions-rétrogradations), soit avec l'extérieur (recrutement, démissions et départs en retraite). Pour la modélisation de ces différents flux, les modèles existants utilisent deux logiques de base, non exclusives¹ :

- Une logique « poussée » : des probabilités de recrutement, de transition ou de sortie exogènes déterminent l'évolution des effectifs dans chaque groupe : ce sont donc ces effectifs qui font office de variables endogènes, le but étant d'en simuler l'évolution spontanée (et qui n'est en générale que potentielle), sous l'influence de comportements ou d'évolutions de carrière données a priori, pour lesquelles une hypothèse de base est en

¹ Nous reprenons cette terminologie de Couronne (1986). Ces principes de projection se retrouvent, explicitement ou implicitement, dans la plupart des travaux existants (en remontant à Henry, 1971). Les deux méthodes sont par exemple utilisées dans le travail de F. Muñoz-Perez et M. Tribalat (1993).

général la reproduction des comportements ou évolutions de carrière ayant prévalu dans le passé.

- Une logique « tirée » : elle consiste à supposer au contraire que ce sont les effectifs futurs qui constituent le jeu d'hypothèses exogènes sur lesquelles l'ensemble de la projection doit se caler, les flux d'entrée-sortie ou de promotion étant sensés s'ajuster pour que se réalisent ces objectifs exprimés en termes d'effectifs. Ceci laisse évidemment de nombreux degrés de liberté à la simulation, puisqu'une même évolution de l'effectif peut être obtenue avec une infinité de combinaisons d'entrées et de sorties. Aussi fait-on intervenir des contraintes supplémentaires. Par exemple, on peut se fixer les taux de sortie. C'est alors la résultante de ces hypothèses sur les effectifs désirés et les sorties qui détermine les promotions et les entrées. Celles-ci se font donc « par appel d'air », en fonction des places créées ou rendues disponibles par départ, d'où l'appellation « tirée ».

Le présent modèle retient lui aussi un mélange de ces deux logiques. Il concerne la simulation d'une population divisée en trois catégories hiérarchisées (rangées par ordre hiérarchique décroissant). Il implique donc la simulation de huit types de flux :

- Des départs aux trois niveaux (par retraite, décès, démission ou licenciement).
- Des recrutements pouvant aussi intervenir aux trois niveaux (recrutement à la base ou en cours de carrière aux deux grades supérieurs).
- Deux flux de promotion (du grade 3 vers le grade 2 et du grade 2 vers le grade 1).

Pour tous ces flux, le modèle prévoit une simulation en deux temps. Le premier temps relève d'une simple logique « poussée ». On simule ainsi un premier ensemble de mouvements spontanés ou exogènes, dits encore « autonomes » : départs en retraite (déterminés par l'âge), décès, départs volontaires (démissions) mais aussi, éventuellement, des flux de recrutements ou promotions « autonomes » (les premiers étant motivés, par exemple, par un besoin de renouvellement minimum à la base). Dans un deuxième temps, on établit une comptabilité des effectifs résultant de ces mouvements autonomes, effectifs que l'on compare à des effectifs désirés. L'écart entre les uns et les autres est alors résolu par une logique tirée, de proche en proche, à partir du grade le plus élevé : lorsque l'effectif du grade i qui résulte de l'étape « poussée » est plus faible que l'effectif désiré, on rétablit cet effectif en intensifiant les promotions depuis le grade juste inférieur et/ou les recrutements directs, selon un dosage choisi a priori. Dans le cas contraire, l'ajustement se fait en procédant à des départs forcés. On appellera « flux d'ajustement » ces flux entrants ou sortants permettant l'adéquation finale aux effectifs et à la structure désirée.

Cette structure générale est en fait assez flexible pour incorporer toute une gamme de cas particuliers, en jouant sur la façon dont sont paramétrés les flux autonomes et en jouant sur la possibilité de bloquer tout ou une partie des flux d'ajustement :

- Supposer des flux autonomes nuls (hors départ à la retraite) équivaudra à un modèle tiré pur.
- Autoriser des flux autonomes non nuls et interdire toute forme de flux d'ajustement conduira au contraire à simuler un modèle poussé pur.
- Un cas intermédiaire intéressant sera celui où le blocage des flux d'ajustement correspondra uniquement à l'interdiction des licenciements. Dans ce cas, le modèle fonctionnera selon une logique tirée tant que les mouvements autonomes conduiront à des déficits d'effectifs *ex ante*, et passera en mode « poussé » dans le cas contraire.

Ces principes généraux sont évidemment applicables à toutes les catégories de modèles de projection, qu'il s'agisse de modèles matriciels ou de micro-simulations. On va maintenant préciser leur mise en oeuvre dans ce second cas de figure.

2. Mise en œuvre dans un cadre de micro-simulation

a) Le fichier de départ

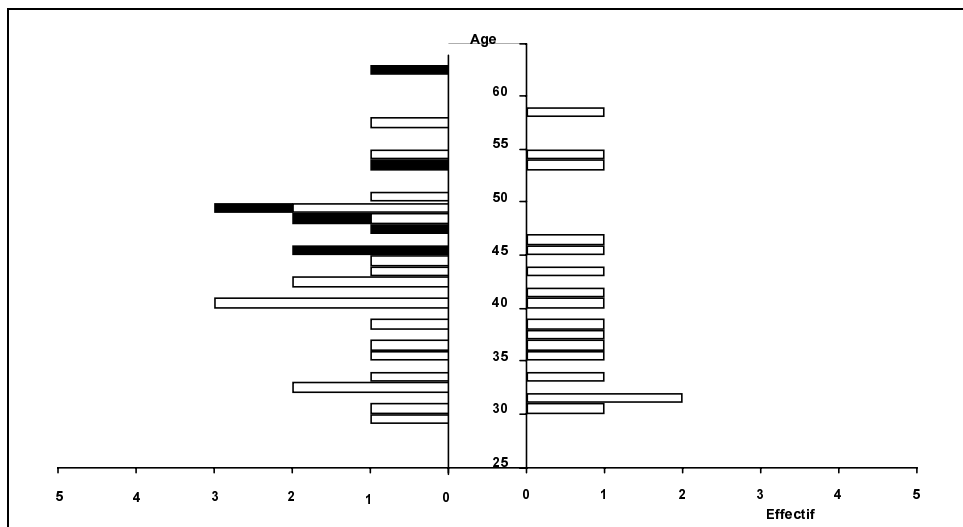
Un exercice de micro-simulation suppose la constitution d'un fichier de données individuelles initiales. Celui utilisé dans ce travail correspond à une population fictive (quoique inspirée d'une population réelle), correspondant à un corps de taille très limité (44 personnes). Cette population est répartie en trois grades que nous dénommerons respectivement, du plus bas au plus élevé, grade 3, grade 2 et grade 1.

Pour chaque individu, on possède les caractéristiques suivantes :

- Âge (ou date de naissance)
- Sexe
- Grade
- Ancienneté dans la population
- Ancienneté dans le grade

La répartition de l'ensemble du personnel par âge et grade est donnée par la figure 1. L'irrégularité de la pyramide est liée à l'histoire dans le recrutement. Elle est d'autant plus visible que l'effectif est faible. Irrégularité par sexe (64 % d'hommes, 36 % de femmes), par âge (35 % des hommes ont plus de 50 ans, c'est le cas de 18 % des femmes). Au niveau de la répartition selon le grade, les personnes aux grades 1 et 2 se retrouvent, tout naturellement, parmi les plus âgées car, d'une part, une fois atteint ce grade, elles y restent jusqu'à leur sortie de la population ; d'autre part, une période d'ancienneté minimale dans le grade inférieur est requise pour pouvoir postuler au passage au grade supérieur.

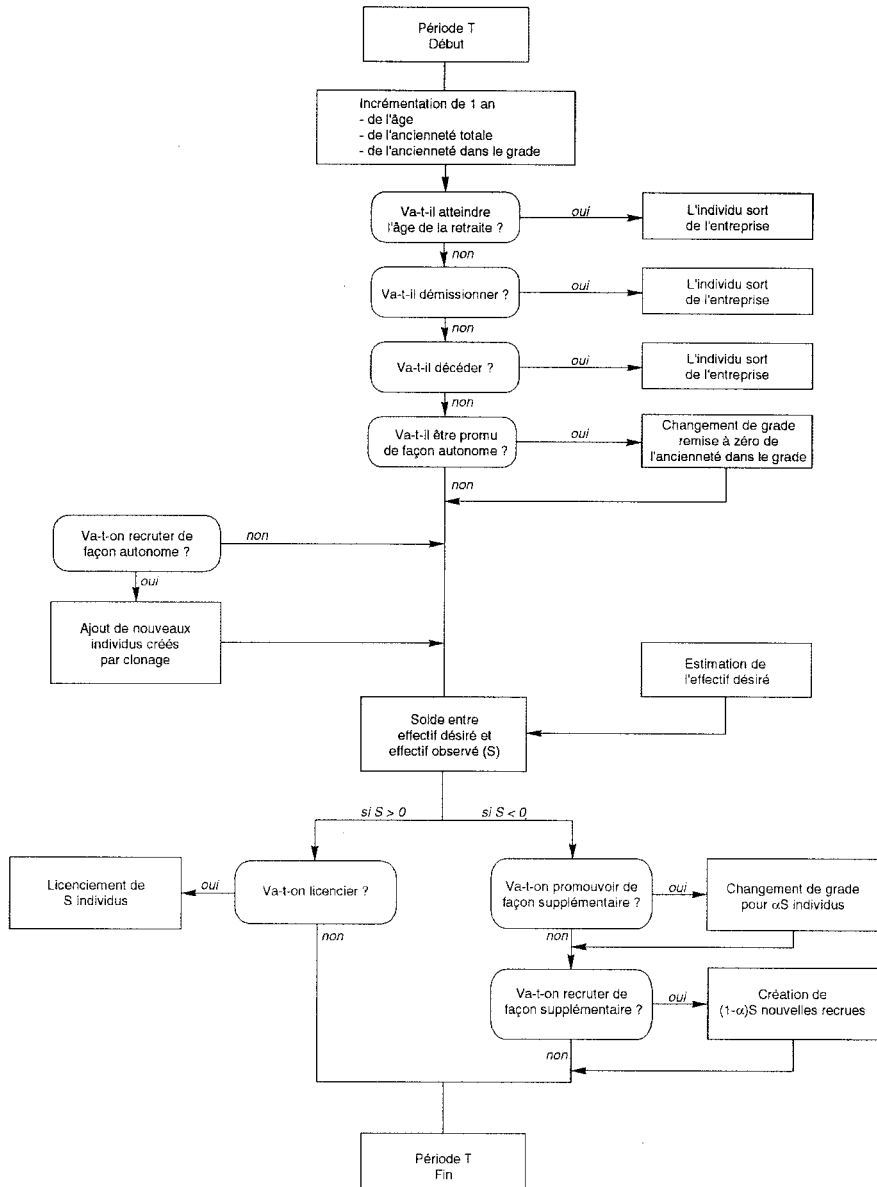
FIGURE 1 : RÉPARTITION PAR SEXE, ÂGE ET GRADE DE LA POPULATION



b) Organigramme de la simulation

La structure générale du programme de simulation est décrite à la figure 2. On distingue deux parties : l'une consacrée aux individus, l'autre à l'entreprise dans son ensemble, comprenant la totalité du personnel.

FIGURE 2 : ORGANIGRAMME DE LA SIMULATION



La partie individu teste pour chaque membre de la population, pour chaque période de simulation, sa probabilité de décéder, de démissionner, et selon le type de gestion retenu, d'être promu de manière « autonome », c'est-à-dire suivant un plan de carrière minimal défini.

La partie entreprise s'intéresse tout d'abord aux recrutements si ceux-ci sont systématiques (autonomes). Ensuite, on calcule l'effectif désiré c'est-à-dire celui qui doit être atteint pour que l'entreprise respecte ses paramètres de croissance désirée. Selon le résultat de la comparaison

entre effectif désiré et effectif observé (ce dernier correspond au nombre des présents dans l'entreprise après les entrées et les sorties autonomes), une alternative se présente :

- si le solde est négatif, possibilité d'ajustement par licenciements ;
- si le solde est positif, possibilité de promotions supplémentaires et/ou recrutements supplémentaires.

Le même programme permet alors de simuler toute une palette de modes de gestion selon que l'on active ou non certains paramètres. Les paramètres individuels tels que décès, démission, retraite sont toujours des flux poussés quel que soit le modèle choisi. Le caractère « poussé », « tiré » ou mixte de la simulation dépendra alors de la façon dont se font les promotions et les recrutements. Si les fonctions de recrutements et de promotions sont totalement fixées a priori, nous sommes dans une logique « poussée », l'évolution de la population se déduit de ces paramètres. A l'inverse, dans une logique « tirée », c'est le solde entre effectif désiré et effectif observé qui détermine les paramètres qui vont intervenir, soit pour combler un déficit, soit pour réduire un surplus d'effectif.

c) Les sorties engendrées par le programme

Le programme utilise une bibliothèque d'utilitaires généraux (Blanchet, 1993), qui facilite en particulier les opérations de tirage aléatoire et de manipulation d'individus dans des listes. Elle facilite également la production des résultats finals. En effet, très souvent, dans une simulation, on obtient comme résultat le fichier brut de la population tel qu'il a été modifié par les paramètres du programme et il faut donc ensuite passer par une étape de tabulation des résultats à l'aide de logiciels statistiques extérieurs au programme de simulation.

Dans le cas présent, les procédures de tabulation peuvent être introduites directement dans le programme. En sortie, on obtient alors immédiatement la série retraçant l'évolution de la variable retenue (par exemple, effectif dans les grades, âges moyens ou promotions...) tout au long de la durée de simulation. Des procédures graphiques permettent également de visualiser de façon simple et immédiate ces évolutions.

C'est ce même système de tabulations qui permet de suivre la carrière d'une ou plusieurs personnes. Une variable indicatrice permanente que l'on nomme « marqueur » peut en effet être utilisée pour repérer un individu ou un groupe d'individus...) avec filtrage par cette indicatrice, on obtient une description automatique de la situation de l'individu ou du groupe au sein de l'entreprise (sa carrière).

d) Clonage et suivi des carrières individuelles

La faible taille de l'échantillon simulé introduit évidemment une forte sensibilité des résultats de la simulation à la séquence des tirages aléatoires sur laquelle elle s'appuie, d'où la nécessité de simulations jointes ou répétées, pour deux raisons :

- Ces simulations donnent une image plus plausible de l'évolution future moyenne de la population.
- La multiplication des simulations indépendantes est le moyen de donner une idée de la dispersion des futurs possibles, que ce soit au niveau agrégé (par exemple, distribution de l'âge moyen dans tel ou tel grade) ou individuel (distribution de l'état dans lequel se trouve l'individu à une date future).

Pour ce faire, deux options sont a priori possibles :

- Une première méthode consiste à démultiplier l'échantillon initial de façon à obtenir une population plus large à laquelle sera appliquée une simulation unique. Cette démultiplication peut être réalisée par une méthode de clonage ou de pseudo-clonage (duplication stricte d'individus existants dans le fichier initial, ou tirage de nouveaux

individus par extra ou interpolation des caractéristiques des individus de ce fichier). L'intérêt de cette méthode est qu'elle ne suppose qu'une modification du fichier initial, sans qu'il y ait à intervenir sur la phase de simulation ou de traitement des résultats, qui sera identique à ce qu'elle aurait été sur le fichier de départ. Cette méthode est celle qui a été utilisée pour la production de tous les résultats suivants qui sont exprimés en termes de moyenne.

- Cette méthode n'est cependant pas la plus satisfaisante car elle peut fausser le jeu des contraintes qui tiennent à la taille absolue : les perspectives de carrière et les risques de blocage ne sont pas les mêmes, *a priori*, dans un corps de taille N et un corps de même structure mais de taille pN avec $p > 1$. Une prise en compte correcte de ces contraintes suppose donc plutôt la réalisation successive de p simulations indépendantes, sur les résultats desquelles on procède ensuite à des calculs de moyennes ou de variances. Cette méthode est celle qui a été utilisée pour ceux des résultats qui suivent et qui sont exprimés en termes de dispersion ou de distribution de probabilités (par exemple, probabilité d'être au grade i à la date t).

3. Résultats

Les résultats qui vont suivre sont une sélection d'un ensemble de résultats présentés par ailleurs (Pennec, 1994). On va se limiter à exposer des simulations relevant de la logique « tirée », avec un jeu d'hypothèses donné par le tableau 1. Parmi ces hypothèses, on va privilégier les variantes sur le taux de croissance des effectifs désirés et sur les modalités de sélection des promus. Les premières distinguent des hypothèses de croissance plus ou moins rapide des effectifs. Les secondes opposent deux modalités de promotion : dans le premier cas, lorsque les places vacantes autorisent la promotion de n personnes du grade i au grade $i-1$, les n personnes retenues sont celles qui ont l'ancienneté la plus forte dans le grade i . Dans le second cas, les n personnes sont tirées de façon totalement aléatoire parmi l'effectif du grade i . On notera que ces deux cas correspondent aux deux cas extrêmes d'une modalité de sélection plus générale, prévue par le programme, qui autorise n'importe quel degré de panachage entre ancienneté et choix², procédure qu'il aurait été évidemment difficile de mettre en œuvre dans le cadre d'un modèle de projection traditionnel.

a) Âge moyen dans les grades et taux de croissance

Pour évaluer les éventuelles répercussions des modifications de paramètres, nous nous attacherons ici principalement aux âges moyens. Les résultats en termes d'effectif ne sont pas présentés, puisqu'ils seront la traduction directe des hypothèses d'évolution exogène qui leur sont affectées. Précisons simplement que les hypothèses de variations retenues sont l'accroissement nul pour la plus basse, +1% pour la moyenne (référence) et +5% pour la plus élevée (cette hypothèse d'école étant évidemment irréaliste).

Les figures 3-1 à 3-3 donnent les évolutions des âges moyens dans les trois grades selon les trois scénarios de croissance des effectifs. L'hypothèse retenue en ce qui concerne les

² Plus précisément, l'algorithme commence systématiquement à ordonner la population du groupe i par ancienneté décroissante. Dans le cas d'un tirage aléatoire (au choix), on tire la promotion du plus ancien avec une probabilité n/p , puis la promotion de tous les suivants avec une probabilité égale au rapport entre le nombre de postes encore à pourvoir et le nombre d'individus restant à examiner dans la population i . Cette procédure garantit effectivement un tirage aléatoire sans remise de n individus parmi les p candidats. Dans le cas de la promotion à l'ancienneté, le tirage consiste à nouveau à balayer la liste par ancienneté décroissante, en promouvant les individus avec la probabilité 1 tant qu'il reste des postes à pourvoir. Les cas intermédiaires sont obtenus en tirant, au rang k , une promotion avec la probabilité n/p tant qu'il reste des postes à pourvoir. Le cas = 1 permet ainsi de réobtenir la règle d'avancement à l'ancienneté et le cas = 0 redonnant le cas de promotion au choix.

promotions est celle de promotions à l'ancienneté. S'agissant du dernier grade, ceci y implique un âge moyen extrêmement élevé (qui tient en partie à une hypothèse très haute sur l'âge de départ en retraite, supposé pouvoir monter à 68 ans pour cette catégorie), et une durée de séjour d'environ 6 ans, soit environ 15% de la vie active totale en régime stationnaire, qui correspond bien à l'effectif relatif de ce groupe. L'hypothèse haute de croissance très rapide (au demeurant irréaliste à long terme) permet un abaissement de cet âge moyen qui paraît faible à première vue (seulement trois années), mais qui est néanmoins très important en termes relatifs, puisque ceci correspond à un doublement du temps passé dans cet état.

Le même rajeunissement est observé dans les autres grades, et mélange à la fois l'effet du rajeunissement général de la pyramide des âges et les modifications des âges à la promotion dont les graphiques agrégés ne sont pas donnés, mais qui vont pouvoir être appréciés à partir des graphiques relatifs à des profils de carrière individuels.

Bien que nous ayons retenu la promotion à l'ancienneté comme hypothèse de référence principale, nous avons calculé, pour les paramètres de référence, les différences qu'entraîne l'utilisation de l'un ou l'autre des deux modes de promotions possibles sur les évolutions de la population étudiée. Assez logiquement, les âges moyens des grades supérieurs et les âges moyens à la promotion sont plus élevés lorsque la promotion est à l'ancienneté que lorsqu'elle se fait au choix (résultats non fournis sur les graphiques). Ce phénomène doit évidemment être compensé au niveau du grade 3, puisque la modification des règles de promotion ne change pas la structure globale tous grades confondus. De fait, l'âge moyen y est plus bas.

FIGURE 3-1 : MODÈLE TIRÉ - ÂGE MOYEN DU GRADE 1 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF

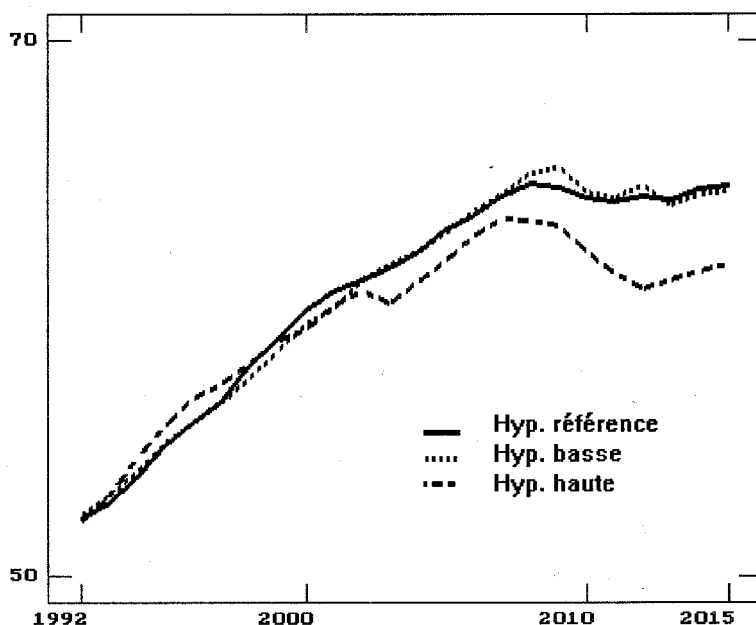


FIGURE 3-2 : MODÈLE TIRÉ - ÂGE MOYEN DU GRADE 2 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF

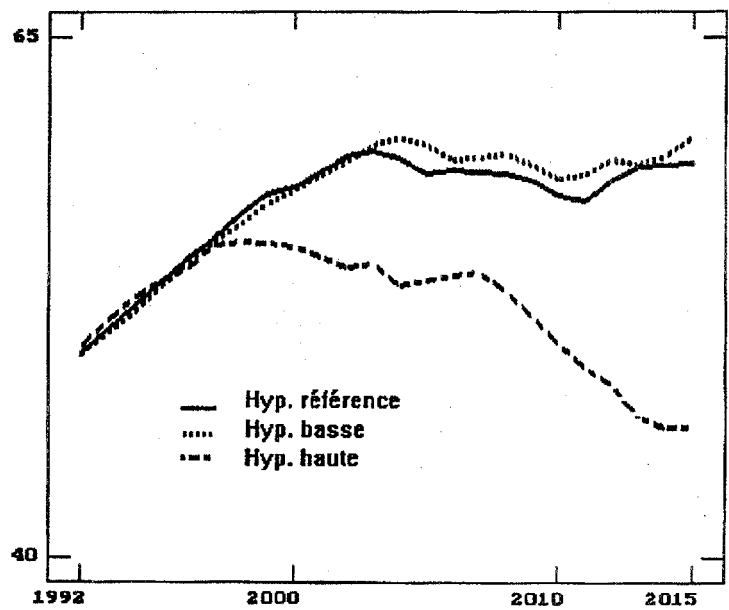
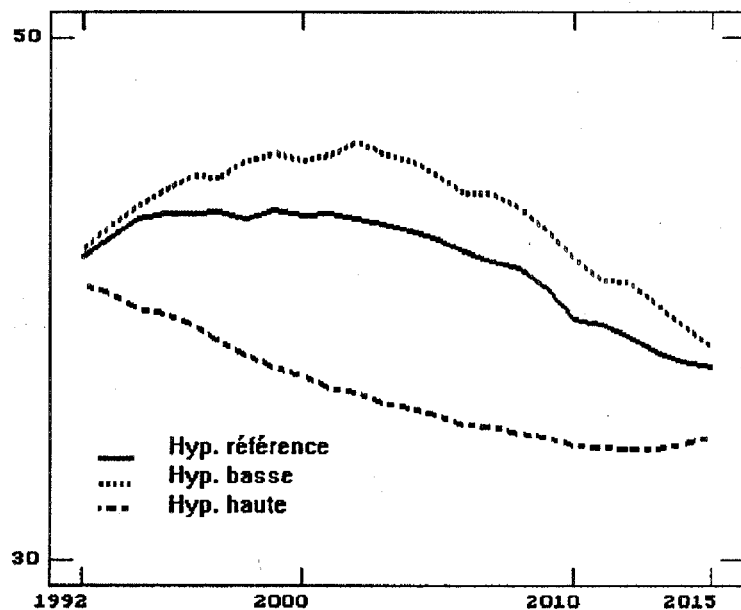


FIGURE 3-3 : MODÈLE TIRÉ - ÂGE MOYEN DU GRADE 3 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF



TABEAU 1 : PARAMÈTRES DES DIFFÉRENTES SIMULATIONS DU MODÈLE TIRÉ

	Effectif	Départ autonome (1)	Départ supplém.	Recrutement autonome	Recrutement supplém.	Promotion autonome	Promotion supplém.	Quota promotion supplém. (2)	Priorité ancienneté (3)
Simulation 1.1	0 %	5 % grade 1 10% grades 2 et 3	oui	non	Grade 3	non	oui	100%	0 1
Simulation 1.2	+ 1 %		oui	non	Grade 3	non	oui	100%	0 1
Simulation 1.3	+ 5 %		oui	non	Grade 3	non	oui	100%	0 1
Simulation 2.1	+ 1 %	5 % grade 1 10% grades 2 et 3	non	non	Grade 3	non	oui	100%	0 1
Simulation 3.1	+ 1 %	5 % grade 1 10% grades 2 et 3	oui	non	Tous les grades	non	oui	50%	0 1
Simulation 4.1	+ 1 %		oui	non	Grade 3	non	oui	50%	0 1

b) Profils individuels de carrière

Ces profils sont établis à l'aide de la procédure de marquage expliquée plus haut. A titre d'illustration, on a marqué un individu particulier. Une tabulation du grade avec filtrage sur ce marqueur, réalisée aux différentes dates sur un jeu de 100 simulations indépendantes donne alors une image directe des *probabilités*, pour cet individu d'être dans les grades 1, 2 et 3 aux différentes dates. Seule l'approche par micro-simulation permet la production simple de ce type de résultats, à la fois probabilistes et établis au niveau individuel.

Les graphiques 4 et 5 donnent l'évolution de la carrière selon les deux types de promotions dans l'hypothèse moyenne. Si la promotion se fait à l'ancienneté, l'agent a plus de 8 chances sur 10 d'être promu au grade 2, mais aucune d'atteindre le grade 1. Si elle est au choix, les probabilités d'atteindre les grades 2 et 1 sont respectivement de 16% et 7%. Autrement dit, l'individu a plus de chances d'atteindre le grade 2 si les promotions se font à l'ancienneté mais comme avec la configuration de la population, il va y arriver à un âge élevé, il n'aura pas le temps d'atteindre le grade 1 avant son départ à la retraite. En revanche, avec la promotion au choix, il peut soit y gagner soit y perdre par rapport à la promotion à l'ancienneté. Il y perd car sa probabilité de ne jamais être promu est plus élevée qu'avec la promotion à l'ancienneté, mais à l'inverse il y gagne car sa promotion peut intervenir à un âge plus jeune et de ce fait, il peut par la suite prétendre atteindre le grade 1.

Nous allons à présent nous attacher à cerner l'influence que peut avoir une modification du taux de croissance des effectifs par grade sur la carrière d'un individu. Lorsque l'on suppose une croissance nulle, on observe peu de répercussion sur la carrière de la personne étudiée. Par contre, une augmentation des effectifs de 5% par an, entraîne une probabilité plus forte d'être promu au grade supérieur. Les figures 6-1 à 6-3 comparent la répercussion du changement de taux de croissance des effectifs sur la carrière de l'individu. Passer d'une croissance de 1% à la stationnarité prolonge d'environ 3 ans la durée de séjour dans le groupe 3, tandis qu'une croissance de 5% réduit cette même durée de 5 ans (figure). La période moyenne de séjour dans le groupe 2, illustrée par la figure 6-2 est avancée en conséquence. En revanche, elle n'est pas plus courte que dans le scénario de croissance limitée à 1%.

FIGURE 4 : MODÈLE TIRÉ - PROBABILITÉ D'ÊTRE AU GRADE 1, 2 ET 3
HYPOTHÈSE DE RÉFÉRENCE (PROMOTION À L'ANCIENNETÉ)

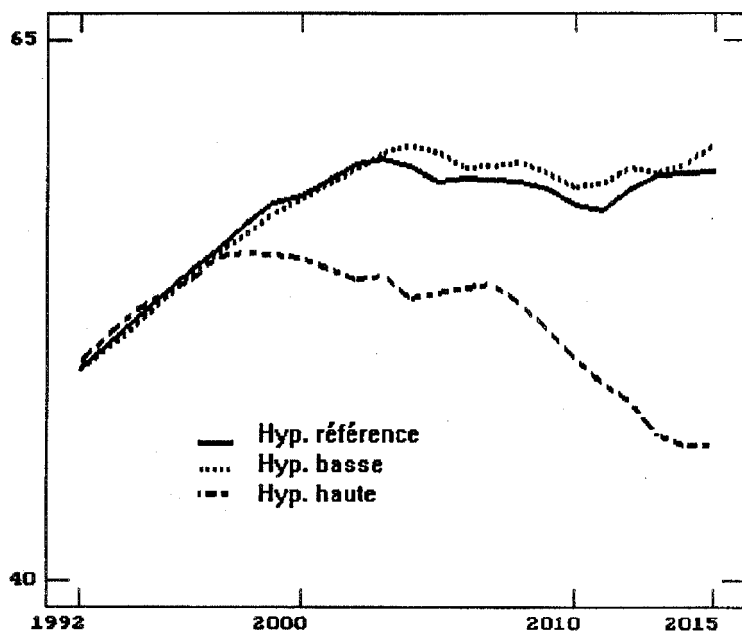


FIGURE 5 : MODÈLE TIRÉ - PROBABILITÉ D'ÊTRE AU GRADE 1, 2 ET 3
HYPOTHÈSE DE RÉFÉRENCE (PROMOTION AU CHOIX)

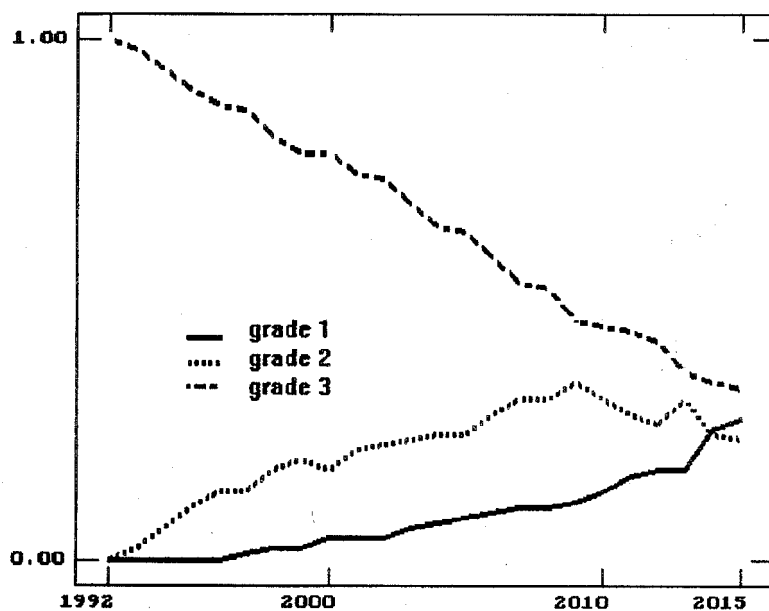


FIGURE 6-1 : MODÈLE TIRÉ - PROBABILITÉ D'ÊTRE AU GRADE 3 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF (PROMOTION À L'ANCIENNETÉ)

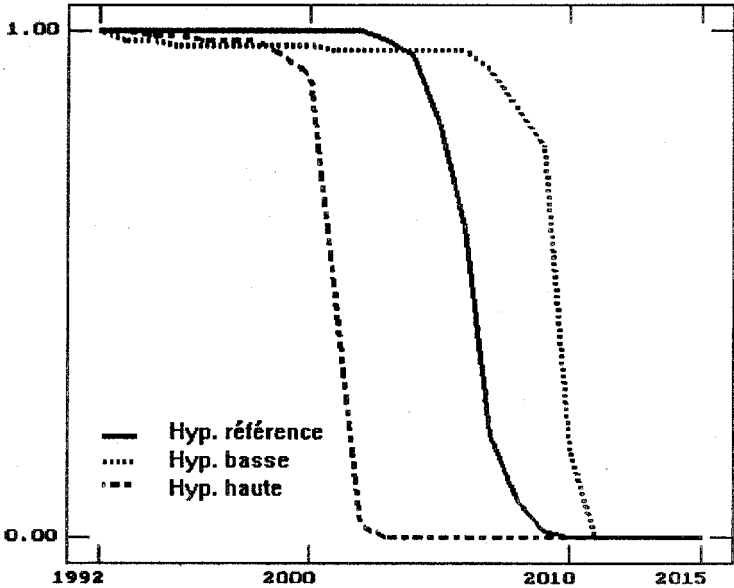


FIGURE 6-2 : MODÈLE TIRÉ - PROBABILITÉ D'ÊTRE AU GRADE 2 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF (PROMOTION À L'ANCIENNETÉ)

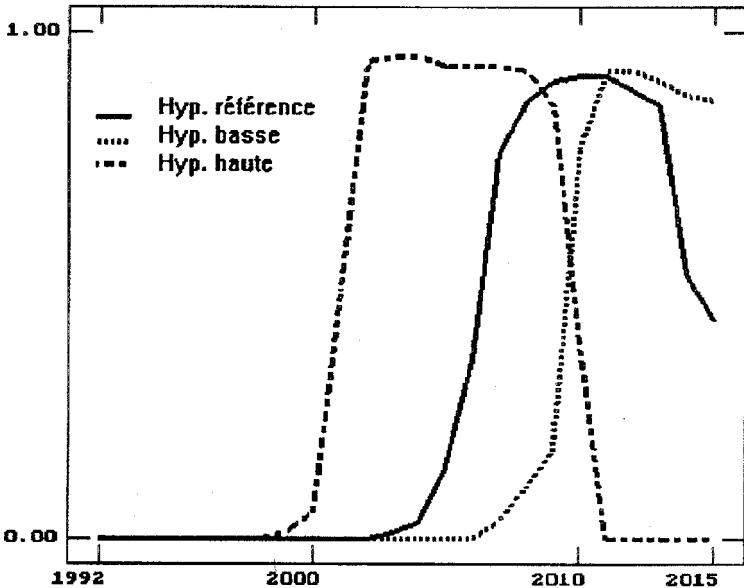
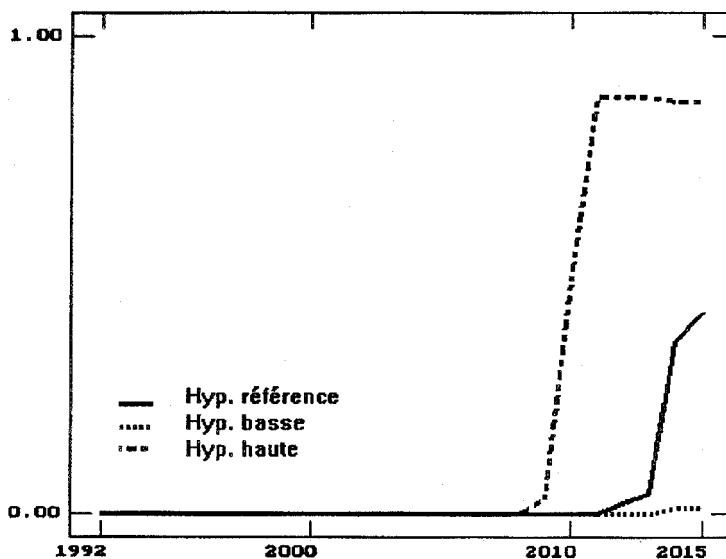


FIGURE 6-3 : MODÈLE TIRÉ - PROBABILITÉ D'ÊTRE AU GRADE 1 LORSQUE L'ON MODIFIE LES TAUX DE CROISSANCE DE L'EFFECTIF (PROMOTION À L'ANCIENNETÉ)



Conclusion

Bien évidemment, l'intérêt de l'outil de micro-simulation n'est pas d'avoir permis la démonstration des effets qualitatifs qu'on vient de lister, qui sont bien connus et intuitifs. Il est de permettre leur quantification systématique, et leur illustration concrète au niveau d'individus bien identifiés, en incluant aussi de larges possibilités de refléter au mieux la complexité des règles d'évolution des carrières qui gouvernent les populations professionnelles réelles. Le but à atteindre, dans cette perspective, serait la constitution d'un outil flexible et général, facilement reparamétrable pour s'adapter à la description de situations concrètes.

BIBLIOGRAPHIE

- BLANCHET, D. et KESSLER, D., (1990), *Programme de perspectives d'un système de retraite professionnel - Manuel d'utilisation*, Fondation pour la recherche économique et financière, 35 p.
- BLANCHET, D., (1993), *Outils informatiques pour la micro-simulation*, Séminaire « Outils et Méthodes », INED, janvier.
- BONNEUIL, N., (1993), Support du cours de gestion du personnel, DEA de démographie économique, IEP, 16 p.
- BOUEYERE, D., (1982), *Le personnel des PTT face aux nouvelles technologies - Perspectives de recrutement et évolution du corps*, Thèse de doctorat de 3ème cycle en démographie, Université de Paris V-René Descartes, 321 p. + annexes.

- COURONNE, P., (1986), *Modèles de gestion prévisionnelle des ressources humaines - Le cas du CNRS*, CNRS, 602 p.
- DINH, Q. C. et LABAT, J. C., (1986), *Projection de la population totale pour la France 1985-2040*, INSEE, Collections de l'INSEE, Série D n° 113, 93 p.
- GENDREAU, F., (1987), La démographie des chercheurs de l'ORSTOM. Analyse et perspectives in *Les projections démographiques*, Actes du colloque, INED, Travaux et documents n°122, pp. 127-145.
- HENRY, L., (1971), Pyramides, statuts et carrières. I. Avancement à l'ancienneté. Sélection, *Population*, n° 3, pp.463-486.
- HENRY, L., (1972), Pyramides, statuts et carrières. II. Avancement au choix, *Population*, n° 4-5, pp. 599-636.
- HENRY, L., (1975), Perspectives d'évolution du personnel d'un corps, *Population*, n° 2, pp. 241-270.
- HENRY, L., (1976), Pyramides, statuts et carrières. III. Corps de petits effectifs, *Population*, n° 4-5, pp. 840-855.
- KEYFITZ, N., (1973), Individual mobility in a stationnary population, *Population Studies*, 27, pp. 335-352.
- LARRIVÉE, D., (1988), Présentation d'un modèle de projection de main-d'oeuvre pour l'entreprise, *Cahiers Québécois de Démographie*, volume 17, n° 2, pp. 233-245.
- MALLET, L. (1991), *Gestion prévisionnelle de l'emploi*, Éditions Liaisons, 200 p.
- MUNOZ-PEREZ, F., TRIBALAT, M., (1993), Perspectives de carrières et évolution du corps des magistrats 1990-2030, *Population*, n° 1, pp. 27-62.
- NORVEZ, A., (1977), *Le corps enseignant et l'évolution démographique. Effectifs des enseignants du second degré et besoins futurs*, INED, Travaux et Documents, Cahier n° 82, 206 p.
- PENNEC, S., (1994), *Applications démo-économiques de la méthode des micro-simulations*, Thèse de doctorat, Institut d'Études Politiques de Paris, 376 p.
- PERETTI, J. M., (1987), *Gestion des ressources humaines*, Éditions Vuibert, 268 pages.
- STEWMAN, S., (1986), Demographic models of internal labor markets, *Administrative Science Quaterly*, n° 31, pp. 212-247.
- VILTER, S., (1991), Vieillissement dans l'entreprise - Théorie et modèles, *Risques*, n° 5, pp. 147-163.

Instruments d'analyse démographique et mesure de la satisfaction du consommateur

Michele COSTABILE

Université de Calabre (Cosenza)

Département Gestion d'Entreprise et Administration Publique

Gianluca MARZOCCHI

Université de Bologne

Département Économie d'Entreprise.

1. Introduction

Au cours des vingt dernières années, l'orientation stratégique de fond des entreprises occidentales s'est peu à peu déplacée du marketing à la qualité totale, à la satisfaction du consommateur, qui devient composante essentielle des ressources immatérielles sur lesquelles l'entreprise base son succès. Il ne s'agit pas, semblerait-il, d'un changement dû à la mode, mais bien d'une évolution naturelle vers la recherche d'avantages compétitifs nouveaux et défendables. (Nelson et Winter, 1982 ; Itami, 1987 ; Winter, 1987 ; Vicari, 1988 ; Prahalad et Hamel, 1990 ; Grant, 1991 ; Baccarani, 1991 ; Valdani et Busacca, 1992 ; Teece, Pisano et Shuen, 1992 ; Rullani, 1993).

En effet, l'augmentation de la pression compétitive, le développement et la diffusion des nouvelles technologies, l'évolution des préférences et des expectatives de la demande, ont encouragé les entreprises à rechercher le renforcement de leur position sur le marché en se basant sur le développement de ressources immatérielles difficiles à imiter par la concurrence. Satisfaction, confiance et fidélité de la clientèle représentent sans doute les ressources immatérielles les plus importantes parmi toutes celles engendrées par les investissements en marketing. Avoir une clientèle très satisfaite et fidèle a des conséquences très positives sur le plan économique et de la gestion de l'entreprise. (Valdani, Busacca et Costabile, 1994 ; Busacca, 1994-a et 1994-b ; Busacca et Costabile, 1994).

Les principes et les modèles de la satisfaction du consommateur sont largement diffusés et semblent amplement partagés par les équipes de gestion des entreprises. Pourtant, le patrimoine méthodologique nécessaire à son développement et à sa gestion est encore à définir. Si nous considérons à cet effet, le stade de développement des connaissances, de nombreux rapprochements interdisciplinaires sont en phase d'expérimentation ; certains d'entre eux semblent d'ailleurs mériter une analyse plus approfondie.

Un des principaux problèmes posés par la gestion et l'amélioration de la satisfaction du consommateur est sans doute la façon de la mesurer. Les études jusqu'ici réalisées ont mis en évidence de nombreuses approches méthodologiques et des indicateurs, relevés sur le terrain ou bien développés à partir des données contenues dans les systèmes d'information des entreprises. Parmi les indicateurs obtenus le plus facilement par les bases de données des entreprises, nous trouvons ceux relatifs à la dynamique des clients existants et à leur fidélité, ce qui représente la dimension comportementale, fondamentale dans la satisfaction du consommateur.

Ce rapport a pour objectif de proposer l'application à l'étude et à la mesure des dynamiques du portefeuille client de certains instruments liés à l'analyse démographique. Cette proposition est basée sur l'analogie qui existe entre des phénomènes démographiques comme la

natalité, le vieillissement ou la mortalité et l'entrée, le séjour et la sortie des clients du portefeuille clients de l'entreprise.

Nous nous concentrerons tout particulièrement sur les indicateurs de calcul de la satisfaction du consommateur qui peuvent être définis sur la base de données secondaires, c'est-à-dire de ces données extraites couramment du système d'information des entreprises même si elles ont une autre finalité plus immédiate. Ces indicateurs peuvent être le taux de fidélité, l'âge moyen prévisionnel du portefeuille clients et leur valeur patrimoniale (life time value). Ils représentent en effet un potentiel d'information formidable, parfois prévisionnel, des indicateurs de performance conventionnels de l'entreprise (part de marché, rentabilité, etc.) et lorsqu'intégrés aux indicateurs de satisfaction et de fidélité de la clientèle, du niveau de ses ressources immatérielles en marketing.

Depuis toujours, les indicateurs de performance d'une entreprise ont été d'autant plus pris en compte que leur portée informative a été immédiate et fiable pour les équipes de gestion aux commandes de l'entreprise. A ce propos, il faut souligner que la valeur de ces indicateurs n'est presque jamais absolue et qu'afin de pouvoir affirmer leur validité informative, il faut effectuer quelques comparaisons interentreprises, intersectorielles et transversales. C'est justement pour ces analyses et ces comparaisons que pourront être employés certains instruments d'enquête démographique. En effet les problèmes les plus importants dérivent des choix méthodologiques pour la construction de ces indicateurs, et influencent leur fiabilité scientifique et gestionnelle. Il semblerait tout particulièrement nécessaire de se concentrer sur leur valeur relative synchronique, c'est à dire sur la validité des comparaisons interentreprises et intersectorielles, et sur leur valeur relative diachronique, c'est-à-dire sur la fiabilité des analyses transversales. Sous le profil méthodologique, il s'agit d'appliquer à l'étude des dynamiques du portefeuille client des instruments d'analyse qui devraient permettre d'isoler l'effet des différentes composantes des phénomènes de rétention et de fidélité : dans le cas, par exemple, de portefeuilles clients composés de différentes durées de la relation avec l'entreprise ou bien dans le cas d'un « effet durée » qui influence la variabilité du taux de fidélité dans le temps, et ainsi de suite.

En réponse justement à ces exigences, nous proposons, du moins de façon expérimentale, l'application de l'analyse par populations et de la standardisation démographique.

Dans le premier cas, en ce qui concerne les études transversales du portefeuille clients, nous pouvons prévoir l'utilisation d'instruments analytiques utilisés en démographie, afin de distinguer les « effets de durée » des « effets de populations » et des effets dus à l'ancienneté de la relation. Il s'agit en effet de mesurer séparément les effets d'un phénomène sur les différentes couches de la population qui pourraient réagir différemment à un même événement suivant une ancienneté différente de la relation avec l'entreprise.

En revanche, la standardisation démographique a pour objectif de rendre comparable deux populations de clients avec des structures différentes en ce qui concerne l'ancienneté de la relation avec l'entreprise, et garantit ainsi la significativité de la comparaison entre indicateurs d'analyse de plusieurs populations « clientèle d'entreprise ».

2. Les conditions de la centralité retrouvée par le client

Dans le cadre des disciplines qui ont pour objet l'étude des comportements et des relations d'échange, il peut paraître inutile, et dans une certaine mesure redondant, de souligner la position centrale occupée par des problématiques relatives à la mesure et à la gestion de la satisfaction de la clientèle.

Définir la satisfaction des besoins de la clientèle (satisfaction du consommateur) comme objectif final des actions d'une organisation représente en effet le fondement même de cette révision radicale des rapports entreprise-marché qui se traduit le plus concrètement du monde

dans le marketing (compris comme discipline d'étude et de processus social). De nombreux exemples montrent comment la plupart des entreprises en activité incorporent de façon explicite une telle orientation dans leur projet (mission statement) et dans leur stratégie de communication.

D'un point de vue plus général, nous ne devons pas être étonnés d'une telle convergence d'opinions : aucune organisation ne peut survivre dans le long terme sans satisfaire les besoins de ses propres clients. Tout simplement les individus n'achètent pas (ou ne rachètent pas) ce dont ils n'ont pas besoin.

La satisfaction du consommateur devient par conséquent l'objectif principal de chaque entreprise, la condition nécessaire (mais pas toujours suffisante) de sa propre survie. En conséquence, chaque intervention sur le marché doit avoir pour origine une pointilleuse connaissance des exigences de sa propre clientèle objectif, exigences qui détermineront l'ensemble des décisions et des comportements qui détermineront l'offre d'un produit service adéquat. La satisfaction de la clientèle cimente ainsi d'une certaine façon entre elles les différentes fonctions organisatives - elle représenterait d'après certains un langage entrepreneurial (Busacca 1994-b) - elle en représente un dénominateur commun et même un paramètre de référence auquel doit se mesurer l'efficacité de chaque initiative.

Comme nous avons déjà pu le remarquer, d'un point de vue philosophico-conceptuel, l'importance de la satisfaction du consommateur est telle que sa diffusion immédiate et complète au niveau de l'entreprise est retenue comme nécessaire. Afin de mieux comprendre l'actualité du sujet, il semblerait utile d'élargir notre cadre analytique en y intégrant quelques considérations de caractère plus particulièrement opérationnel.

Le concept de marketing et le caractère central des besoins du client et de leur complète satisfaction, commencent à se diffuser dans les années 50. Il s'est écoulé depuis un laps de temps apparemment suffisant pour permettre une sédimentation presque complète de la culture et des techniques de gestion qui leur sont attachées. En fait, des événements importants ont rendu au thème de la satisfaction de la clientèle son importance et sa vivacité. Le riche débat qui s'est développé à la fin des années 70 (et qui est encore d'actualité) à partir de concepts comme la qualité totale et la qualité globale, en est un exemple efficace.

Actuellement le regain d'attention, non seulement pour la modélisation du concept de la satisfaction du consommateur, mais aussi pour l'élaboration des techniques et des instruments qui permettent de l'étudier, confirme cette tendance d'un renouveau d'intérêt pour cet argument.

Nous pouvons interpréter cette situation de deux façons : la première a une certaine connotation critique et fait appel aux difficultés sous-jacentes à une complète acceptation du concept de la satisfaction du consommateur comme objectif de l'entreprise ; la deuxième est plus positive et témoigne en revanche de l'importance de cette thématique et d'un effort de recherche continu visant à élaborer des aspects toujours plus raffinés de ce phénomène, cohérents par ailleurs avec les profonds changements des micro variables compétitives qui ont caractérisé les dernières années.

Il s'agit de phénomènes qui doivent être rappelés, même brièvement. Il s'agit tout d'abord de la progressive déshomogénéisation des marchés dont les contours sont toujours moins définis et nets, dans un contexte d'apparition de micro-agrégats de clients fortement spécifiques, porteurs de besoins et de valeurs hautement différenciées. Le second phénomène, qui est en quelque sorte lié au premier, est l'émergence de la part du client d'une large demande de participation et d'implication dans la définition des paramètres de l'offre : en d'autres termes, d'une demande d'interaction et d'individualisation d'un rapport caractérisé auparavant par une standardisation élevée. Le troisième phénomène est le passage à un état de maturité des principaux secteurs industriels (automobile, informatique, électronique de grande consommation), tout en mettant parallèlement l'accent sur la défense des parts de marché détenues et sur l'utilisation de la clientèle

déjà existante. Il faut rappeler à ce propos que dans un secteur mûr, le comportement du consommateur peut prendre deux aspects : la banalisation de la consommation, et par conséquent une tendance à l'infidélité et la différenciation des préférences et donc des expectatives plus importantes. Dans tous les cas, essayer d'augmenter la satisfaction du consommateur et les indicateurs de fidélité devient un objectif prioritaire de l'entreprise.

Même si elle est incomplète, cette brève analyse réussit à donner une première explication à la récente attention portée à la satisfaction du consommateur dans le cadre du rapport client-fournisseur : ce n'est plus un instrument de mise en application de « l'orientation générale au marché » de la fonction commerciale, mais c'est un indicateur de la capacité de l'entreprise à construire et à maintenir une relation-interaction avec la clientèle, non plus considérée dans un sens agrégé mais comme un ensemble d'éléments unitaires. Il est encore plus intéressant, pour nous tout au moins, de constater comment l'effet conjugué des phénomènes préexposés a contribué à un changement important du système des valeurs existant dans le concept de marketing : auparavant l'accent était surtout mis sur la volonté de satisfaire les besoins d'un nombre maximum de clients prêts à acheter, alors que la tendance s'est maintenant déplacée vers la maximisation de la satisfaction (et donc de la fidélité) des clients existants. On passe de la création d'acheteurs vers celle de relations. La « customer retention » c'est à dire la capacité de maintenir et de consolider son propre patrimoine clients de façon absolue et donc a fortiori vis-à-vis de la concurrence, devient l'indicateur le plus immédiat de l'efficacité de son intervention sur le marché.

Deux exemples peuvent faciliter la compréhension des processus mis en place, tout en suggérant en même temps le terrain où un tel changement a trouvé son application la plus immédiate. L'utilisation de ce que l'on appelle « loyalty programs », programmes de loyauté, est de plus en plus répandu et ils jouent même un rôle d'entraînement dans des secteurs comme la distribution et les services de masse. Comme le nom l'indique, il s'agit de programmes d'encouragement à l'utilisation de produits ou de services précis de la part de ceux qui déjà les achètent. Nous connaissons les hypothèses théoriques d'une telle approche : la rentabilité du client n'est pas constante face à l'ancienneté de la durée du service fourni par l'entreprise, mais au contraire elle augmente proportionnellement à la durée du rapport (sa fidélité). Les motivations résident dans le faible impact des coûts fixes d'acquisition par client, dans les faibles coûts de gestion de chaque client dus à l'expérience, dans le volume accru d'achats par unité de temps, dans une plus grande disponibilité de la part de l'entreprise à l'adoption de nouvelles propositions (cross-buying). Au delà d'une incitation à l'achat adressée à tout le parc de clients, l'objectif de tels programmes est de toucher tout particulièrement ces groupes de clients qui sont des « grands utilisateurs » et qui se caractérisent par des volumes d'achat supérieurs à la moyenne. Ces initiatives sont supportées par les « Customer Information Files », bases de données clients, (cif), dont l'utilisation dépasse en réalité la gestion d'un simple loyalty program, au point de faire apparaître nécessaire son développement autonome. Même si sur un plan formel il s'agit d'une réédition sophistiquée du concept classique du fichier clients, la réalité de la mise en place s'en éloigne fortement. Une fois dépassé le simple objectif de la description de l'identité comptable du client, le concept de celui-ci s'enrichir désormais a de nouvelles connotations qualitatif-comportementales : nous pouvons imaginer pour chaque client la construction d'une fiche avec ses préférences établie d'après ses achats antérieurs, nous pouvons estimer le volume futur de ses achats, au cours de l'année comme au cours de sa vie commerciale (le customer lifetime value) et il devient possible de mettre en place des procédures de routine automatiques pour le ralentissement de la fréquence d'achat. En un mot, il devient possible de traduire sur un plan d'action cette « mass customization » qui a été plus d'une fois mise en évidence comme une des principales expressions des changements actuels dans la relation offre-demande (Pine, 1993).

C'est justement dans cette direction que nous pouvons identifier de nouvelles opportunités de rapprochements entre démographie et marketing. Jusqu'à présent, les points de contacts entre les deux disciplines se sont limités, comme nous le savons, à deux grands domaines : le décodage des principales macrotendances démographiques à l'aide de clés d'évaluation des opportunités et menaces de l'environnement (la population est utilisée dans un sens quantitatif) ; la définition de politiques de segmentation (la population est ici entendue dans un sens qualitatif). Le premier point de vue a des conséquences immédiates, non seulement sur le front de l'estimation des volumes agrégés de la demande, mais aussi sur la gestion de la dynamique des marchés, actuels ou futurs (parmi les exemples les plus immédiats, nous pouvons citer les conséquences d'une variation du taux de natalité, ou la présence non plus marginale des minorités ethniques et religieuses, ou les changements dans la composition du noyau familial). Le second point de vue (la segmentation), permet une lecture plus raffinée de ces phénomènes. Et ce, non seulement en termes d'utilisation directe des indicateurs démographiques comme variables de segmentation (approche entièrement consolidée et certainement utile en première phase de lecture du marché, mais souvent incapable de rendre clairs des phénomènes qui déterminent les vraies motivations d'achat), mais aussi comme clés d'explication de leur utilisation comme « descripteurs » de segments définis selon des variables fortement corrélées aux bénéfices recherchés du produit par les individus qui composent ces populations (benefit segmentation).

Comme nous pouvons le remarquer, l'utilisation de ces techniques démographiques est cohérente avec la vocation historique du marketing décrite ci-dessous : la satisfaction des besoins d'un maximum de clients potentiels. En ce sens, le rôle précurseur de la démographie est un résultat certainement déterminant. Si nous reconsidérons pourtant le rapport entre démographie et marketing à la lumière de la nouvelle importance accordée à la tutelle du patrimoine représenté par les relations de long terme avec les clients existants, alors c'est l'objet même de l'analyse démographique qui change d'aspect. L'objet de l'enquête n'est plus simplement la population, mais aussi ce sous-ensemble déterminé par la clientèle existante au moment considéré. Il va de soi que même l'unité d'analyse change : ce n'est plus l'individu « biologique » mais son équivalent virtuel représenté par le concept de « client ».

Les analogies sont évidentes mais il semble pourtant nécessaire de les souligner : l'acquisition d'un nouveau client équivaut à la « naissance », la fin du rapport à la « mort », le déplacement représente la « migration ». Pour chacun de ces événements, il est possible de construire des tableaux capables de fournir les valeurs caractéristiques de cette même population (en termes d'espérance de vie, de taux de mortalité, de distribution des causes de mortalité, etc.) soit au niveau de l'entreprise, soit à celui du secteur. Une fois ce parallèle accepté, le nombre d'instruments démographiques applicables à l'analyse du portefeuille clients devient significatif et consistant. Cette approche ne présente pas en soi, et d'ailleurs ce n'est pas son intention de le faire, les traits d'une révolution méthodologique. Ce qui pourtant constitue un élément nouveau est la possibilité d'appliquer des techniques d'analyse démographique à des contextes organisationnels dans lesquels la disponibilité technologique permet de maintenir un suivi individuel de chaque client, dépassant ainsi la dimension agrégée typique de cette discipline.

Il est évident que le terrain dans lequel un tel effet conjugué peut donner des résultats immédiats est celui qui caractérise les secteurs dans lesquels des liens de nature structurelle ont depuis longtemps imposé un investissement dans des bases de données clients plus ou moins sophistiquées : banques, assurances, compagnies aériennes, gestion des cartes de crédit, éditeurs de magazines sont quelques uns des exemples les plus immédiats en ce qui concerne les biens de consommation ; à son tour, l'univers entier du business-to-business répond de façon physiologique à de tels requis.

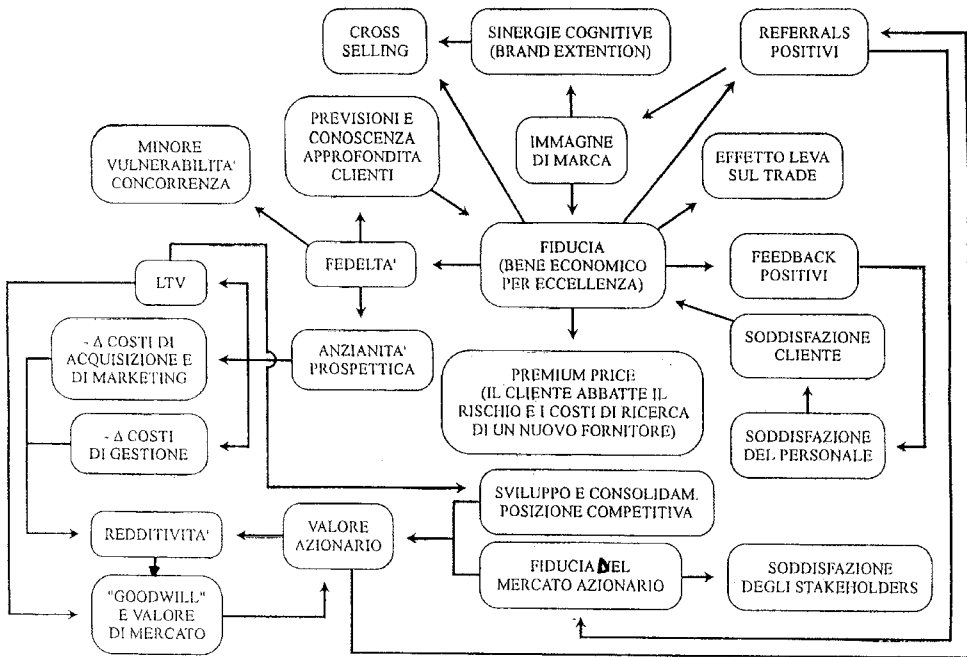
3. Quelques indicateurs clé de la satisfaction du consommateur

Sur la base de ce qui a été dit dans le dernier paragraphe, nous pouvons soutenir que dans la gestion d'une entreprise, et à plus forte raison dans le cadre des disciplines gestionnelles, la place centrale du client et de sa satisfaction doivent être considérées comme le présupposé fondamental pour le développement et la survie même de l'entreprise. Un tel principe a été récemment renforcé par les conditions évolutives de la demande et de la concurrence dont nous avons déjà parlé.

Le regain d'attention pour la satisfaction du consommateur a d'ailleurs poussé à un examen plus articulé de ses implications positives pour l'économie et la gestion des entreprises. La figure 2.1. présente un diagramme d'influence de la satisfaction du consommateur sur les résultats de l'entreprise pris dans un sens large : résultats economico-financiers, organisation-institutionnels, et résultats concernant la relation avec la demande et son potentiel générateur d'avantages compétitifs.

FIGURE 1 : DIAGRAMME D'INFLUENCE DE LA CS

IL PALLOGRAMMA DELLA SODDISFAZIONE



En synthétisant à l'extrême, les catégories de résultats influencés positivement par la satisfaction de la clientèle peuvent être regroupées comme suit :

1. relations avec la demande et plus généralement pouvoir de marché ;
2. satisfaction, motivation et performance du personnel ;
3. niveau de connaissance des clients, de leurs attentes et des parcours de développement de l'entreprise ;
4. résultats économiques et financiers.

A la lumière d'une telle richesse d'implications, l'orientation vers la satisfaction du consommateur ne peut plus représenter une simple déclaration d'intentions, mais doit évoluer vers une approche conceptuelle et méthodologique articulée qui impliquerait tous les niveaux de gestion de l'entreprise. La nouvelle orientation vers la satisfaction du consommateur se base donc sur un effort constant de mesure du niveau de satisfaction, de la fidélité qui s'en suit dans la relation, de la valeur de cette fidélité et de l'efficacité des processus de l'entreprise qui contribuent à déterminer les processus d'évaluation de la clientèle et les performances économico-compétitives de l'entreprise. Les multiples raisons qui rendent indispensable un recours constant à la mesure de telles ressources de l'entreprise (satisfaction, confiance et fidélité) sont contenues dans les fondements épistémologiques de l'économie d'entreprise. La bonne gestion de l'entreprise requiert en effet comme fondement principal le support effectif de la mesure, c'est-à-dire de ce processus de connaissance des phénomènes de l'entreprise et du marché et de leurs relations dynamiques (Onida, 1971).

Si nous reformulons ce principe plus simplement, il nous est possible d'affirmer que si un phénomène de l'entreprise ou du marché n'est pas mesurable et qu'il a été mesuré, alors il devient difficilement gérable. Ceci autant d'un point de vue strictement tecnico-administratif que de la communication, en ce qui concerne les aspects organisationnels et de marketing, et surtout d'un point de vue entrepreneurial en ce qui concerne l'amélioration stratégique de la position de l'entreprise.

En faisant spécifiquement référence à la satisfaction du consommateur et à ses implications économico-gestionnaires positives, nous pouvons affronter le processus de mesure suivant deux approches : en relevant sur le terrain les domaines de satisfaction / insatisfaction ou en faisant confiance à des données et indicateurs déjà disponibles par le biais du système d'information de l'entreprise. C'est d'une certaine façon compris comme un choix de compromis entre les relevés sur le terrain et l'utilisation des données internes, ce qui représente la solution la plus équilibrée pour obtenir des résultats dignes de foi et d'une portée informative sûre, tout en ne négligeant pas l'efficacité de cette même activité de mesure. A ce propos, il est pourtant indispensable de souligner comment ces diverses modalités d'approche comportent des différences significatives en termes de temps et de coûts de la mesure, de l'utilité des indicateurs obtenus pour la satisfaction du consommateur, et du value informative.

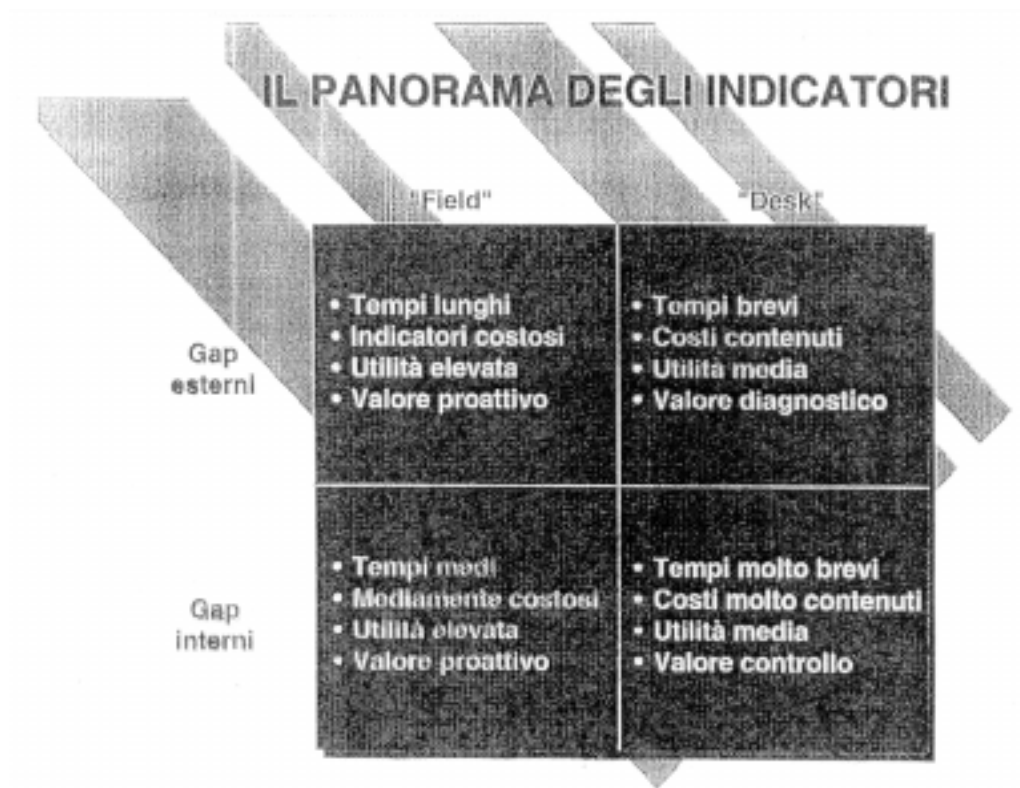
Les mesures faites à partir d'enquêtes directement menées sur le terrain contribuent sans aucun doute à une plus grande valeur informative des indicateurs et de leur concours à l'amélioration de la satisfaction du consommateur. Dans le même temps pourtant, de telles mesures requièrent des temps plus longs en moyenne et des investissements supérieurs par rapport aux mesures « desk », fondées comme nous le savons sur les indicateurs déjà présents à l'intérieur du système d'information de l'entreprise, développés sur la base de données existantes. Les mesures field dans les domaines de satisfaction/insatisfaction peuvent être conduites en ayant recours à des méthodologies de recherche qualitative (interviews personnelles en profondeur et focus group) ou quantitatives (analyses univariées ou multivariées) ; il est en même temps possible d'enquêter sur les niveaux de satisfaction globale des clients de l'entreprise et de leur intention d'achat. Le patrimoine méthodologique développé dans le cadre des recherches de marketing est évidemment utilisable avec une grande efficacité et ce quelque soit le problème de mesure de la satisfaction du consommateur. Les analyses de positionnement (discriminant analysis et multidimensional scaling), les mesures des comportements (indice de Fishbein), le relevé des préférences et de l'utilité perçue dans l'offre de l'entreprise (factor et conjoint analysis), représentent quelques exemples seulement des méthodologies appliquées à la mesure de la satisfaction du consommateur.

Comme déjà évoqué, le principal avantage des informations provenant des recherches sur le terrain est donné par leur valeur « pro active » c'est-à-dire que le relevé direct du phénomène permet (presque toujours) de définir immédiatement les actions d'amélioration que l'entreprise

peut entreprendre pour augmenter le niveau de la satisfaction du consommateur. En revanche, les indicateurs construits sur des relevés internes (desk) ont une portée informative mineure, surtout en termes d'approfondissement de la connaissance générale du phénomène. Il s'agit en effet de mesures indirectes et moins analytiques que les écarts entre satisfaction attendue et perçue. Nous parvenons à de tels indicateurs en utilisant des données recueillies avec une autre finalité.

Les propriétés des indicateurs desk sont un coût inférieur, la disponibilité immédiate et l'exclusivité des informations, la non intrusivité de la récolte des données et de leur élaboration. A la figure 2.2, nous montrons un schéma de synthèse des différentes approches de mesure, mettant en évidence les caractères distinctifs de chacun d'eux.

FIGURE 2 - LE PANORAMA DES INDICATEURS



Si nous nous concentrons sur la clientèle de l'entreprise comme population à analyser, les indicateurs principaux sur lesquels travailler sont généralement obtenus en ayant recours aux bases de données de l'entreprise. Parmi les indicateurs les plus importants nous avons :

- l'âge moyen de la clientèle, indice historique du portefeuille client ;
- le taux annuel de rétention de la clientèle, ou « CRR » (Customer Retention Rate), entendu comme le pourcentage de clients qui restent tels entre deux périodes de gestion ; c'est un indicateur utile sur la capacité de l'entreprise à fidéliser ses clients ;
- l'âge moyen évalué du portefeuille client : c'est la période au cours de laquelle on prévoit de faire durer la relation entre l'entreprise et ses clients actuels ;

- le Life Time Value (LTV) : la valeur économique de chaque client - actuellement en portefeuille - au cours de sa vie commerciale.

L'âge moyen du portefeuille clients est en général calculé selon deux façons. La première, la plus simple, consiste à faire la somme du nombre d'années depuis lesquelles chaque client est fidèle à l'entreprise, divisé par le nombre total de clients. C'est une moyenne et par conséquent cet indice est biaisé par son possible manque de fiabilité. C'est-à-dire qu'il assume une signification et des valeurs diverses en fonction de la variance dans la distribution de l'âge de chaque client. Pour cette raison, une méthode plus fiable est donnée par le calcul de la moyenne pondérée de l'âge (AMPo) :

$$AMPo = \frac{Ci \times Ai}{N}$$

où Ci indique le nombre de clients fidèles depuis Ai années, et N représente le nombre total de clients en portefeuille.

Le Customer Retention Rate représente le complément de 1 du taux de mortalité des clients. Dans l'état actuel de la science, nous le déterminons en calculant le nombre de clients en portefeuille en début d'année et nous faisons la différence entre ce chiffre et le nombre de clients en fin d'année diminué du nombre de nouveaux clients acquis au cours de la période, et rapportant le tout au nombre de clients en portefeuille en début d'année :

$$CRR = \left[CIA - (CFA - NCA) \right] / CIA \times 100$$

où CIA est le nombre de clients en début d'année, CFA le nombre de clients en fin d'année et NCA le nombre de nouveaux clients acquis au cours de l'année ce qui, soustrait à CFA, permet d'éviter de confondre deux effets imputables à des causes diverses. En fait, dans la gestion des bases de données des entreprises pour construire le CRR, on a recours à des opérations booléennes qui impliquent les « champs-âge » des différents « records - clients » des fichiers clients (customer file). L'opération décrite ci-dessus peut être effectuée d'une année sur l'autre afin de déterminer le CRR instantané. De tels indices ont toutefois des limites importantes, qui dérivent du laps de temps très court pendant lequel le phénomène est observé. On comprend facilement, en effet, que n'importe quel phénomène contingent, une manœuvre tactique en marketing par exemple, peut fausser l'ampleur d'un tel indicateur si il se réfère à une seule année de gestion. Pour dépasser ces limites du CRR instantané, il est possible d'avoir recours au CRR historique. Celui-ci représente un indicateur construit comme moyenne des CRR instantanés au cours d'un intervalle de temps raisonnable (3 à 5 ans), et permet de dépasser le caractère contingent du CRR instantané, tout en offrant à la direction de l'entreprise une mesure plus fiable de comment est en train de varier la capacité de l'entreprise à maintenir fidèles ses propres clients (ceci est certainement vrai dans les comportements, à vérifier pour les attitudes).

Sur la base des résultats obtenus avec le calcul du CRR, il est possible de parvenir à l'estimation d'un indicateur important : l'âge moyen prévisionnel (AMPr). Celui-ci peut en effet être calculé en mesurant le taux de rotation des clients en portefeuille et en le projetant dans le temps d'après la formule :

$$AMPr = 1 / (1 - CRR)$$

où AMPr indique l'âge moyen prévisionnel et CRR le Customer Retention Rate, c'est-à-dire le taux annuel de rétention de la clientèle comme précédemment défini. Évidemment, la prévision concernant la durée du portefeuille clients peut varier soit en valeur absolue, soit en

fiabilité en fonction de la configuration du CRR (instantané ou historique) qui est utilisé dans le calcul de cet indicateur.

Enfin, le Life Time Value est le produit des données sur la consommation et sur la rentabilité des clients en portefeuille avec leur âge moyen prévisionnel. Même en ce cas là, nous pouvons procéder avec des valeurs moyennes ou bien avec des valeurs pondérées et en segmentant les clients en portefeuille. Dans le deuxième cas, il est possible de suivre deux parcours. Le premier consiste dans la pondération de la consommation des clients regroupés par classe d'ancienneté dans le portefeuille clients. Le deuxième en revanche propose de segmenter le portefeuille client par classe de consommation et à l'intérieur de chaque classe procéder à la mesure du CRR, de l'âge moyen prévisionnel et donc du Life Time Value. A ce propos, il n'est pas possible de défendre la validité d'un critère plutôt que d'un autre, mais nous pouvons définir a priori les critères de segmentation du portefeuille clients plus cohérents avec les objectifs de l'analyse, c'est-à-dire avec l'homogénéité de leur système d'évaluation (satisfaction du consommateur) de la valeur de la relation (classe de consommation annuelle) ou de sa durée (ancienneté). Si nous voulons obtenir une valeur absolue de l'indicateur plus correcte, il est nécessaire de calculer le Net Present Life Time Value (NPLTV), c'est-à-dire la valeur financière du client obtenue en actualisant ses flux d'achats prévisionnels pour la période pendant laquelle il est réaliste de le considérer comme faisant partie du portefeuille.

Tous ces indicateurs fondent leur valeur informative sur l'indiscutable lien qui existe entre satisfaction, confiance et fidélité que les clients démontrent envers l'entreprise. Si nous regardons bien cette fonction informative, il faut souligner que de tels indicateurs auraient une faible signification si leur valeur n'était pas soumise à une analyse dynamique et à une comparaison précise avec les concurrents et les entreprises des autres secteurs. En revanche, leur construction est faite le plus souvent sans aucune prise en compte des principes fondamentaux de l'analyse transversale des phénomènes, en ignorant en fait que la valeur d'une mesure peut varier selon les contextes temporels et que à deux moments différents des agents causaux différents peuvent influencer un même phénomène avec des intensités diverses. Toujours dans une optique de comparaison, il faut noter que de tels indicateurs pourraient présenter des valeurs différentes lorsque la comparaison est faite entre deux populations de clients avec une structure d'âge différente (âge moyen). Chacun de ces indicateurs mérite donc quelques approfondissements en ce qui concerne ses modalités de calcul, avec l'objectif de démontrer que leur niveau de fiabilité peut être significativement amélioré si nous avons recours à quelques instruments d'analyse démographique.

4. Satisfaction du consommateur et démographie : quelques hypothèses d'utilisation

En l'état actuel, les études sur la mesure de la satisfaction du consommateur semblent présenter d'importantes opportunités de possibles améliorations, surtout en ce qui concerne la fiabilité et la significativité des indicateurs développés sur la base des systèmes d'information de l'entreprise. Et c'est justement dans ce contexte et sur un tel ensemble d'indicateurs que nous estimons possible d'appliquer à propos quelques instruments typiques de l'analyse démographique.

Les principaux indicateurs desk sont basés sur des bases de données concernant le portefeuille client de l'entreprise (customer database ou encore customer information file). En se basant sur l'hypothèse de fond de ce rapport, l'analogie entre portefeuille clients et population, il est possible de réexaminer les modalités de construction de tels indicateurs.

Une première série de problèmes concerne la détermination et l'utilisation du concept de la Customer Retention Rate (CRR). Il y a en effet une importante différence par rapport au calcul du taux de mortalité dérivant de la démographie exprimé en général comme :

$$TMC = CPPAi / CP \text{ en milieu d'année}$$

où TMC représente le taux de mortalité client (le CRR est par conséquent calculé comme le complément de 1 de ce rapport), CPPAi indique le nombre de clients en portefeuille perdus au cours de l'année i et CP le nombre total de clients en portefeuille. Comme le calcul de cet indice, le calcul du CRR présente une limite liée à sa construction : le balayage transversal des périodes d'observation, correspondant à des intervalles discrets et écartés (en l'occurrence d'un an), ne permet pas de relever l'importance des phénomènes de natalité et de mortalité (clients acquis et perdus au cours d'une période inférieure à celle du relevé), au cours de cette période. En termes spécifiquement opérationnels, ceci signifie surestimer systématiquement le taux de rétention et en particulier se poser dans la condition de ne pas pouvoir identifier des abandons qui de par leur précocité pourraient être associés à des niveaux élevés d'insatisfaction.

Une seconde limite au CRR calculé de cette façon est liée à l'impossibilité de savoir quels clients sont restés fidèles ou bien quelle « génération » (dans le sens chronologique) détermine la donnée du CRR du portefeuille clients perdu dans l'ensemble. De cette perte d'information dérive une série de lacunes concernant l'évaluation des effets que les programmes de marketing et ceux de la satisfaction du consommateur ont produit sur la fidélité des clients dans les moyen et long termes. En se référant surtout au CRR instantané, on sent intuitivement que si nous nous limitons à une valeur contingente du CRR nous risquons de ne pas relever l'effet réel de l'activité de l'entreprise sur la fidélité de ses clients. Il pourrait par exemple nous échapper l'importante information du retour sur les effets différentiels que le changement de la politique de l'entreprise peut obtenir sur la fidélité des clients « vieux » par rapports aux clients « jeunes » et ainsi de suite. Le CRR instantané est donc peu fiable si nous voulons l'utiliser à des fins de mesure de l'espérance de vie d'un portefeuille clients.

Si nous considérons que le CRR représente une mesure propédeutique au calcul d'une série d'autres indicateurs extrêmement importants du point de vue de la portée informative, ceux de l'âge moyen prévisionnel et le Life Time Value dans leurs diverses configurations, cela vaut la peine de concentrer sur la détermination de tels indicateurs les efforts d'enrichissement analytico-méthodologiques qui constituent l'objectif central de ce rapport. Tout particulièrement, nous présenterons quelques instruments méthodologiques qui permettent d'améliorer la significativité des comparaisons intertemporelles, interentreprises et intersectorielles.

L'analyse par populations du portefeuille clients

L'analyse démographique offre des instruments qui permettent d'améliorer de façon significative la fiabilité de l'analyse.

L'approche suivie jusqu'ici ne prend pas, en effet, pas en compte l'unité fondamentale de l'analyse démographique : la population. En simplifiant à l'extrême ce concept (destiné à la communication des gens de marketing et non des démographes), l'analyse par populations consiste en l'examen des phénomènes ayant trait aux sous-groupes d'une population, déterminés selon leur homogénéité chronologique, par exemple l'année de naissance.

Les processus exposés ci-dessus proposent presque toujours des mesures qui lient, en les confondant, des phénomènes qui ont pris naissance (et effet) à des moments différents. Par exemple le CRR « médiateur » de l'ancienneté moyenne du portefeuille clients ne tient pas compte des différences qui pourraient subsister en termes de satisfaction, de confiance et de fidélité entre les différentes générations de clients. Tout cela, sans considérer que l'espérance de vie (âge moyen prévisionnel) pourrait être plus fiable si elle était déterminée par la population. De la même façon, la détermination du LTV ne prend pas en compte que des clients avec des âges de portefeuille différents pourraient présenter une variance élevée pour le niveau de consommation, comme il est souvent démontré par les indices de cross-selling qui augmentent au fur et à mesure de l'augmentation de la durée de relation entre entreprise et client. En suivant une telle logique, il nous est donc possibles de déterminer le LTV comme suit :

$$LTV = Li \times Ci \times AMPi$$

où Li représente le nombre de clients en portefeuille à partir de la i -ème période (exprimé en années), Ci la consommation moyenne annuelle des clients acquis au cours de la i -ème période et $AMPi$ l'âge moyen de portefeuille prévisionnel calculé par la population de clients, née au cours de la i -ème période. L'avantage offert par l'analyse par population apparaît alors évident lorsqu'il faut calculer les effets du cross-selling plutôt que la relation entre ancienneté et pénétration sur le client. En somme, la mesure des indicateurs de la satisfaction du consommateur différenciée par population permettrait une plus grande fiabilité des indicateurs et des évaluations économique-entrepreneuriales qui se basent le plus souvent sur des données de satisfaction de la clientèle et sur la valeur du portefeuille clients (essentiellement par la mesure du goodwill).

Pour plus de détails, ce problème peut être étudié en faisant spécifiquement référence à la littérature développée pour la business demography. De nombreux auteurs en effet ont souligné l'abus méthodologique substantiel rencontré de façon récurrente quand, sur des données de nature cross-sectional, nous superposons une interprétation de type évolutif (Emerson, 1977 ; Reynolds et Rentz, 1981 ; Rentz, Reynolds et Stout, 1983). En d'autres termes, ceci se vérifie quand les différentes valeurs prises par une variable déterminée en faisant référence aux différentes classes démographiques d'une population (par exemple le taux d'utilisation d'une catégorie de produit par classe d'âge) est expliquée en hypothétisant un changement des styles de consommation lié au vieillissement de cette même population. Il s'agit évidemment comme souligné par une large littérature sur le sujet, d'une interprétation qui se base sur des présupposés restrictifs et pas nécessairement corrects : d'autres modèles fondés sur des prévisions des comportements de consommation complètement différentes pourraient être parfaitement adéquats aux données.

Le problème, comme nous le savons, est dû à l'impossibilité de séparer complètement l'un de l'autre les effets d'âge, de durée et de population qui dans leur interaction contribuent à déterminer le modèle d'observation à notre disposition. Dans ce contexte spécifique pris comme objet d'enquête, un tel problème est affronté à travers une prospective légèrement différente de celle illustrée ci dessus : l'objectif n'est pas celui d'analyser la contribution de l'analyse par population à une relecture plus juste des données cross-sectional, mais bien de faciliter la comparaison et l'interprétation correcte des indices de la satisfaction du consommateur obtenus à des moments différents. Les deux problèmes sont méthodologiquement identiques : si l'analyse cross-sectional d'une population composite présente des ambiguïtés interprétatives liées au fait de confondre les deux effets d'âge et de population, il est aussi vrai que la comparaison transversale de cette même population en deux moments différents tend à confondre en son sein les effets d'âge et de durée.

Il est alors utile de résumer les points de comparaison entre les catégories analytiques d'âge, de population et de durée dans le contexte spécifique de mesure de la satisfaction du consommateur. Plus précisément le parallèle exposé fait l'hypothèse de certains a priori théoriques exposés ci dessous :

Satisfaction du consommateur et effet âge

Nous faisons l'hypothèse d'une augmentation naturelle de la satisfaction liée à la prolongation de la relation avec une entreprise déterminée ; le débat théorique n'est pas unanime à définir le sens du lien causal entre satisfaction et fidélité mais l'existence d'une corrélation est incontestable.

Satisfaction du consommateur et effet durée

Nous représentons l'impact des actions gestionnelles entreprises afin de modifier de façon délibérée et immédiate les dimensions du produit ou du service qui caractérisent la qualité perçue par le client. Il peut être considéré comme variable de « traitement » contrôlée par l'entreprise. Il va de soi que l'effet de durée peut être déterminant même à partir de phénomènes plus généraux, et pas directement influençables par l'entreprise.

Satisfaction du consommateur et effet population

Il est déterminé par les différences dans les évaluations de la satisfaction liée uniquement aux variations de la composition dans les comportements et les attitudes d'une population entendue comme classe actuelle de clients. On peut faire l'hypothèse d'un effet de population qui puisse refléter l'augmentation naturelle des attentes dues à une hausse progressive du niveau moyen de l'offre tiré vers le haut par une pression compétitive croissante.

Dans la comparaison transversale des données de la satisfaction du consommateur afin de mettre en valeur les effets de durée liés à des initiatives ou à des programmes d'amélioration, il devient donc important de séparer les effets de durée purs des effets d'âge et de population. Si ceci n'est pas fait, le risque se profile de définir des rapports de cause à effets faussés en attribuant à l'entreprise et à ses actions de marketing des mérites ou démérites en réalité inexistantes. L'hypothèse de travail exposée ici est que l'analyse par populations puisse devenir un instrument puissant d'identification des effets nets des actions de l'entreprise (effets durée) séparés des processus naturels d'approfondissement des intensités relationnelles du rapport client-fournisseur (effet d'âge).

Un exemple numérique construit sur des données hypothétiques peut aider à clarifier la signification de ces différences. Pour des raisons évidentes de simplicité, nous considérons seulement deux périodes et deux typologies de clientèles (population) classées sur la base du rapport avec l'entreprise. Dans le tableau ci-dessous, nous reportons les valeurs des indices de la satisfaction du consommateur mesurée dans les deux unités de temps (1993 et 1994).

	1993	1994
Client type 1	70	60
Client type 2	90	80
Indice moyen de CS	80	70

Une première analyse basée comme d'habitude sur la seule transversale semble relever une nette diminution dans le niveau général de satisfaction soit relatif à la clientèle dans son ensemble, soit relatif aux simples composantes divisées par classe d'âge. Il s'en suit un jugement fortement critique sur la capacité démontrée par l'entreprise à satisfaire les attentes de sa propre clientèle.

En réalité, en utilisant des schémas interprétatifs suggérés par l'analyse des populations, nous pouvons proposer une lecture complètement différente. En effet, les données reportées dans le tableau peuvent dériver de l'action conjuguée d'un effet d'âge égal à 10 qui se cumule à un effet de populations de même valeur et respectivement comparables aux différences transversales et le time lag. Il est intéressant de remarquer qu'une telle interprétation fait l'hypothèse d'un effet de durée égal à 0 : ce qui signifie qu'une baisse des indices de la satisfaction du consommateur peut se produire indépendamment des actions de gestion directement ou indirectement menées pour en influencer le niveau.

Il est nécessaire de répéter que l'analyse par population veut être proposée ici comme une nouvelle approche méthodologique de l'analyse des données qui dérivent de la base de données clients et non pas comme une technique statistique formelle : son utilité se base sur des prérequis théoriques qui, à leur tour, doivent être l'objet d'attentions précises. Ceci dit, il semble évident que l'apport à l'interprétation des phénomènes complexes sous-jacents à la satisfaction du consommateur ne puisse être négligé en aucune façon.

Toujours à propos de la mesure du CRR, quelques recherches récentes sur la satisfaction du consommateur (Bridelli et Berni, 1991; Ainslie et Pitt, 1992) ont mis l'accent sur la possibilité d'utiliser des valeurs de rétention comme des indicateurs de performances de l'entreprise en des termes absolus et relatifs (benchmarking). En fait, surtout si nous voulons l'utiliser comme moyen de comparaison, le CRR peut être utilisé seulement dans le cas où la structure d'âge du portefeuille clients a été standardisée (ajustée pour tenir compte des structures d'âge différentes et particulières). La comparaison entre les taux de rétention relatifs et deux populations différentes de clients peut fournir des résultats profondément différents. Il serait injustifié pourtant de corrélérer une telle différence à l'existence de facteurs causaux qui agissent sur les deux populations de référence. La différence en effet pourrait être imputable aux taux de mortalité par classe d'âge différents mais physiologiques. La composition par classe d'âge du portefeuille clients pourrait avoir un fort impact sur le niveau de fidélité, considérant que très souvent au fur et à mesure de l'augmentation de la durée de la relation, les coûts de transferts augmentent aussi, de même que le nombre de cas de comportement d'achat inertiel de la part du client et de façon positive la capacité de l'entreprise à satisfaire des exigences même très spécifiques de chaque client.

Le manque de relevés des différences dans les structures d'âge du portefeuille client pourrait donc conduire à une estimation non réaliste des effets compétitifs d'un programme d'amélioration de la satisfaction du consommateur ou bien d'une quelconque action de marketing dirigée essentiellement vers les clients de l'entreprise. A ce propos il est intéressant d'interpréter les programmes d'amélioration de la satisfaction du consommateur selon la mesure des actions pour l'augmentation du niveau de vie d'une population et d'en mesurer donc l'effet en termes d'augmentation de la fidélité et de permanence dans le portefeuille. Nous pouvons mesurer aussi les augmentations dans la probabilité de survie d'une population.

Encore une fois, l'analyse différenciée par population est essentielle. Il est évident en effet que les effets des programmes de satisfaction du consommateur doivent être relevés de façon à mesurer les différences entre les clients « nés » dans les conditions d'offre déjà améliorées par rapport à ceux qui ont joui d'une telle amélioration seulement à partir d'un certain moment de leur relation avec l'entreprise.

La standardisation du portefeuille clients

Comme précédemment évoqué, la mesure des indicateurs CRR, AMP_r et LTV a une portée informative élevée à partir du moment où leur construction permet de parvenir à des résultats engageants soit à la fin du contrôle transversal des performances de l'entreprise, soit dans le cadre des comparaisons interentreprises et intersectorielles.

En procédant en ce sens, le premier problème est lié à la comparaison de deux populations différentes de clientèles de l'entreprise.

En effet, il peut se passer que deux portefeuilles clients caractérisés par des structures d'âge différentes, il s'agit bien de l'âge de la relation donnent lieu à des indicateurs de fidélité différents justement parce qu'ils ont des structures différentes, et non pas en conséquence d'un réel différentiel de satisfaction offert par l'entreprise.

Par exemple, des entreprises qui ont commencé à opérer sur le marché à des moments historiques différents pourraient présenter des âges moyens de la clientèle (historiques ou

prévisionnels) très différents ; et aussi un turnover différent puisque les coûts (réels et psychologiques) de transfert (émigration) sont positivement corrélés à la durée de la relation. Dans de telles conditions, il est fréquent pour les biens qui requièrent un effort d'évaluation important avant l'achat (les biens de fidélité - Nelson 1974) deux niveaux différents de customer retention ne peuvent être utilisés comme indicateurs de performance différente de l'entreprise et encore moins comme paramètres de patrimoine de fidélité accumulé par l'entreprise.

Le problème est donc de rendre comparables les populations de clients sur lesquelles on a l'intention de travailler pour l'analyse des phénomènes liés à la satisfaction du consommateur et qui sont par conséquent corrélés à l'âge de la relation entre entreprise et clients. L'analyse démographique en situation analogue propose le recours à la standardisation.

La standardisation consiste à la redéfinition de la structure d'âge de plusieurs populations qui doivent être comparées. L'objectif est de neutraliser les effets de différentes compositions par âge dans le calcul des indicateurs synthétiques auxquels nous avons fait référence précédemment, ou bien l'évaluation comparée des autres indicateurs comme le Cross Buying ou la Life Time Value.

En termes opérationnels, la méthode pour rendre comparables deux populations de clientèles consiste à pondérer les différentes phases d'âge avec des poids standards qui montrent l'incidence des clients avec un âge déterminé et à remplacer les poids en pourcentages spécifiques. De cette façon, nous attribuons à deux populations de clients spécifiques une même structure par âge. Par exemple, si nous voulons construire les indicateurs d'âge moyen prévisionnel et de life time value et si nous voulons les comparer pour deux portefeuilles clients différents, la standardisation prévoit des indicateurs en amont, âge moyen pondéré et CRR. Un tel contrôle intervient lorsque, par exemple, nous construisons l'indice d'âge moyen pondéré sur les coefficients de pondération c'est-à-dire sur la définition du nombre de clients qui est multiplié par chaque classe d'âge de relation.

Les poids spécifiques des portefeuilles clients des différentes entreprises considérées sont ainsi rendus homogènes.

Si nous posons pour les deux entreprises et l'âge moyen pondéré de leur portefeuille clients, nous pouvons calculer :

$$AMPo = \frac{Ci \times Ai}{N}$$

et

$$AMPo = \frac{Ci \times Ai}{N}$$

où Ci indique le nombre de clients fidèles depuis un nombre d'années égal à Ai et N le nombre total de clients en portefeuille. La standardisation opère en substituant au poids Ci spécifique des entreprises un poids standard Ci .

Les deux indicateurs standards sont :

$$AMPo = \frac{Ci \times Ai}{N}$$

$$AMPo = \frac{Ci \times Ai}{N}$$

La standardisation peut être mise en place par la substitution des poids spécifiques c'est-à-dire le nombre relatif de clients pour chaque classe d'âge avec des poids standards identiques

dans les portefeuilles clients concernés. La définition des poids standards peut être faite soit en adoptant la distribution par âge des clients des entreprises considérées, soit en définissant une distribution idéale qui d'après le bon sens des démographes (Livi Bacci, 1981) devrait présenter des valeurs intermédiaires à celles réelles des différentes entreprises à l'étude.

Le principal résultat d'un processus de standardisation consiste en la réduction des différences dues à la structure d'âge du portefeuille clients (C_i et C_j) et à l'importance relative conséquente qui est donnée aux différentiels que les entreprises, objet de comparaison, présentent dans les indicateurs de cross-buying, dans la capacité de rétention et dans le life time value de leur portefeuille clients.

Tous les indicateurs sont ainsi comparables dans la mesure où leur éventuelle différence n'est pas imputable à une structure d'ancienneté différente du portefeuille clients mais plutôt à leur capacité effective à satisfaire le client, à engendrer la confiance et à fidéliser la clientèle. En fait, à partir des indicateurs standardisés, il est possible de procéder à des relations de comparaison qui permettent de mesurer les différentiels nets. Par exemple, en confrontant le life time value de deux entreprises dont les indicateurs de AMPo et CRR ont été soumis à un processus de standardisation :

$$\frac{LTV_i}{LTV_j}$$

nous pouvons obtenir trois résultats. Si le rapport est supérieur à 1, il est possible d'affirmer que la capacité de l'entreprise à satisfaire et à construire des relations de confiance avec ses propres clients est supérieure à celle de l'entreprise et vice versa si l'indicateur est inférieur à 1. Dans le cas où il est égal à 1, les entreprises sont équivalentes.

En résumé, la standardisation permet de déterminer quelle part de la différence entre deux phénomènes est imputable aux différences de structure dans l'âge du portefeuille clients et différences dans les attitudes et les comportements qui caractérisent la relation entre l'entreprise et ses propres clients.

Conclusion

Les opportunités offertes par l'application des instrument d'analyse démographique à la construction des indicateurs de la satisfaction du consommateur apparaissent incontournables.

Sur la base des hypothèses formulées dans le présent rapport, il apparaît évident qu'ils peuvent offrir une importante contribution à la mesure de la satisfaction du consommateur sous un double profil. D'un côté, en effet, ils consentent à développer des indicateurs dont la valeur absolue est très fiable. De l'autre, ils permettent de rendre fiables les comparaisons.

Il est alors souhaitable de lier les finesses méthodologiques destinées à l'augmentation de la fiabilité et de la significativité des indices et leur utilisation, non seulement dans le cadre de la gestion des entreprises, mais aussi dans le système complexe des relevés de valeurs, systèmes toujours à la recherche de mesures fiables du potentiel qui engendre les ressources immatérielles du marketing.

Contribution à la prévision de la population étudiante universitaire de Belgique francophone

Luc DAL

GÉDAP, Centre d'études de Gestion Démographique pour les Administrations Publiques
Université catholique de Louvain

1. Le contexte général

Caractérisée par un accès assez libre, en comparaison de la situation qui prévaut dans d'autres pays voisins, l'université belge connaît depuis la fin de la Seconde Guerre Mondiale une croissance spectaculaire du nombre de ses étudiants. Plusieurs facteurs, en interaction mutuelle, peuvent expliquer cette croissance : parmi ceux-ci, on citera la demande accrue de formation supérieure émanant de la population, les exigences de plus en plus élevées du monde du travail, l'évolution du marché de l'emploi, la crise économique qui pousse les jeunes à retarder leur entrée dans la vie professionnelle, la démocratisation de l'enseignement supérieur, ...

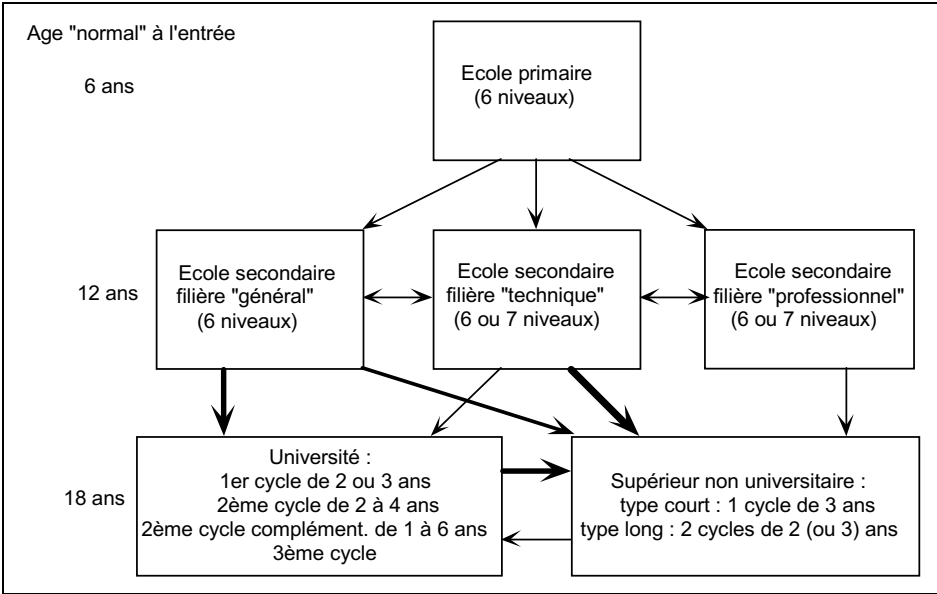
Cette croissance rapide peut aussi trouver une explication partielle dans la réglementation des conditions d'accès inchangées depuis la fin des années cinquante. Celle-ci ont été mises en place dans un souci de démocratisation et d'ouverture de l'université et elles sont en fait très peu restrictives, puisque tout élève qui termine avec fruit le cycle de l'enseignement secondaire, dans la filière « générale » ou « technique », peut légalement s'inscrire dans une université et une faculté de son choix (exception faite de la faculté des sciences appliquées où un examen d'entrée est organisé par les universités) ; de plus, un système de « passerelles » permet, sous certaines conditions, aux étudiants titulaires de certains titres de l'enseignement supérieur non universitaire d'accéder directement en première année du second cycle de base universitaire. Enfin, pour être complet, il faut signaler l'existence des formations organisées pour le public adulte, bien qu'en termes d'effectifs relatifs, ces dernières attirent un public peu important.

Quoiqu'il en soit, actuellement, l'essentiel de la population universitaire (plus de 95 %) est constitué d'étudiants ayant suivi un parcours « classique » et qui entrent donc en première année universitaire au terme de l'enseignement secondaire.

La figure ci-dessous synthétise les principaux canaux d'accès en première année dans l'enseignement universitaire belge, tout au moins pour les étudiants issus de l'enseignement secondaire belge.

Pour l'ensemble du pays et depuis 1946, le nombre des étudiants universitaires est ainsi passé de 18 120 à 124 046 en 1991-92, soit une multiplication des effectifs par presque un facteur 7 en 45 ans. Ce mouvement s'est particulièrement accéléré depuis 1970, période pendant laquelle les effectifs ont été multipliés par un facteur 1,5. Ces chiffres concernent l'ensemble de la Belgique (i.e. les Communautés flamande et française).

FIGURE 1. : PRINCIPAUX CANAUX D'ACCÈS À L'UNIVERSITÉ



Dans le cas particulier de la Communauté française, cette croissance spectaculaire est synthétisée par le tableau suivant qui présente l'évolution des effectifs totaux d'étudiants universitaires depuis l'année académique 1972-73 jusqu'en 1991-92.

1972-73	1991-92	Accrois. absolu	Accrois. relatif	Accrois. annuel moyen
41 383	59 766	18 383	44 %	1,85 %

Face à une telle augmentation du nombre d'étudiants, il est légitime de s'interroger sur l'évolution future de la population universitaire de la Communauté française de Belgique¹.

A cet effet, nous proposons de calculer des projections de nouveaux étudiants en première année du premier cycle de base dans les universités francophones belges et ce à l'horizon de l'année académique 2005-2006. Ces projections seront basées sur des scénarios dont les hypothèses résulteront, elles-mêmes, naturellement de l'analyse du passé.

Au plan strictement méthodologique, établir des perspectives d'entrées dans le système universitaire n'a évidemment de sens que dans la mesure où celles-ci s'appuient sur la population soumise au risque, c'est-à-dire sur la population susceptible d'entrer dans ce système et, comme signalé, sur l'étude du passé du phénomène. Aussi, nous nous référerons aux projections de population de l'Institut National de Statistique² et ³.

¹ Rappelons que les universités sont subsidées par la Communauté française et que leur financement dépend directement de la taille de leur population.

² Ministre des Affaires Économiques : Institut National de Statistique, Perspectives de Population 1992- 2050.

³ Les autres sources de données qui ont été utilisées sont :
- les publications du Ministère de l'Éducation Nationale (jusqu'en 1986-87),
- les publications du Ministère de la Communauté française depuis 1987),
- les annuaires de la Fondation universitaire,
- les annuaires et la base de données individuelles du Conseil des Recteurs francophones.

Le cadre dans lequel nous travaillons par la suite est défini comme suit :

- nous nous intéressons exclusivement à la Communauté française du pays. Celle-ci est composée de la population de la Région wallonne et de 90 % de la population de la Région bruxelloise (qui représente plus ou moins 40 % de la population belge),
- nous envisageons uniquement les effectifs d'étudiants et non pas les inscriptions que ceux-ci prennent puisqu'un étudiant peut prendre éventuellement plusieurs inscriptions au rôle.
- nous ne distinguons pas les différentes institutions d'enseignement universitaire, mais nous travaillons à un niveau agrégé⁴.

En outre, il est nécessaire de préciser quelques termes de la terminologie utilisée en particulier en démographie scolaire. On désignera ainsi par :

- *cohorte ou génération civile* : l'ensemble des individus nés au cours d'une même année civile.
- *étudiant de première génération* : tout étudiant qui s'inscrit pour la 1^{ère} fois en première année dans l'enseignement universitaire de la Communauté française.
- *taux de participation ou taux de fréquentation* : le pourcentage d'individus d'une classe d'âge donnée inscrits dans le système considéré au cours d'une année académique donnée.
- *taux d'accès ou taux d'entrée* : le pourcentage d'individus d'une classe d'âge donnée inscrits pour la première fois dans le système considéré au cours d'une année académique donnée.

Ainsi, par exemple, si la cohorte 1974 compte 1000 individus et que, parmi ceux-ci, 80 sont inscrits à l'université en 1992-93, le taux de participation, ou de fréquentation, de la classe d'âge de 18 ans de cette cohorte est donc de 8 %. Si parmi ces 80 étudiants, 50 s'inscrivent pour la toute première fois à l'université, on parlera alors de 50 étudiants de première génération, en ce sens qu'il s'agit de 50 nouveaux étudiants dans le système universitaire, et le taux d'accès des 18 ans, de la génération 1974, est de 5 %.

2. Analyse des tendances de 1972-73 à 1992-93

La croissance des effectifs dans l'enseignement supérieur est spectaculaire et ce, quelle que soit la forme d'enseignement envisagée (les deux formes organisées étant l'universitaire et le supérieur non universitaire). Si l'enseignement universitaire prédomine en chiffres absolus, la vitesse de croissance est cependant relativement plus importante dans le supérieur non universitaire. La figure 2 illustre cette évolution.

Cette forte croissance doit être mise en relation avec la diminution du nombre des naissances observée en Belgique depuis 1964, phénomène qui devrait, en toute logique, avoir un impact négatif, ou tout au moins avoir un effet « modérateur » sur le volume de la population susceptible d'entrer à l'université.

De fait, les observations et les perspectives de la population âgée de 17 à 20 ans indiquent une tendance à la décroissance depuis 1979 ; une période de stabilisation (qui débute en 1994) succède à cette période de dépression et elle est suivie d'une timide croissance à partir de l'an 2000.

La croissance du nombre d'étudiants, contradictoire avec le mouvement précédent, peut s'expliquer a priori par l'effet de deux facteurs :

1. un accès de plus en plus large de la population ;
2. une rétention plus importante à l'intérieur du système pouvant elle-même s'expliquer par :
 - une augmentation des taux de redoublement qui aurait pour effet de retenir plus longtemps les étudiants dans le système ;

⁴ Il est certain qu'il existe des différences entre les différentes institutions, ne fût-ce que par le public qu'elles accueillent.

- une prolongation du séjour universitaire des étudiants (formations complémentaires telles que licences spéciales, second diplôme, ...).

FIGURE 2 : ÉVOLUTION DU NOMBRE D'ÉTUDIANTS (UNIVERSITÉ ET SUPÉRIEUR NON UNIVERSITAIRE) EN COMMUNAUTÉ FRANÇAISE DE BELGIQUE, PAR ANNÉE CIVILE DE L'INSCRIPTION.

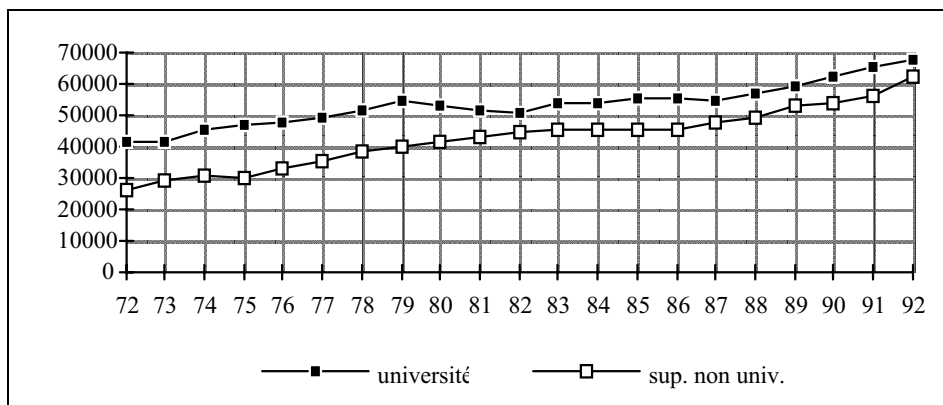
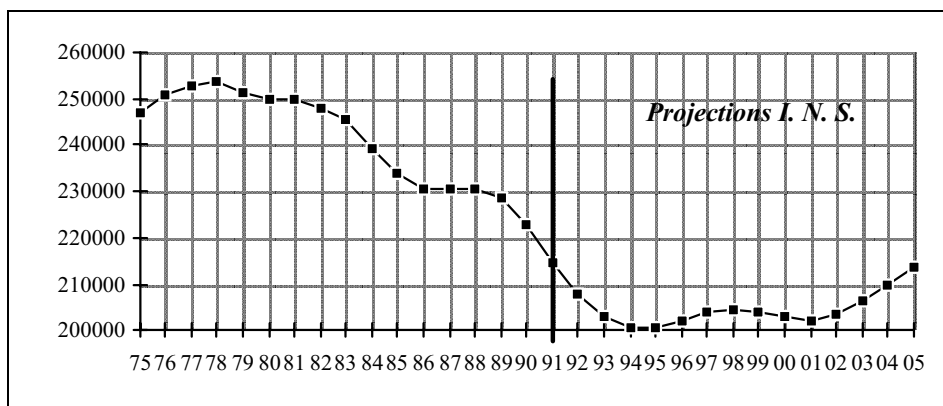


FIGURE 3 : ÉVOLUTION DE LA POPULATION ÂGÉE DE 17 À 20 ANS EN COMMUNAUTÉ FRANÇAISE DE BELGIQUE, PAR ANNÉE CIVILE : OBSERVATIONS ET PROJECTIONS.



Dans la réalité, l'analyse des transitions internes de l'université⁵ montre que les taux de redoublement sont relativement stables au cours du temps et que le prolongement du séjour universitaire ne permet pas d'expliquer lui seul l'évolution spectaculaire qui vient d'être observée.

⁵ Voyez à ce propos : DAL L., DUPIERREUX J.M., GUYOT J.L., KITA-PHAMBU P., TILLE Y. (sous la direction scientifique de Mme Wattelar Ch. et de Mrs Beguin A., Delcourt J., Dræsbeke J.J., Devooght J. et Hecquet I.), *Étude prospective de la population étudiante des universités de la Communauté française de Belgique* rapport de recherche final, Volume I et II, photocopié, Institut de Démographie de L'U.C.L., Laboratoire de Méthodologie du Traitement des Données de l'U.L.B., Louvain-la-Neuve / Bruxelles, janvier 1994, 568 pages et annexes (700 pages environ) et DAL L., DUPIERREUX J.M., GUYOT J.L., KITA-PHAMBU P., TILLE Y. (sous la direction scientifique de Mme Wattelar Ch. et de Mrs Beguin A., Delcourt J., Dræsbeke J.J., Devooght J. et Hecquet I.), *Étude prospective de la population étudiante des Dræsbeke de la Communauté française de Belgique*, rapport de recherche final, document de synthèse, photocopié, Institut de Démographie de L'U.C.L., Laboratoire de Méthodologie du Traitement des Données de l'U.L.B., Louvain-la-Neuve / Bruxelles, janvier 1994, 97 pages.

On se focalisera par conséquent sur le premier facteur pour étudier l'accès croissant des jeunes à l'université, et on analysera en fait quelle est la population qui s'inscrit pour la première fois à l'université, en d'autres termes, qui sont les étudiants de première génération.

A priori, on pourrait croire que les jeunes âgés de 17 à 21 ou 22 ans constituent ce substrat démographique dans lequel l'université va recruter ses étudiants de première génération. Dans ce cas, il suffirait de mettre en relation les inscriptions observées avec le substrat démographique. Le coefficient reliant les deux sous-populations est alors le « taux d'accès des étudiants de première génération ». Il suffirait alors d'analyser la tendance chronologique de ce taux, de l'extrapoler, de relier ces extrapolations au substrat démographique et le problème serait ainsi facilement résolu.

En réalité, le problème est beaucoup plus complexe qu'il n'y paraît et les perspectives d'étudiants de première génération, si elles s'appuient sur le substrat démographique, dépendent également d'une grande diversité de facteurs et ne peuvent en aucun cas se réduire à un modèle « mécanique » simpliste.

Il est utile de préciser que le public universitaire, qu'il soit de première génération ou non, est un public hétérogène, en ce sens qu'il est composé de résidents belges (il s'agit de Belges et d'étrangers qui ont effectué leur scolarité en Belgique) ainsi que d'étrangers qui viennent en Belgique pour y accomplir leurs études supérieures. Nous nous intéressons dans les lignes qui suivent à la première sous-population (de loin la plus importante), la seconde étant, par nature même, soumise à des effets non prévisibles et étant nettement moins importante en chiffres relatifs.

Le graphique 4 présente l'évolution des effectifs d'étudiants de première génération, diplômés de l'enseignement secondaire belge, selon le sexe.

Le volume de la population susceptible d'entrer à l'université (17 à 20 ans) et celui du nombre d'étudiants de première génération présentent clairement des évolutions contradictoires et ces deux évolutions paradoxales se traduisent par une croissance marquée du taux annuel d'accès à l'université. Ainsi, le taux annuel d'accès, défini par le nombre d'étudiants de première génération âgés de 17 à 20 ans rapporté à la population correspondante, suit une évolution chronologique comme indiqué à la figure 5.

FIGURE 4. : ÉVOLUTION DES EFFECTIFS D'ÉTUDIANTS DE PREMIÈRE GÉNÉRATION.

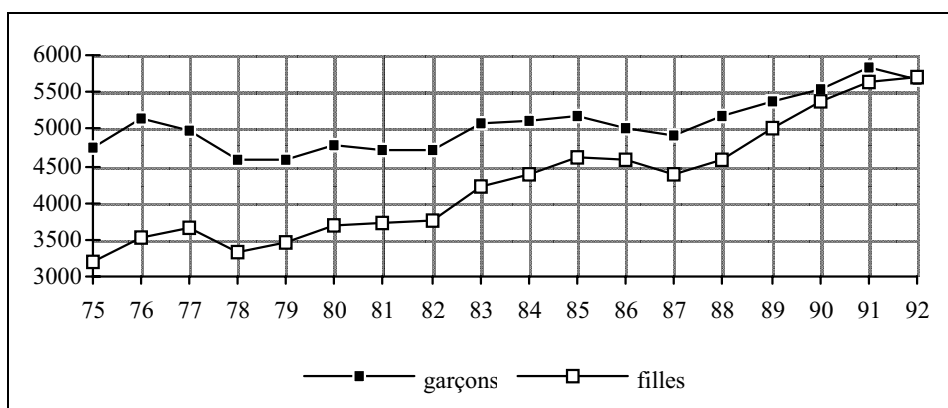
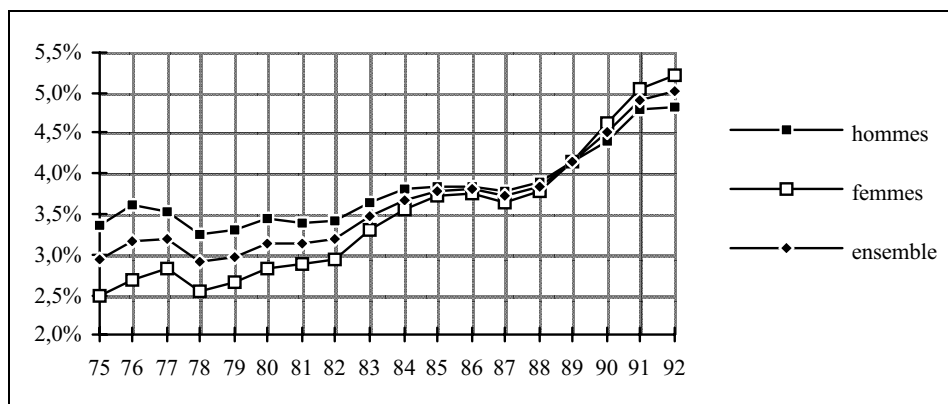


FIGURE 5. TAUX ANNUEL D'ACCÈS À L'UNIVERSITÉ



Globalement, la tendance de ces courbes est nettement à la croissance. Celle-ci est plus marquée dans la population féminine et il est intéressant de noter qu'à partir de 1989, le taux d'accès annuel des filles rejoint celui des garçons, pour ensuite le dépasser. C'est clairement cette sous-population qui a donc bénéficié le plus de l'ouverture du monde universitaire. Sans entrer dans des détails, ce phénomène est évidemment à mettre en rapport avec les divers mouvements d'émancipation féminine. La question du choix des études qu'elles entreprennent n'est pas abordée ici, mais on retiendra qu'essentiellement, la féminisation a lieu dans les facultés de sciences humaines et médicales. La plupart des pays industrialisés connaissent le même phénomène.

3. Analyse des parcours scolaires

Rappelons que l'objectif fixé est d'établir des perspectives du nombre d'étudiants de première génération à l'université. A partir de ces perspectives totales, une répartition de ces étudiants dans les différents domaines d'études et facultés est déterminée, sur base des observations du passé, et ensuite par un modèle assez complexe de type markovien, calcule le futur du système dans ces diverses composantes ainsi qu'une série d'indicateurs tels que les temps de parcours moyens, le coût par diplôme décerné, Ce volet de la perspective n'est pas abordé ici.

En ce qui concerne notre objectif, nous avons retenu une approche originale, en ce sens que nous « suivons » chacune des cohortes nées depuis 1954 au travers de son parcours dans les différentes composantes du système scolaire donnant accès à l'université⁶.

Les populations de ces cohortes sont ensuite connectées à la population susceptible de se trouver dans les différents états que nous allons étudier. Nous disposons des observations passées de la population francophone du pays ainsi que des perspectives établies par l'Institut National de Statistique. Précisons à ce propos que les futurs étudiants de l'an 2005 sont déjà nés et qu'en travaillant de la sorte, nous ne nous hasardons pas trop quant à ces perspectives de population.

Le volume total des écoliers (en distinguant les deux sexes) qui entrent en première année primaire est ainsi suivi au cours de leur scolarité jusqu'à leur éventuel accès à l'enseignement

⁶ Ce choix est imposé par le fait que nous ne disposons pas de données individuelles aux niveaux primaires et secondaires, mais celles-ci existent dans le monde universitaire.

supérieur à l'aide de tables des parcours scolaires définies ci-dessous⁷. Chaque cohorte est étudiée de la sorte. Il est évident que cette manière de travailler ne permet pas de distinguer les élèves redoublants des autres élèves, selon leur âge et leur niveau, mais cette contrainte est imposée par les publications officielles qui ne désagrègent pas les populations selon ces critères.

A partir de ces tables, qui constituent le principal matériau statistique, nous avons essayé de trouver sur base de l'information qu'elles contiennent quels étaient les déterminants de l'option « université ». On notera que cette table contient des données qui semblent erronées : leur présence ne peut s'expliquer que par des erreurs d'encodage qui délibérément n'ont pas été modifiées.

FIGURE 6. EXEMPLE DE TABLE DES PARCOURS SCOLAIRES : POPULATION MASCULINE NÉE EN 1966

Année scolaire	École primaire						École secondaire : filière général						
	1	2	3	4	5	6	1	2	3	4	5	6	7
1971	957	7		1		1							
1972	28693	1059	53	3									
1973	5750	23372	960	13									
1974	940	6686	21451	964	9								
1975	204	1783	7102	19853	923	35							
1976	73	382	2230	6818	19512	878							
1977	23	142	752	2922	7413	17255	745	8					
1978	17	32	215	895	3431	6884	15214	601	3				
1979	7	9	55	253	1055	2907	8047	11868	481	40			
1980	5	2	10	46	201	676	3328	6752	7744	387	7	2	
1981			2	5	19	56	666	2991	3544	6060	324	4	
1982							100	743	1384	3139	4771	296	
1983							9	129	347	1370	2679	4317	16
1984							2	15	88	549	1220	2100	152
1985								2	14	126	462	1061	195
1986									3	15	146	405	35
1987											26	136	64
1988													
1989													
1990													

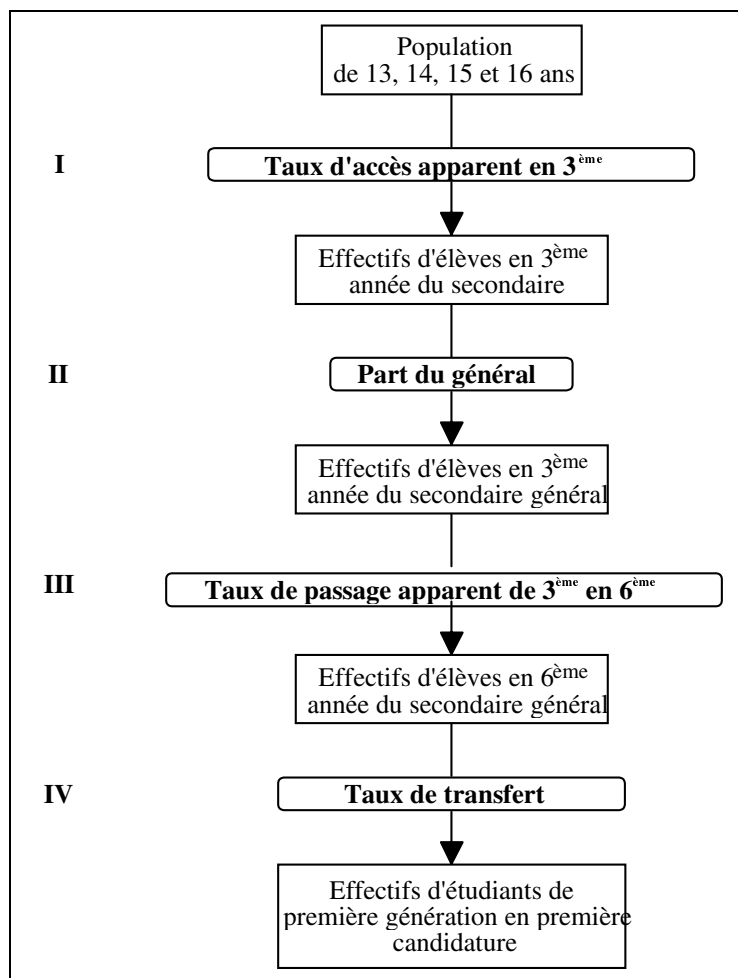
Sauf cas exceptionnels tels que la maladie ou l'immigration et les problèmes de connaissance de la langue française qui peuvent en découler, l'entrée dans le système a lieu en première année primaire à l'âge de 6 ou de 7 ans.

⁷ Ces tables sont établies directement à partir des données publiées par le Service Statistiques du Ministère de l'éducation, avec une reconstitution des données pour les années scolaires 1988-89 et 1989-90, années pour lesquelles les données dans leur structure par âge n'ont pas été publiées.

Les élèves qui effectuent un parcours sans redoublement se « déplacent » sur la diagonale du tableau qui est déterminée par leur position d'entrée. Par contre, ceux qui redoublent se déplacent verticalement, puisqu'ils restent au même niveau au cours de l'année scolaire suivante mais qu'ils vieillissent d'un an.

En ce qui concerne l'école primaire, la scolarité est obligatoire et même si des retards peuvent déjà s'y accumuler, les mesures légales empêchent de plus en plus le redoublement des écoliers⁸.

FIGURE 7. ORGANIGRAMME DE LA PROJECTION.



Le point de départ de notre analyse va se situer en fait en classe de troisième année du secondaire. Cette année est en effet une étape cruciale dans le futur scolaire des élèves puisque

⁸ A cet égard, les récentes mesures légales interdisent purement et simplement le redoublement à l'école primaire et lors de la transition entre la première et la deuxième année d'enseignement secondaire.

c'est à ce stade que l'orientation vers l'une des trois grandes filières (générale, technique, professionnelle) y est prise. Après ce « cap », les réorientations sont rares.

L'état d'avancement des élèves à ce stade constitue donc en quelque sorte une « synthèse » de leur passé scolaire antérieur et à cette étape, « les compteurs sont remis à zéro ». Compte tenu de l'objectif final, nous nous intéresserons aux seuls élèves qui terminent leur scolarité dans des délais jugés raisonnables, c'est-à-dire avec un maximum de deux ans de retard par rapport à l'âge normal de sortie du secondaire. L'analyse des âges des étudiants de première génération a en effet montré clairement que la plupart de ces derniers sont dans un des états suivants :

- avance d'un an,
- à temps,
- retard d'un an.
- retard de deux ans par rapport à l'âge normal d'entrée à l'université (18 ans).

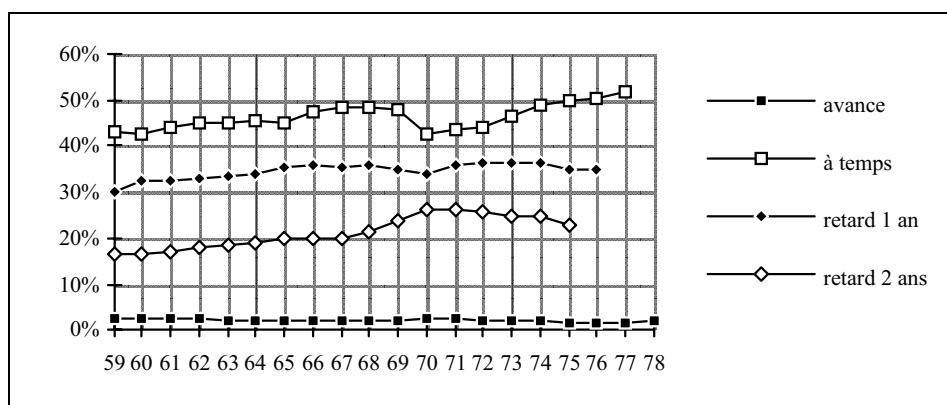
On s'intéressera par conséquent au parcours de ces seuls élèves (en avance, à temps, en retard d'un et de deux ans) au travers de l'organigramme suivant :

Le « détour » par l'étape notée « II » est indispensable en vertu des conditions d'accès à l'université : seuls les élèves en provenance du général et du technique y ont accès et le passage par cette étape permet de les « isoler » des élèves inscrits en professionnel.

Les taux qui définissent le modèle sont des taux apparents, en ce sens qu'ils résultent de la comparaison de deux sous-populations issues de la même cohorte, en deux instants différents statut à statut ; par exemple, on compare le volume d'élèves en retard d'un an en classe de sixième au volume d'élèves en retard d'un an en classe de troisième de la même cohorte, mais trois ans plus tard. Il est clair que le numérateur peut avoir été « alimenté » au cours de ces trois ans par des élèves qui étaient sans retard et qui ont doublé, par exemple, la classe de cinquième. De même, il est possible que des élèves aient quitté le dénominateur entre la troisième et la sixième, soit par abandon, soit par redoublement de classe. Nous sommes forcé de travailler de la sorte par manque de données individuelles.

L'évolution des taux d'accès apparents en classe de troisième, toutes filières confondues, en fonction des états d'avancement sont les suivants :

FIGURE 8. TAUX D'ACCÈS APPARENTS EN CLASSE DE TROISIÈME ANNÉE DU SECONDAIRE, SELON L'ANNÉE DE NAISSANCE.



La croissance des taux d'accès apparents des élèves de 13 et 14 ans est un indice de la meilleure performance du système : en effet, si on trouve plus d'élèves âgés de 13 ou de 14 ans en classe de troisième année, c'est que nécessairement, ces élèves ne sont pas en retard; si tel était le cas, ces élèves se retrouveraient, un ou deux ans plus tard en classe de troisième mais en ayant du retard et on observerait dès lors une augmentation du taux d'accès apparent des classes plus âgées. Or, on observe que le taux d'accès apparent de la classe d'âge de 16 ans commence à diminuer.

Au sein de la troisième année du général, il est intéressant et utile d'analyser les évolutions des catégories d'élèves : « en avance », « à temps », « en retard d'un an » et « en retard de deux ans ». Cette analyse permettra de préciser quel public choisit la filière « générale ».

FIGURE 9. ÉTAT D'AVANCEMENT DES ÉLÈVES EN CLASSE DE TROISIÈME ANNÉE SECONDAIRE GÉNÉRALE SELON L'ANNÉE DE NAISSANCE.

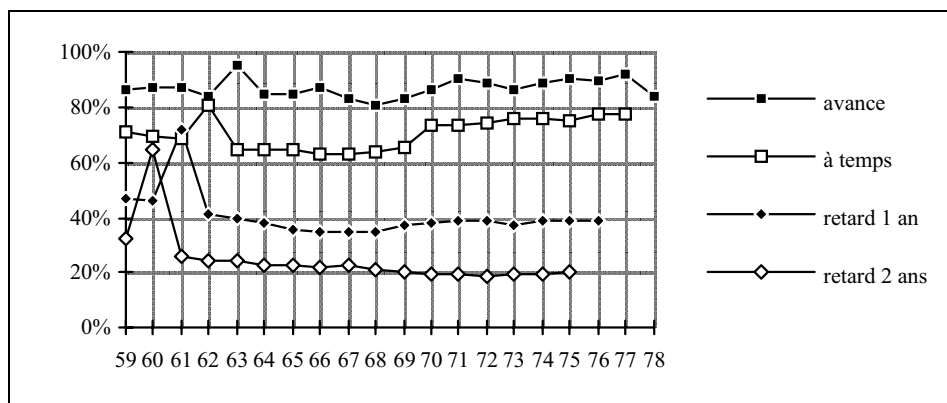
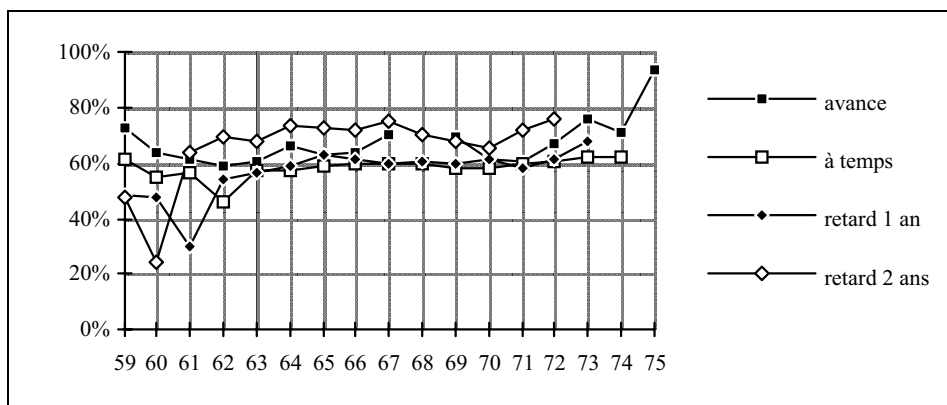


FIGURE 10. TAUX DE PASSAGE APPARENTS ENTRE LA TROISIÈME SECONDAIRE GÉNÉRALE ET LA SIXIÈME SECONDAIRE GÉNÉRAL.



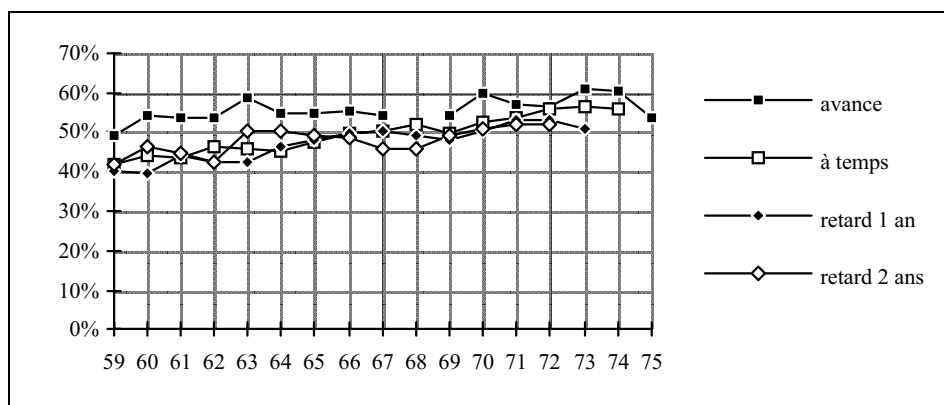
Il est frappant de constater que moins les élèves ont de retard par rapport à l'âge normal, plus ils choisissent (ou sont orientés) dans la filière générale : la majorité des élèves en état d'avancement « en avance » et « à temps », choisissent cette filière générale, ceci indépendamment de leur sexe. Une proportion de 30 % des élèves en retard d'un an s'orientent

dans le général, et une minorité (moins de 25 %) des élèves en retard de deux ans optent quant à eux pour cette filière.

La différence de ces proportions à l'unité se retrouve obligatoirement dans les filières technique et professionnelle, et ceci fournit directement un indice de la hiérarchisation qui existe entre les trois filières : il apparaît que plus les élèves ont du retard, plus ils choisissent les filières technique et professionnelle.

Le dernier point-clé sur lequel repose la projection est celui des taux de transfert apparents de la sixième année générale vers la première candidature. De manière générale, ces taux sont élevés. Des différences existent entre les sexes, mais la croissance est présente dans les deux cas.

FIGURE 11. TAUX DE TRANSFERTS APPARENTS ENTRE LA SIXIÈME ANNÉE DU SECONDAIRE « GÉNÉRAL » ET LA PREMIÈRE CANDIDATURE EN FONCTION DE L'ANNÉE DE NAISSANCE.



L'analyse différentielle de ces taux par sexe a montré qu'au sortir du secondaire, 60 % des garçons des générations civiles récentes s'orientent vers l'université contre 50 % pour les premières générations observées. Chez les filles, on observe aussi une croissance et elles rattrapent les garçons en termes de taux de transfert apparent.

4. Projections d'étudiants de première génération en première candidature à l'université

A ce stade, il n'est pas inutile de rappeler que la démarche suivie pour effectuer les projections s'articule sur quatre points-clés qui sont :

- l'accès en troisième année du secondaire,
- le choix de la filière,
- la transition de 3^{ème} générale en 6^{ème} générale,
- le transfert de 6^{ème} générale vers la 1^{ère} candidature.

Sur base des analyses menées précédemment, nous avons formulé des hypothèses quant à l'évolution des différents paramètres qui interviennent à ces quatre étapes.

Il résulte de l'analyse précédente que les taux qui varient le plus sont

- le taux d'accès en 3^{ème},
- le taux de transfert de 6^{ème} générale vers la 1^{ère} candidature.

Les autres taux (choix de la filière et transition de 3^{ème} générale en 6^{ème} générale) restent statistiquement stables, bien que présentant parfois des fluctuations : dans ce cas, celles-ci ne concernent que des populations faibles en chiffres absolus et dont la variation, somme toute, n'influence pratiquement pas le résultat final.

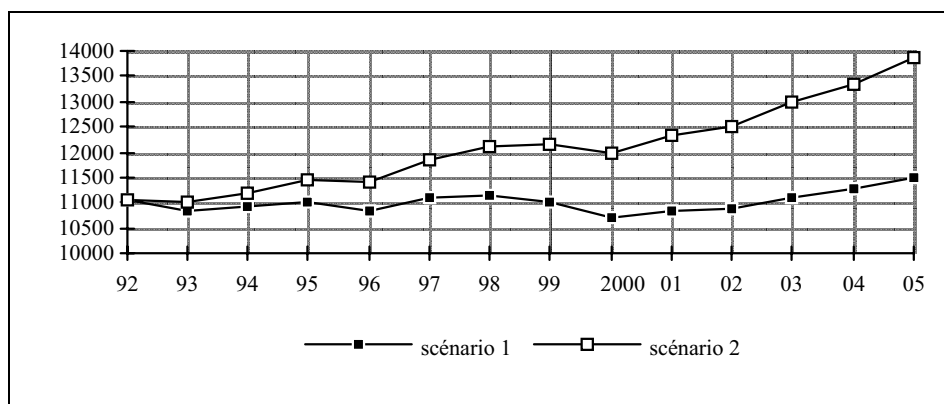
C'est donc sur les deux premiers paramètres que nous avons « joué » dans l'exercice de projection, les deux autres (part du général et transition 3^{ème} - 6^{ème}) restent constants (et valant la moyenne des observations).

A l'horizon retenu, l'année académique 2005-2006, nous avons fixé des valeurs limites et entre ces valeurs et la dernière observation, nous avons procédé à une interpolation linéaire (cf. tableaux annexes).

En dernière étape, nous avons tenu compte des étudiants ayant un retard de plus de deux ans au moyen d'un facteur correctif. Ces étudiants représentent respectivement 7 % du total de garçons et 4,5 % chez les filles.

Différents choix de taux apparents sont évidemment possibles et nous présentons ci-dessous les deux scénarios extrêmes que nous avons retenus et qui reposent donc sur des hypothèses d'évolution raisonnées.

FIGURE 12. PROJECTION D'ÉTUDIANTS DE PREMIÈRE GÉNÉRATION EN PREMIÈRE CANDIDATURE : SCÉNARIOS EXTRÊMES.



Le scénario 1 consiste simplement en la reconduction à l'horizon 2005-2006 des différentes valeurs des taux à leurs dernières observations sur le substrat démographique. Il rend donc compte simplement de l'impact du seul facteur démographique, toutes autres choses égales par ailleurs. Le scénario 2 consiste à améliorer les performances du système scolaire, dans le sens d'une croissance de l'accès en classe de troisième, pour les élèves à temps. Il s'agit donc d'un scénario dans lequel les élèves ont été réorientés en technique ou professionnel avant d'accéder en classe de troisième. Il suppose également une croissance modérée des taux de transfert apparent vers l'université, ce qui correspond à une demande accrue d'enseignement universitaire.

Quoiqu'il en soit, dans les années à venir, le nombre d'étudiants futurs ne devrait donc pas diminuer de manière significative. Au contraire, il est plutôt vraisemblable que le monde universitaire connaisse encore une croissance de ses effectifs à l'avenir et dès lors, de nouveaux modes de financement des universités devront être trouvés.

5. Points de repères internationaux

Bien que la comparaison d'indicateurs de différents systèmes d'enseignement soit une opération assez délicate, on observera que les taux d'accès à l'enseignement supérieur de la Belgique se situent en bonne position par rapport aux autres pays de l'O.C.D.E., puisque plus

de 25 % d'une classe d'âge entre à l'université, et qu'au total, près de 50 % de la population d'une classe d'âge entreprend actuellement des études supérieures.

FIGURE 13. TAUX D'ACCÈS À L'ENSEIGNEMENT SUPÉRIEUR DES PAYS DE L'O.C.D.E. - 1991.

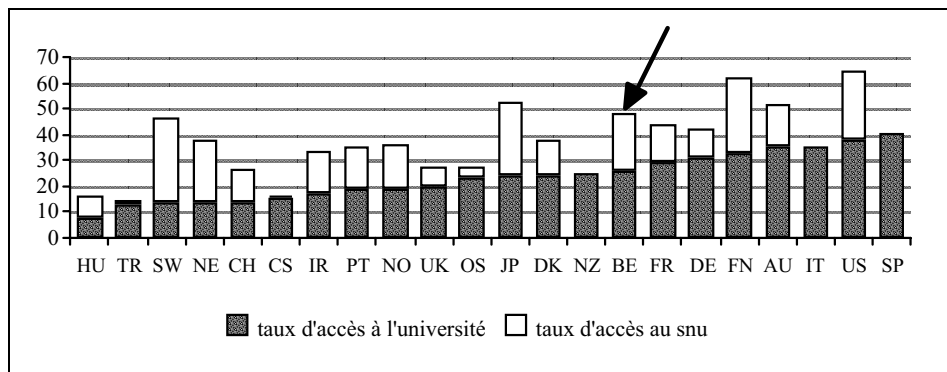
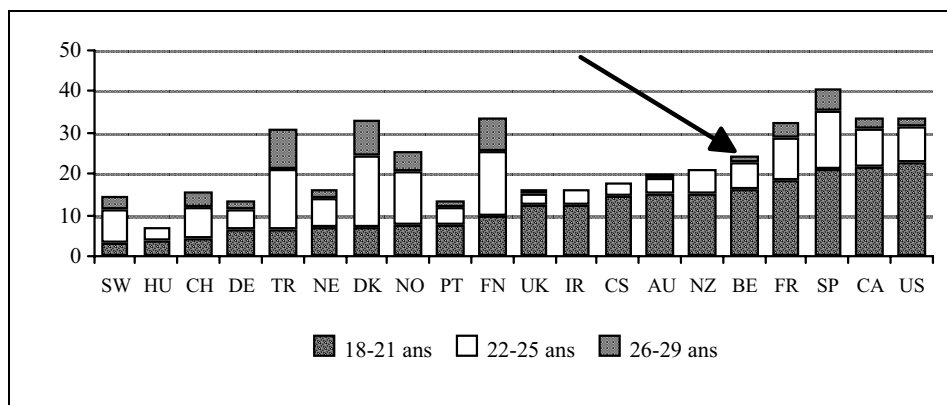


FIGURE 14. TAUX DE PARTICIPATION À L'ENSEIGNEMENT SUPÉRIEUR DES PAYS DE L'O.C.D.E. - 1991.



Les taux de participation des étudiants plus âgés sont assez faibles mais par contre, ceux des étudiants plus jeunes sont assez élevés et ceci doit être interprété comme étant un indice du retard relativement peu important des étudiants en Belgique.

Conclusions

Quel que soit le scénario choisi, tout porte à croire que le nombre d'étudiants de première génération dans les universités devrait augmenter dans les années à venir.

La participation accrue de la population à l'enseignement général, les performances du système secondaire en nets progrès et la demande croissante de formation supérieure universitaire qui se traduit par des taux de transfert apparents en croissance différents, l'ensemble de ces effets sont parvenus à contrebalancer largement ceux de la diminution des effectifs des population susceptibles d'accéder à l'université. En comparaison avec les autres pays de l'O.C.D.E., la situation de la Belgique est assez favorable.

Même si un certain nombre de facteurs influençant favorablement l'accès à l'enseignement supérieur peuvent et doivent encore être améliorés, la situation de la Belgique,

en termes d'accès à l'université, de la Belgique est bonne par comparaison avec les autres pays de l'O.C.D.E., et tout porte à croire qu'il devrait en être de même pour les années à venir.

ANNEXE : Valeurs limites retenues à l'horizon 2005.

PART RELATIVE DE LA 3ÈME GÉNÉRALE

	Garçons	Filles
en avance d'un an	78,6 %	87,7 %
à temps	75,5 %	79,0 %
en retard d'un an	37,8 %	39,6 %
en retard de deux ans	19,6 %	21,0 %

TAUX DE TRANSFERT DE 3ÈME EN 6ÈME

	Garçons	Filles
en avance d'un an	87,4 %	76,1 %
à temps	61,0 %	63,4 %
en retard d'un an	66,1 %	68,9 %
en retard de deux ans	80,7 %	70,6 %

ACCÈS EN TROISIÈME SECONDAIRE

	Observation	Amélioration
Garçons en avance d'un an	2 %	2 %
Garçons à temps	47 %	53 %
Garçons en retard d'un an	37 %	33 %
Garçons en retard de deux ans	27 %	25 %
Filles en avance d'un an	3 %	3 %
Filles à temps	56 %	60 %
Filles en retard d'un an	33 %	30 %
Filles en retard de deux ans	19 %	15 %

TAUX DE TRANSFERT

	Observation	Amélioration
Garçons en avance d'un an	50 %	70 %
Garçons à temps	60 %	65 %
Garçons en retard d'un an	56 %	63 %
Garçons en retard de deux ans	59 %	62 %
Filles en avance d'un an	46 %	65 %
Filles à temps	53 %	65 %
Filles en retard d'un an	45 %	53 %
Filles en retard de deux ans	43 %	50 %

Estimation de population locale pour la gestion publique et le marketing

Giuseppe De BARTOLO, Rosina CHIAPPETTA

Département d'Économie Politique
Université de Calabre

Introduction

Comme on le sait, les événements démographiques sont étroitement liés aux facteurs socio-économiques, par suite, dans le cadre de la programmation des interventions dans ce domaine, on se trouve souvent obligé de lier les décisions à la population de référence. La définition du type et des dimensions de ces interventions nécessite, au préalable, la connaissance de la situation présente aussi bien que l'évaluation des besoins futurs¹. Tout cela implique une conscience de la dimensions de la population mais surtout la connaissance de sa structure selon certains caractères ; en outre, dans le but de saisir les besoins futurs, il est nécessaire de faire des prévisions démographiques.

Souvent, aussi bien l'analyse que la prévision doivent être réalisées sur des zones très limitées, à la suite de l'autonomie de plus en plus forte des entités subrégionales, comme les provinces, les communes, les unités sanitaires de base et les subdivisions scolaires. Ces entités, pour bien programmer les investissements de leur ressources, ont l'exigence de trouver un équilibre socio-démographique et par conséquent il leur est nécessaire, dans leurs choix et décisions, d'avoir les informations les plus fiables au sujet de la démographie du territoire de leur compétence.

A ce niveau territorial très restreint, les analyses et les prévisions démographiques ne sont pas faciles à réaliser à cause de la difficulté d'obtenir les données de base nécessaires ; et si la désagrégation est trop poussée, on ne connaît même pas la dimension démographique. Pour tout cela, on se trouve obligé d'utiliser des méthodes d'estimation et de prévision qui sont conçues pour faire face aux exigences susdites.

Les sources qui peuvent fournir des données pour estimer la population des petites zones sont d'abord des sources classiques : recensement, état civil ; mais on peut exploiter aussi les fichiers de certains organismes publics ou privés pour obtenir des données symptomatiques (comme les inscriptions à l'école primaire, les immatriculations des automobiles, les données sur les permis de construire, les dépôts bancaires, etc.) à travers lesquelles on peut évaluer la dimension démographique.

1. Caractères des méthodes d'estimation pour micro-zones

Nous passerons en revue certaines méthodes d'estimation de la population qui pourraient être employées pour les petites zones. Même si, parmi ces méthodes, certaines ont été construites pour combler des situations particulières (par exemple, pour les États-Unis), elles sont reportées ici car elles pourraient être adaptées aux autres situations². Les estimations concernant la population se distinguent en :

¹ Cfr De Rita, G., Ruberto, A., L'utilizzazione delle previsioni demografiche nella programmazione regionale degli interventi sociali, *Rivista Italiana di Economia, Demografia e Statistica*, V, XXII, n. 1-4, 1968.

² Aux États-Unis, en effet, comme il n'existe pas, comme dans de nombreux pays européens, un registre de la population, la nécessité s'est révélée, plus qu'ailleurs, d'estimer la population durant les années intercensitaires au niveau de grandes et petites aires. Pour faire face à cette exigence, on a élaboré des méthodes non démographiques, de plus en plus perfectionnées avec le temps, qui emploient des variables symptomatiques du niveau de la population.

- Estimations inter-censitaires, appelées ainsi car elles tentent d'estimer le chiffre de la population dans les années comprises entre deux recensements, prenant en considération justement les informations de ces derniers.
- Estimations post-censitaires, qui utilisent les données du recensement précédent.
- Estimations acensitaires c'est-à-dire non basées sur les données censitaires.

Elles peuvent être classifiées aussi par rapport à d'autres caractéristiques : par exemple, elles peuvent être réalisées pour un pays entier, pour les subdivisions géographiques plus importantes ou bien pour une variété d'aires d'un même territoire (aires urbaines et rurales, typologies de villes) ; elles peuvent se référer à toute la population ou à certaines sous-populations différenciées par sexe, âge, état civil, à la population résidente ou à la population présente.

Pour les petites zones, la méthodologie dépend fortement des données disponibles et souvent, elle doit être adaptée à la situation concrète : une méthode conçue pour une certaine configuration de données et pour une situation donnée n'est pas nécessairement applicable à une autre typologie de données et à un autre milieu géographique. En outre, au niveau de petites zones, la qualité inégale de données réduit fortement la précision des estimations. Des méthodes d'estimation construites avec un input peu flexible sont peu adaptables aux champs des données très variables des petites zones. Il faut aussi souligner qu'à la faiblesse de la méthodologie, dont on a conscience, ne s'oppose pas toujours une prudence accrue de la part des utilisateurs des estimations et des prévisions.

2. Quelques méthodes d'estimation

En négligeant les techniques qui emploient des fonctions mathématiques spéciales, et qui pourraient être employées pour faire des interpolations et des extrapolations, dont les caractéristiques et les limites sont bien connues, nous examinerons ici, très brièvement, certaines méthodes qui emploient des variables considérées comme symptomatiques des variations de la population. Ces variables, étant souvent disponibles aussi pour des zones très petites, peuvent être utilisées pour permettre l'estimation de la population d'une zone très restreinte³.

Rapport population/variable

Sans spécifier le nom de la variable symptomatique, l'estimation de la population d'une zone au temps t_2 , en connaissant la valeur de la population et de la variable à un temps t_1 (par exemple au moment du recensement), pourrait être obtenue par :

$$P_{t_2} = \left(S_{t_1} + U \right) \times \frac{P_{t_1}}{S_{t_1}}$$

où P_{t_2} est l'estimation de la population au temps t_2 , S_{t_1} est la valeur de la variable à une occasion précédente et U est la variation nette de la variable dans la période entre la dernière évaluation et le moment de l'estimation.

Étant donné que les valeurs de la variable au moment de l'estimation sont employées pour estimer la population, la précision de l'estimation est fonction de la stabilité dans le temps du rapport variable/population ou de la connaissance des changements de ces rapports.

Souvent les données symptomatiques sont produites pour des raisons très variées et elles ont une couverture géographique qui ne correspond pas à celle de la zone dont on veut estimer

³ A ce sujet voir notamment : Murdock, S.H., Ellis, D.R., *Applied Demography, An Introduction to Basic Concepts, Methods and Data*, Westview Press, Boulder, 1991.

la population. Par conséquent, soit l'échelle temporelle, soit la couverture géographique des variables doivent être convenablement modifiées. Naturellement ces précautions sont extensibles à tous types d'estimation examinés ici. Comme exemple une liste des variables symptomatiques sera donnée après.

Méthode basée sur les habitations

Dans cette méthode, on utilise les données sur les habitations. Les variables plus fréquemment employées sont le nombre et les variations des permis de construire et le nombre et les variations des consommateurs d'électricité. Dans ce cas, l'estimation de la population assume la forme suivante :

$$P_{t_2} = (H_{t_1} + U) \times \frac{P_{t_1}}{H_{t_1}}$$

où P_{t_2} est la population de la zone estimée à un temps postérieur au recensement, H_{t_1} est le nombre des habitations occupées dans la zone considérée au dernier recensement, U est la variation nette de la variable pendant l'intervalle recensement - moment de l'estimation,

$\frac{P_{t_1}}{H_{t_1}}$ est le nombre moyen des personnes par habitation au dernier recensement.

Pour utiliser cette méthode de la façon la plus correcte, il est nécessaire de connaître les données sur les démolitions et aussi sur les nouveaux permis, les premières étant plus fiables que les deuxièmes. Si l'on utilise les données sur les consommateurs d'électricité, on a besoin d'avoir des informations sur la qualité de ces données, sur les usages multiples et sur le nombre d'habitations corrélées à ces usages. Les avantages qu'on a avec cette méthode, c'est que la corrélation entre le nombre d'habitations et la population est forte ; les désavantages sont dus au fait qu'il est difficile d'évaluer les variations du nombre moyen des personnes occupant une habitation. Il a été prouvé que la qualité de cette méthode est susceptible d'amélioration ⁴.

Rapports élémentaires

Si l'on indique par R_1, R_2, \dots, R_n les rapports de certaines variables symptomatiques entre le moment de l'estimation et une date précédente, la moyenne de ces rapports peut être égale au rapport entre les populations aux mêmes dates. C'est-à-dire :

$$\frac{P_{t_2}}{P_{t_1}} = \sum_{i=1}^n \frac{R_i}{n} \Rightarrow P_{t_2} = \sum_{i=1}^n \frac{R_i}{n} \times P_{t_1}$$

Par exemple, si l'on utilise comme variables les naissances et les décès on a :

$$P_{t_2} = P_{t_1} / \left(\left(\frac{n_{t_2}}{n_{t_1}} + \frac{m_{t_2}}{m_{t_1}} \right) \div 2 \right)$$

Méthode des taux bruts

Cette méthode consiste à rapporter, à deux dates différentes n_1 et n_2 , les taux bruts de natalité (ou de mortalité) d'une sous-zone et à égaliser tout cela au même type de rapport pour la zone plus grande. Après, en connaissant les événements (par exemple les naissances) de la sous-

⁴ Smith, S. K., A Review and Evaluation of the Housing Unit Method of Population Estimation, *Journal of American Statistical Association*, v. 81, n. 394, 1986.

zone au moment n_2 , on peut estimer la population totale de la sub-zone au même moment. En indiquant par Q^{sub} un taux brut de la sous-zone et par Q^{sup} le même taux pour la zone, on peut écrire :

$$\frac{Q^{sub} n_2}{Q^{sub} n_1} = \frac{Q^{sup}_{n_2}}{Q^{sup} n_1} \Rightarrow .. Q^{sub} n_2 = Q^{sup} n_1 \times \frac{Q^{sup}_{n_2}}{Q^{sup} n_1}$$

$$\Rightarrow \text{Population totale sous - zone au moment } n_2 = \frac{n \text{ événements sous - zone au moment } n_2}{Q^{sub} n_2}$$

Méthode des proportions

Cette méthode consiste à égaliser les rapports entre la population de la sous-zone et celle de la zone dans deux temps différents, dans l’hypothèse que le rapport sous-zone /zone ne varie pas dans la période considérée. Pour deux temps t_1 et t_2 on a :

$$\frac{P^{sub}_{t_2}}{P^{sub}_{t_2}} = \frac{P^{sub}_{t_1}}{P^{sub}_{t_1}}$$

dont on peut obtenir l’estimation de la population de la sous-zone à l’époque t_2 .

Méthode composée

Ici la population totale est obtenue comme somme des estimations de la population des différentes classes d’âge, estimations obtenues au moyen des variables symptomatiques liées aux classes mêmes. Dans le tableau suivant, on a reporté quelques-uns de ces variables.

Classes d'âges	Variables
0- 5	Naissances des six dernières années
6- 10	Inscriptions à l'école primaire
11- 13	Inscriptions à l'école moyenne
14-49 (sexe féminin)	Taux global de fécondité générale
14-49(sexe masculin)	Rapport entre les sexes
50-65	Taux de mortalité des personnes de 50 ans et plus
65+	Données sur les personnes retraitées

Méthode de régression

Cette méthode peut être vue comme un instrument d'estimation dans lequel on emploie des indicateurs symptomatiques multiples avec poids variables. Voici une liste des variables symptomatiques les plus souvent utilisées :

- naissances
- décès
- inscriptions à l'école primaire
- permis de construire
- usagers de téléphone privés
- usagers électriques
- données sur les votes
- immatriculations des automobiles

– dépôts bancaires

qui peuvent servir pour estimer une relation de régression multiple du type :

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n$$

où y est la valeur de la population, les x_i représentent les variables symptomatiques, b_i les poids obtenus sur la base de l'expérience passée.

La relation précédente est évaluée pour une période antécédente et est supposée valable pour un moment donné ; par conséquent, la connaissance des valeurs des variables symptomatiques permet d'obtenir une estimation de la population. Pour que cette méthode puisse être employée, il faut connaître les valeurs des variables pour chaque année.

La méthode de régression est utilisée pour évaluer les rapports entre la valeur de la population d'une sous-zone et celle de la zone (Ratio-Correlation Method), plutôt que pour faire une estimation directe de la population⁵. Plus précisément le RCM est une technique de régression multivariée où les variables indépendantes sont en général des rapports entre indices de « composition », calculés en temps différents.

Si pour simplifier on se limite à une seule variable symptomatique, qu'ici nous avons indiquée par S, l'équation de régression à estimer aura la forme suivante :

$$\frac{P_{t_2}^{sub} \div P_{t_2}^{sup}}{P_{t_1}^{sub} \div P_{t_1}^{sup}} = b_0 + \frac{S_{t_2}^{sub} \div S_{t_2}^{sup}}{S_{t_1}^{sub} \div S_{t_1}^{sup}}$$

La méthode Ratio-Correlation parmi les techniques d'estimation de population est la plus employée parce qu'elle utilise une pluralité d'indicateurs symptomatiques des variations de la population et en plus elle permet d'employer les estimations de la zone plus grande, qui sont en général robustes. Sa faiblesse au contraire est due au fait qu'on doit accepter passivement toutes les hypothèses du modèle de régression. Différentes variantes de cette méthode ont été proposées, mais aucune ne s'est révélée la meilleure dans toutes les situations.

De l'examen des méthodes d'estimation de la population que nous avons considérées, sommairement, les plus simples, il est possible de tirer quelques leçons. Les techniques symptomatiques sont très utiles pour estimer la population d'une petite aire, soit qu'on ait, souvent, la disponibilité immédiate de nombreuses variables, soit parce que ces techniques sont désormais expérimentées depuis longtemps. La plus grande faiblesse de ces méthodes réside dans la difficulté d'identifier les changements qui interviennent dans les relations entre les variables et la population. Aucune méthode n'est la meilleure dans toutes les situations à cause des grandes différences qui existent entre pays, dans les sources et dans la qualité des données. Par exemple, aux U.S.A., on a observé que les données sur les inscriptions scolaires sont des variables très sensibles des groupes d'âges juvéniles, la même chose vaut pour les décès et l'âge sénile. Dans d'autres pays, ces relations peuvent ne pas être aussi fortes. Le choix des sources et des méthodes pourrait varier en fonction de la nature des statistiques et des estimations voulues. Nous voyons enfin que le choix d'une méthode plutôt que d'une autre est

⁵ Pour un exemple d'estimation de la population italienne qui utilise la régression multiple, voir De Bartolo, G. , *Stima della popolazione presente delle provincie italiane mediante il metodo della regressione multipla*, Saggi di Statistica Metodologica ed Applicata, Contributi del Centro Interdipartimentale di Studi Statistici, v. I, Università degli Studi della Calabria, 1976. Dans cette étude, cinq variables symptomatiques ont été employées, naissances, décès, mariages, inscriptions scolaires élémentaires, relevées au niveau provincial. La relation a concerné la variation de la population dans la période 1961-71, et les variations, à la même période, des variables symptomatiques adoptées. L'équation de régression estimée a été donc utilisée pour évaluer la population présente des provinces italiennes en 1974.

une opération très délicate car, quelque soit la meilleure méthode, cela ne sera clair qu'au moment où l'événement se sera vérifié⁶.

3. Les ordures comme source de connaissance en démographie L'expérience du Garbage Project⁷

Les ordures représentent une source d'information non seulement dans le domaine scientifique mais aussi dans le domaine commercial, social et en particulier dans le domaine démographique, car elles sont un document direct, minutieux et incontestable des habitudes de celui qui les a produites, au-delà même de ses convictions⁸.

En examinant des ordures solides, on peut saisir les préférences d'achat d'un ménage. Leur quantité permet aussi de parvenir à une estimation de la population et donc, c'est un moyen pour évaluer la taille démographique d'une très petite zone, comme un hameau, un quartier et pour évaluer aussi les caractéristiques démographiques des ménages de ces zones.

L'étude des ordures comme moyen de connaissance de la composition sociale d'une population est une discipline relativement nouvelle, qui s'est développée aux États-Unis depuis 1973 et qui est l'oeuvre d'un groupe d'anthropologues de l'Université de l'Arizona.

En se référant à cette expérience, très fructueuse et de longue durée, et en particulier au Garbage Project (GP) développé dans cette université, il a été mis en évidence que, par l'examen des ordures solides, il est possible d'avoir une idée précise du schéma de consommation de certains groupes particuliers (Latino-américains, Anglo-saxons). En outre, on a pu constater que ces types de corrélation sont très stables et qu'ils ont une validité générale soit pour ce qui concerne le comportement d'achat, soit pour ce qui concerne la liaison entre masse de population et masse d'ordures solides.

En 1986 le Bureau du Recensement américain a demandé le soutien du Garbage Project pour tenter de résoudre le problème chronique de la sous-estimation de la population dans certaines zones critiques : ghettos, zones à forte immigration clandestine, etc. Un des buts du GP a été d'évaluer si, en général, il était possible de reconstruire la structure d'un quartier par sexe et par âge sur la base de la quantité et de la qualité d'ordures solides. Ce projet a été confié à des archéologues qui sont familiarisés avec ce genre de recherche. Dans ce cas particulier, ils avaient comme but la détermination d'un nombre « magique » qui, appliqué à certaines quantités de certaines types d'ordures solides, aurait pu permettre une bonne estimation de la population. Il faut souligner que les chercheurs du GP avaient l'avantage d'avoir à leur disposition toutes les informations sur la consommation alimentaire et, surtout, sur la production d'ordures de 63 familles observées sur une période de temps suffisamment longue. Grâce à cette base de données, il a été possible d'avoir une idée correcte du schéma de production d'ordures d'une famille américaine type.

Le premier objectif était de passer de la quantité d'ordures ou de la quantité d'un certain type d'ordures à l'estimation de la population totale : c'est-à-dire déterminer le nombre « magique » qui, multiplié par le poids total d'ordures produites, ou multiplié par le poids d'un des seize types d'ordures classifiées par un certain groupe de familles, puisse permettre une estimation de la population. Pendant cette recherche, on s'est aperçu que les catégories « total ordures solides » et « plastique » ont le plus fort pouvoir prédictif. Réduite à l'essentiel, voici la relation trouvée qui lie la production d'ordures de plastique à la population :

⁶ Purcell, N. J., and Kish, L., Postcensal Estimates for Local Areas (or Domains), International Statistical Review, 48 (1980), p. 7.

⁷ Cfr Rathje, W. and Murphy, C., Garbage Demographics, American Demographics, May 1992. -Rubbish, The Archaeology of Garbage, Harper Collins Publisher, 1992.

⁸ Viale, G., Un mondo usa e getta, Feltrinelli, Milano, 1994, pp. 28-36.

$$Y = 0,2815 X$$

où **X** représente la quantité de plastique en livres, produite pendant cinq semaines et **Y** la taille du ménage.

Les chercheurs reconnaissent que l'estimation de la population totale obtenue soit en employant la catégorie « total ordures solides », soit la catégorie « plastique » a un écart maximal de plus ou moins 2,5% par rapport à la valeur réelle de la population.

La recherche, visant à estimer la structure par sexe et par âge, a été plus difficile. La classe d'âge qu'on peut plus facilement estimer est celle des enfants. Les couches pour nouveau-nés sont un bon indicateur de cette tranche d'âge parce que la régularité et le volume de ce genre de déchets aident à éliminer les distorsions statistiques.

Voici la relation qui a été trouvée et qui relie la population au nombre d'enfants :

$$Y = 0.01506 X$$

où **X** représente le nombre de couches pour nouveau-nés trouvées dans la poubelle pendant une semaine et **Y** la population infantile correspondante.

En réalité, il y a des déchets « exclusifs » qui pourraient permettre d'estimer la proportion entre les sexes, lier les enfants aux vieux, mais ils n'ont pas la même force prédictive que les couches pour nouveau-nés, car ils sont produits en quantité et fréquence insuffisantes pour ce type d'analyse.

Les chercheurs du Projet ont réussi à déterminer une estimation de la proportion des nouveau-nés sur la base des paramètres suivants :

- taux moyen d'ordures jouets et boîtes : 2,52 par enfant par semaine ;
- taux moyen d'ordures vestimentaires et boîtes : 0,87 par enfant par semaine.

Il a été possible également d'obtenir la proportion de femmes adultes dans une population sur la base des catégories suivantes d'ordures féminines :

- produits hygiéniques féminins : 1,58 par femme par semaine ;
- produits de beauté : 0,86 par femme par semaine ;
- articles d'habillement : 0,62 par femme par semaine.

L'objectif du Projet qui voulait trouver des indicateurs robustes de la présence masculine dans les ordures solides a échoué : en effet, les femmes peuvent boire et manger comme les hommes ; elles fument des cigarettes comme les hommes. Les mégots de cigares ou les boîtes de lingerie masculine pourraient être des indicateurs exclusifs pour le sexe masculin, mais leur présence dans les poubelles n'est pas en quantité suffisante pour établir une relation forte.

Pour résoudre le problème de la détermination numérique de la population masculine, une voie indirecte a été indiquée : de la population totale on soustrait d'abord les nouveau-nés, ensuite les enfants et enfin les adultes de sexe féminin. L'estimation ainsi obtenue est accréditée avec un écart de plus ou moins 10% vis-à-vis de la vraie valeur.

Le Bureau de Censur a jugé qu'il était dangereux, pour un bureau gouvernemental d'employer des gens pour examiner les ordures et il a donc renoncé, à ce moment là, à utiliser cette méthodologie de correction. Mais ce projet pourrait être appliqué en l'an deux mille.

Conclusions

Dans cette communication, nous avons souligné l'importance de la composante démographique dans la programmation des interventions sociales et économiques. Cette activité de programmation, qui est aujourd'hui de plus en plus fréquente au niveau local (communes, unités sanitaires de base, districts scolaires) ne nécessite pas seulement la connaissance de la population, mais aussi, souvent, de sa structure selon certaines caractéristiques. A ce niveau, les données pour faire une analyse démographique adéquate aux objectifs, sont souvent insuffisantes et tout cela affaiblit considérablement les efforts pour ancrer la programmation à des bases solides. Ainsi, nous avons examiné certaines méthodes d'estimation de la population qui nous paraissent utiles pour les petites zones, et dont nous avons fait une analyse critique. Enfin, nous avons souligné les potentialités informatives des ordres pour saisir la composition sociale, démographique d'une population, et aussi pour connaître son comportement d'achat. Cette approche pourrait avoir des utilisations plus générales dans un future proche.

La démographie comme aide à la gestion locale

Thierry EGGERICKX

GÉDAP - Centre d'étude de Gestion Démographiques pour les Administrations Publiques
Université Catholique de Louvain

La population constitue l'essence même d'une commune et la finalité première de l'institution communale est de mettre tout en oeuvre pour assurer le bien-être de ses habitants. Dans ce cadre, nos objectifs sont de sensibiliser les responsables locaux aux réalités démographiques actuelles et de leur permettre de les intégrer à la gestion de leur commune.

D'une manière générale, la variable population est plus ou moins intimement liée à tous les autres domaines de la gestion locale. Par exemple, elle est déterminante pour la prévision des ajustements nécessaires en matière d'équipement communal (crèches, écoles, homes, infrastructures de loisirs...) et la connaissance des variabilités démographiques au sein de la commune permettra d'optimiser leur localisation spatiale. *« L'ouverture d'une nouvelle classe primaire, la suppression d'une classe maternelle devront tenir compte de l'évolution prévisible de la population en âge scolaire, de sa répartition spatiale mais également de celle des jeunes ménages susceptibles d'engendrer les futurs élèves. La situation d'une maison de repos, d'un club du troisième âge devra prendre en compte la proximité des aînés de la commune au même titre que des éléments relatifs à l'organisation de l'espace comme l'aménagement du territoire et les moyens de communication »* (Poulain, 1988-1).

De même, un dépistage et une localisation précise des populations à risque - personnes isolées, personnes âgées, familles monoparentales, population étrangère... - concourent certainement à accroître l'efficacité de la politique sociale. Sans entrer dans le détail, les politiques communales liées au logement, à la sécurité du citoyen, à l'environnement, ... les finances communales entretiennent également des liens plus ou moins étroits avec la démographie.

Les observations démographiques doivent donc s'intégrer dans la panoplie des outils d'aide à la décision. Mais avant d'explicitier la démarche entreprise depuis quelques années à l'U.C.L. par M. Poulain et son équipe, deux questions importantes doivent être posées :

- Une démographie des populations locales est-elle techniquement possible ?
- Quelles sont les spécificités de la démographie communale ?

1. Une démographie des populations locales est-elle techniquement possible ?

En filigrane de cette question se posent celles de la disponibilité et de la fiabilité des données, de l'adéquation des méthodes d'analyse classiques de la démographie. Enfin, au-delà de l'aspect technico-scientifique du problème, on peut s'interroger sur la motivation des acteurs, les scientifiques et les responsables de la gestion locale.

1.1 Le Registre national comme support indispensable de la démographie locale en Belgique

Hormis dans les pays qui disposent d'un registre de population, centralisé ou non, les statistiques démographiques de base ne sont que rarement disponibles, de façon permanente, à l'échelon communal. La structure par âge des populations communales n'est généralement connue qu'au moment du recensement, soit en moyenne tout les dix ans, tandis que les mouvements migratoires, selon l'âge et le sexe, brillent par leur absence, au même titre que les naissances ventilées selon l'âge de la mère et les décès répartis selon l'âge et le sexe. L'inexistence de certaines données de base dans les publications statistiques officielles constitue bien entendu un obstacle de taille à la démographie locale. En outre, compte tenu de la dynamique démographique dont se caractérisent les populations locales, les pays où la

plupart des informations sont fournies par le recensement sont pénalisés par la périodicité - le plus souvent décennale - de celui-ci.

Pour l'étude démographique des populations locales, le registre de population constitue un atout indéniable. Par opposition au recensement qui propose un cliché statique de la population à un moment donné, le registre de population est une source dynamique qui garde en mémoire l'ensemble des faits démographiques ponctuant la vie d'une personne et qui, en outre, peut être interrogée à tout moment. Si les faits d'état civil - naissances, mariages et décès - sont correctement enregistrés, c'est généralement au travers des changements de résidence que les problèmes de fiabilité se posent (Poulain, 1994). Par négligence des déclarants et de l'administration ou, compte tenu des avantages et des inconvénients qui en résultent pour les deux parties, les registres de population pèchent souvent par une sous-déclaration des départs. Seules les vérifications post-censitaires peuvent apporter les corrections nécessaires. Quelle est la situation de la Belgique ?

La Belgique s'est dotée, officiellement depuis 1983, d'un outil d'une richesse démographique inestimable : le Registre national. Cette banque de données, informatisée et centralisée, construite sur base des registres de population communaux, est perpétuellement enrichie des modifications - naissances, décès et changements de résidence - qui surviennent dans la population de chaque commune de Belgique. *Celles-ci peuvent donc connaître à n'importe quel moment l'état et l'évolution les plus récents de leur population.* Si les données existent, il reste encore à s'interroger sur leur fiabilité. Au cours des premières années de son existence, l'un des principaux problèmes du Registre national, fut l'existence de doubles comptes qui touchaient environ 1% de la population (INS, 1990). A la faveur du renouvellement des cartes d'identité, le fichier national fut pratiquement épuré de ces doubles compte. La centralisation et l'informatisation de l'information accroissent sa fiabilité en assurant une vérification des données à chaque interrogation du fichier. Dans le cas de la mobilité interne, il y a désormais une adéquation exacte entre les entrées et les sorties car chaque inscription dans une nouvelle commune entraîne automatiquement une radiation dans la commune de départ. A l'inverse, une sous-déclaration des émigrations subsiste dans le cadre de la mobilité internationale, certaines personnes omettant de déclarer leur départ ou certaines administrations « oubliant » de les rayer de leur fichier. En Belgique, comme dans d'autres pays, il convient de rappeler qu'il existe des seuils « fatidiques » de population en deçà desquels les localités ne peuvent descendre, sous peine de perdre certains avantages (Poulain et al., 1995).

Enfin, l'autre faiblesse du Registre national concerne les personnes ayant plusieurs résidences possibles. Il ne s'agit plus ici d'omissions ou d'enregistrements erronés, mais d'ambiguïtés liées à la notion même de résidence légale. Cela concerne par exemple les jeunes émancipés qui conservent leur résidence légale chez leurs parents, le ménage qui possède deux résidences et qui privilégie sa domiciliation dans l'une d'elle pour des raisons fiscales indépendamment de la durée de résidence, ou encore des personnes âgées qui demeurent dans des homes, alors que leur résidence légale est maintenue ailleurs (Poulain, 1994).

En dépit de ces quelques problèmes, on peut admettre qu'aujourd'hui le Registre national a atteint un niveau de fiabilité remarquable (INS, 1990). C'est en tout cas une banque de données de premier ordre pour la démographie locale.

1.2 Les méthodes classiques de l'analyse démographique et la démographie locale

Le Registre national permettant une observation suivie et exhaustive de la population de chaque commune, rien ne s'oppose à l'application des méthodes démographiques classiques, qu'il s'agisse de l'analyse transversale ou longitudinale. Néanmoins, la démographie communale, et a fortiori infra-communale, se heurte au problème des petits nombres et par extension, à la fiabilité statistique de certains indicateurs démographiques. Ce problème concerne avant tout les mesures

les plus représentatives de la mortalité (espérance de vie, quotient de mortalité...), et à un niveau moindre, celles de la fécondité. Quelques précautions élémentaires s'imposent :

- fixer un seuil minimal de population en deçà duquel certains indicateurs ne pourront être calculés ;
- considérer les événements - naissances, décès et migrations - sur plusieurs années d'observation ;
- calculer les intervalles de confiance qui permettront d'estimer les résultats à leur juste valeur.

1.3 La motivation des démographes et des responsables locaux

Hormis la faisabilité technico-scientifique, il n'y a pas de démographie locale possible sans la bonne volonté du monde scientifique, d'une part, de celui des décideurs locaux, d'autre part. Or, non seulement on déplore une participation presque inexistante des démographes dans les organes de gestion locale, mais, de façon générale, ils ne sont guère nombreux à s'intéresser à la dimension locale des réalités démographiques. La démographie locale a souvent mauvaise réputation, car assimilée aux travaux inlassablement répétitifs des bureaux d'étude. C'est évidemment méconnaître l'extrême diversité du champs démographique local, c'est ignorer que derrière les inévitables moyennes nationales et régionales, se dissimulent des comportements très variés. N'est-ce-pas au plus près des gens et des choses que l'on peut le mieux comprendre comment et pourquoi les changements démographiques se manifestent et se développent ? Enfin, pour le « migratologue » et le spécialiste des phénomènes d'urbanisation, les communes constituent un terrain d'investigation essentiel.

Parallèlement au manque d'intérêt apparent des démographes pour le monde local, on peut également déplorer la méconnaissance des gestionnaires locaux face aux potentialités de l'approche démographique et statistique. Bien entendu, tout bourgmestre est attentif à l'évolution du chiffre de population de sa commune. Une population en croissance n'est-elle pas souvent interprétée comme le signe de la bonne santé démographique d'une commune et de sa gestion efficace, argument électoral oblige ! Ensuite, le chiffre de population, et plus précisément certains seuils déterminent le nombre de conseillers communaux, le nombre d'échevins, le traitement du secrétaire communal et du Bourgmestre... Mais combien, par exemple, ont une perception réellement précise de la répartition spatiale de la population et de ses spécificités démographiques (type de famille, structure par âge, par nationalité...) et socio-économiques (niveau de vie, profession, éducation...) ? Combien sont au courant du potentiel prévisionnel, à court et moyen termes, des perspectives démographiques ? N'est-ce-pas là un outil de décision de tout premier ordre ?

Il y a donc une démarche de sensibilisation aux réalités démographiques et aux potentiels de l'analyse démographique et statistique, à entreprendre auprès des autorités communales. Les démographes doivent susciter cette demande et l'expérience que nous avons acquise en Wallonie et en Région Bruxelloise, au cours de ces dernières années, nous a conforté dans le sentiment que les responsables locaux n'étaient pas insensibles aux arguments des démographes, à condition bien entendu d'utiliser un langage, dans tous les sens du terme, intelligible.

2. Les spécificités de la démographie locale

2.1 Le rôle prédominant des migrations

Les mouvements migratoires ont une influence déterminante, non seulement sur l'évolution, à la hausse ou à la baisse, du chiffre de la population d'une commune, mais également sur ses caractéristiques démographiques (selon l'âge, le sexe et la nationalité), socio-économiques et

culturelles. A l'échelle des communes, les mouvements migratoires se caractérisent par leur fréquence, leur poids dans la croissance démographique et leur sensibilité conjoncturelle.

2.1.1 La fréquence des mouvements migratoires à l'échelon local

Pour les communes, les migrations sont des événements démographiques beaucoup plus fréquents que les naissances ou les décès. En Belgique, au cours de l'année 1992, 921 000 personnes ont effectué un changement de logement alors que l'on dénombre 124 000 naissances et 104 000 décès (Poulain, 1994). Plus on descend la hiérarchie des unités administratives (pays, région, province, arrondissement, commune, quartier), plus l'intensité des mouvements migratoires est élevée. L'explication de cette particularité est double mais néanmoins complémentaire. Le premier volet prend en considération l'impact de la distance. Plus la distance entre le lieu de départ et celui d'arrivée est grande, moins on enregistre de migrations entre ces deux lieux. En d'autres termes, les migrations se font avant tout sur de petites distances (Baccaini, 1994). Ainsi, parmi les 824 740 changements de logement enregistrés en Belgique en 1991, 53 % ont été réalisés à l'intérieur de la même commune alors que 22 % ont occasionné le transfert d'un arrondissement à un autre. En France, lors de la période censitaire 1982-1990, la proportion de population mobile diminuait de près de 10 % par an, si on considérait les changements de logement, à 6 % dans le cadre des changements de commune, à 3 % pour les changements de département et à 2 % pour les changements de région (Pumain, Courgeau, 1993).

Le second volet de l'explication, qui n'est pas indépendant du premier, prend en considération le nombre des unités administratives de référence. En effet, pour un pays, l'évolution du chiffre de la population ne pourra être affecté que par le bilan des migrations externes (ou internationales), celles qui ont, comme lieu destination ou comme lieu d'origine, un autre pays. Les migrations internes au pays n'interviennent pas dans le calcul car chaque événement n'est pas sanctionné par le gain ou la perte d'un habitant. Le bilan démographique des 3 régions de la Belgique sera tributaire des migrations internationales mais aussi des échanges migratoires interrégionaux. Pour une commune belge, les migrations internationales interviennent toujours, mais il faudra également tenir compte des mouvements avec les 588 autres communes.

La fréquence des mouvements migratoires, à l'échelon communal et intra-communal, en fait des éléments déterminants de la dynamique démographique.

2.1.2 Les migrations et la dynamique démographique des communes et de leurs quartiers

Le poids du bilan migratoire dans la croissance démographiques des communes et des quartiers

Aujourd'hui, ce sont avant tout les phénomènes migratoires qui déterminent l'évolution démographique des communes et des quartiers. Une étude, réalisée sur 162 communes wallonnes et bruxelloises et couvrant les années 1982-1987, a montré que dans 75 % des cas, l'impact du bilan migratoire sur l'évolution du chiffre de la population de la commune l'emportait sur celui du bilan naturel. Le coefficient de corrélation entre le bilan migratoire et le taux d'accroissement de la population était de 0,87 pour 0,44 entre ce dernier et le bilan naturel (Eggerickx, 1990). Les mêmes calculs ont été réalisés sur base de 136 quartiers des communes d'Aiseau-Presles, Braine-le-Château, Charleroi et Namur. Les résultats sont tout aussi significatifs avec un indice de corrélation de 0,94 entre le bilan migratoire et le taux d'accroissement de la population et de 0,35 entre celui-ci et le bilan naturel (figures 1 et 2).

FIGURE 1. LA CORRÉLATION ENTRE LE TAUX D'ACCROISSEMENT DE LA POPULATION ET LE BILAN MIGRATOIRE : QUARTIERS D'AISEAU-PRESLES, BRAINE-LE-CHÂTEAU, CHARLEROI ET NAMUR

(1989-1994) (136 OBSERVATIONS) ($R = 0,94$)

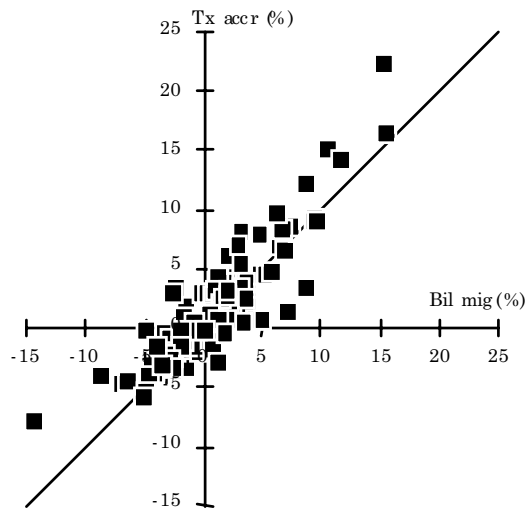
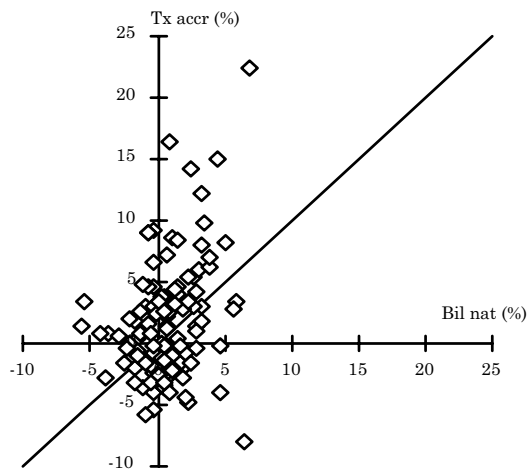


FIGURE 2. LA CORRÉLATION ENTRE LE TAUX D'ACCROISSEMENT DE LA POPULATION ET LE BILAN NATUREL : QUARTIERS D'AISEAU-PRESLES, BRAINE-LE-CHÂTEAU, CHARLEROI ET NAMUR (1989-1994)

(136 OBSERVATIONS) ($R = 0,35$)



Les migrations et le renouvellement des populations locales

L'intensité des migrations n'est pas sans incidence sur le renouvellement des populations locales. Pour 40 communes wallonnes, observées entre 1988 et 1993, nous disposons du cursus démographique des habitants recensés en début de période d'observation, selon qu'ils sont restés dans la commune, qu'ils ont émigré ou qu'ils sont décédés. Les personnes nées et celles immigrées pendant les années 1988-1993 sont ventilées selon les 3 mêmes critères. Nous connaissons ainsi la part relative de la population sédentaire (présente en t et $t+5$) et celle de la « population nouvelle » (les personnes nées et celles immigrées entre t et $t+5$ et présentes en $t+5$). Il est donc possible, d'une part, de calculer un indice de renouvellement de la population en rapportant simplement les effectifs de la « population nouvelle » sur la population totale en $t+5$, et d'autre part, de préciser l'influence respective des mouvements naturel et migratoire sur le renouvellement de la population.

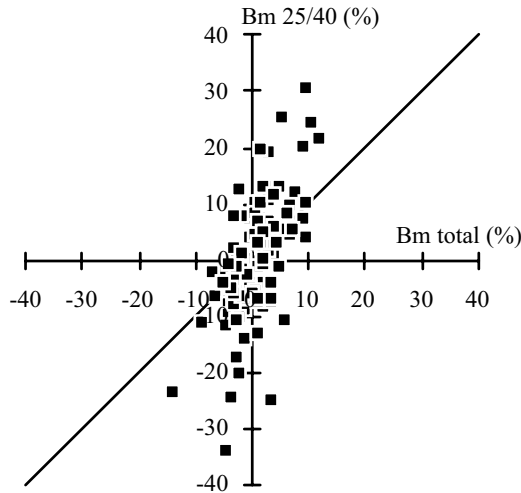
Sur une période de 5 ans, on observe en moyenne 20 % de « population nouvelle » parmi les communes considérées et la mobilité y joue un rôle trois fois plus important que le mouvement naturel. Cette influence déterminante du mouvement migratoire sur le renouvellement des populations locales se vérifie dans chaque commune. En descendant à l'échelle des quartiers, le renouvellement de la population peut toucher près de la moitié de la population, avec une responsabilité quasi-exclusive du mouvement migratoire (Poulain, 1994).

L'effet indirect des migrations sur la structure par âge de la population

Compte tenu de l'effet sélectif de la migration selon l'âge, les mouvements migratoires peuvent être à l'origine de glissements de structure par âge au sein de la population totale. L'intensité du bilan migratoire total est souvent liée à l'attraction ou à la répulsion exercée par une commune sur l'un ou l'autre groupe d'âges particulier. Ainsi, sur un échantillon de 162 communes bruxelloises et wallonnes, observées entre 1982 et 1987, le coefficient de corrélation linéaire entre le bilan migratoire total et celui des personnes âgées de 25 à 40 ans en 1987 était de 0,78. Ce même coefficient était de 0,60 entre le bilan migratoire total et le bilan migratoire des 50-60 ans (Eggerickx, 1990). Le même calcul a été réalisé pour les 136 quartiers des communes d'Aiseau-Presles, Braine-le-Château, Charleroi et Namur, entre 1989 et 1994. Le coefficient de corrélation linéaire entre le bilan migratoire total et celui des 25/40 ans est de 0,78 (figure 3), alors qu'il n'est guère significatif si l'on prend en considération le bilan migratoire des personnes âgées de 50-60 ans ($r = 0,33$).

En conclusion, le rôle fondamental joué par la migration constitue l'une des spécificités de la démographie locale. Les mouvements migratoires déterminent dans une large mesure l'évolution du chiffre de la population des communes, redessinent sans cesse leur portrait démographique, modifient la composition des ménages et des familles et par delà, transforment progressivement les mentalités et les spécificités propres à chaque lieu. Enfin, outre leur fréquence, les migrations se distinguent de la natalité et de la mortalité par leurs fluctuations. Alors que les composantes du mouvement naturel s'inscrivent dans des tendances historiques longues, les migrations sont des phénomènes très conjoncturels, beaucoup plus sensibles. C'est donc sur les comportements migratoires que se focaliseront la plupart des décisions politiques qui, au niveau local, ont une portée démographique (Poulain, 1994). Les propos du Bourgmestre de la ville de Charleroi, tenus lors de la Chaire Quetelet 1988, sont révélateurs des enjeux que suscitent les mouvements migratoires via, dans ce cas, la politique de l'aménagement du territoire.

FIGURE 3. LA CORRÉLATION ENTRE LE BILAN MIGRATOIRE DES PERSONNES ÂGÉES DE 25-40 ANS EN 1994 ET LE BILAN MIGRATOIRE TOTAL : QUARTIERS D'AISEAU-PRESLES, BRAINE-LE-CHÂTEAU, CHARLEROI ET NAMUR (1989-1994) (136 OBSERVATIONS) ($R = 0,78$)



« Ce sont les habitants qui fournissent aux communes leurs moyens financiers par le biais des centimes additionnels à l'impôt sur les personnes physiques et des taxes non rémunératoires appelées jusqu'il y a peu taxes indirectes. Ils influencent aussi le niveau de subside, notamment au sein du Fonds des Communes.

Ce sont aussi les habitants qui constituent l'assiette institutionnelle de la commune puisque le nombre des conseillers communaux et des échevins est déterminé par référence au nombre d'habitants.

Mais il semblerait que l'or noir se fasse plus rare [...] aucune des 5 grandes villes belges n'a échappé au problème du dépeuplement [...] La perte de population est due en partie à la dénatalité, phénomène sur lequel l'aménagement du territoire n'a que peu d'influence [...] Par contre, l'influence est plus directe lorsqu'il s'agit d'oeuvrer à la fixation des populations dans les grands ensembles. Autrefois, avant la crise des finances publiques, on pouvait attirer ou simplement maintenir la population par une fiscalité communale attrayante, par des infrastructures culturelles ou sportives prestigieuses, par des services diversifiés et bon marché. L'âge d'or étant à jamais révolu, les municipalistes ont dû trouver des politiques de substitution et l'aménagement du territoire en est une.

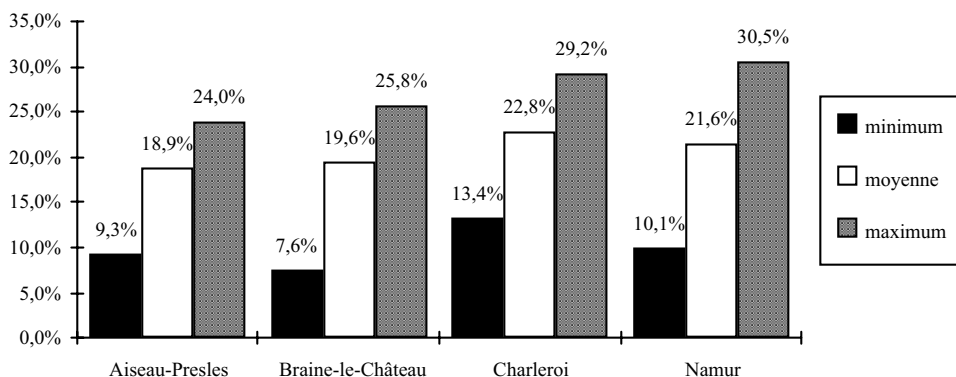
Aujourd'hui on travaille sur le cadre de vie, sur le logement, sur les transports en commun et sur l'image de la ville en vue de procurer aux « citoyen clients », un niveau de satisfaction les poussant à venir s'installer dans telle ville ou tout simplement à ne pas la quitter » (Van Cauwenberghe, 1988).

2.2 La diversité des structures et des comportements démographiques à l'échelon infra-communal

L'expérience nous a démontré que dans la plupart des cas, les différences de structures et de comportement démographiques que l'on pouvait observer entre les quartiers d'une même commune étaient nettement plus prononcées que celles qui distinguaient cette commune, dans son ensemble, de ses voisines. Les quelques exemples qui suivent suffisent à illustrer notre propos (figures 4, 5 et 6). Nous pouvons caractériser les 4 communes choisies de la manière

suivante : Charleroi, ancien centre industriel, est aujourd'hui une métropole de plus de 200 000 habitants qui a subi de plein fouet la crise de la désindustrialisation ; Namur, ville historique, a une population supérieure à 100 000 habitants et est aujourd'hui la capitale de la Wallonie ; Aiseau-Presles est une ancienne commune industrielle de 10 000 habitants alors que Braine-le-Château est une petite commune « huppée » de 8 000 habitants, située à la périphérie de l'agglomération bruxelloise.

FIGURE 4. LA PROPORTION DE PERSONNES ÂGÉES DE PLUS DE 60 ANS, AU 1-1-1994



En dépit de ces différences, on constate que les valeurs moyennes des indicateurs démographiques choisis sont assez semblables. Inversement, les valeurs minimum et maximum mettent en évidence de larges différences à l'intérieur du périmètre communal. A l'échelon infra-communal, c'est l'hétérogénéité démographique qui prime et la gestion communale se doit d'agir en conséquence. Si la diversité constitue une richesse en soi, elle peut également engendrer des tensions et l'apparition de problèmes démo-socio-économiques.

FIGURE 5. L'INDICE CONJONCTUREL DE FÉCONDITÉ, ENTRE 1989 ET 1994

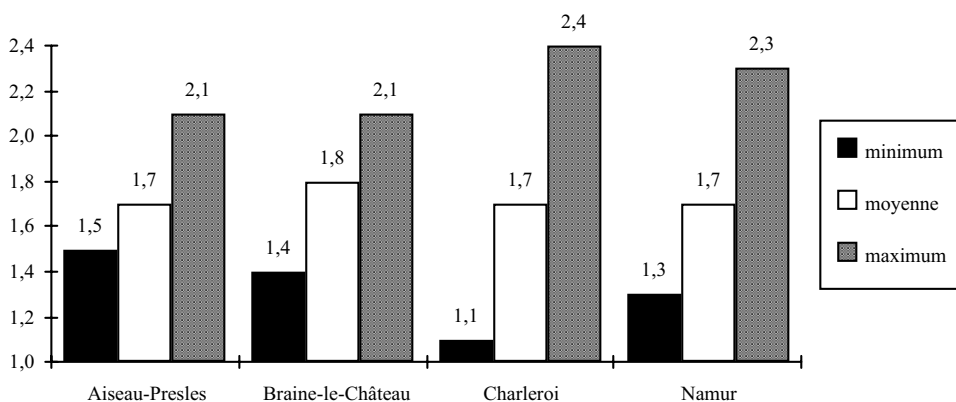
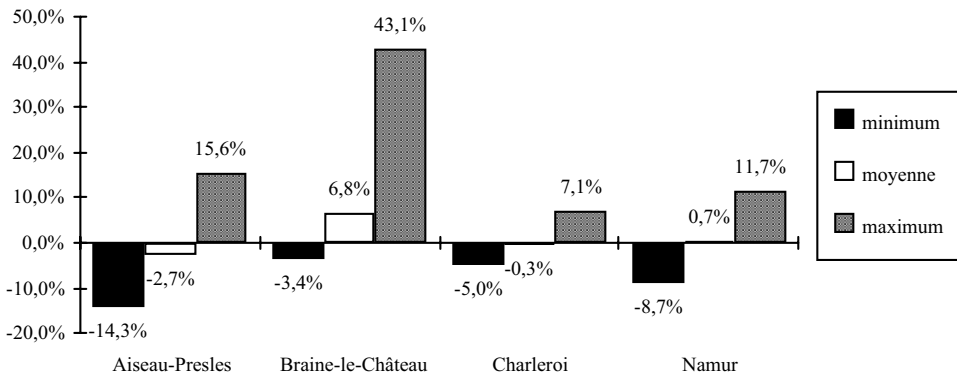


FIGURE 6. LE TAUX D'ACCROISSEMENT MIGRATOIRE, ENTRE 1989 ET 1994



3. Démographie et gestion locale : le rôle du démographe

En tant que démographe, notre action auprès des communes se polarise autour de trois axes interactifs :

- informer les responsables locaux sur les caractéristiques démographiques de leur commune et de ses quartiers ;
- mettre en évidence des situations « anormales » ou qui risquent de le devenir ;
- répondre à une problématique bien précise.

3.1 La description des caractéristiques démographiques d'une commune et de ses quartiers

Notre première tâche est d'informer les responsables communaux sur la situation démographique de leur commune et de ses quartiers. Une gestion locale optimale doit impérativement reposer sur une connaissance actualisée des réalités démographiques et de ses différences à l'intérieur du périmètre communal.

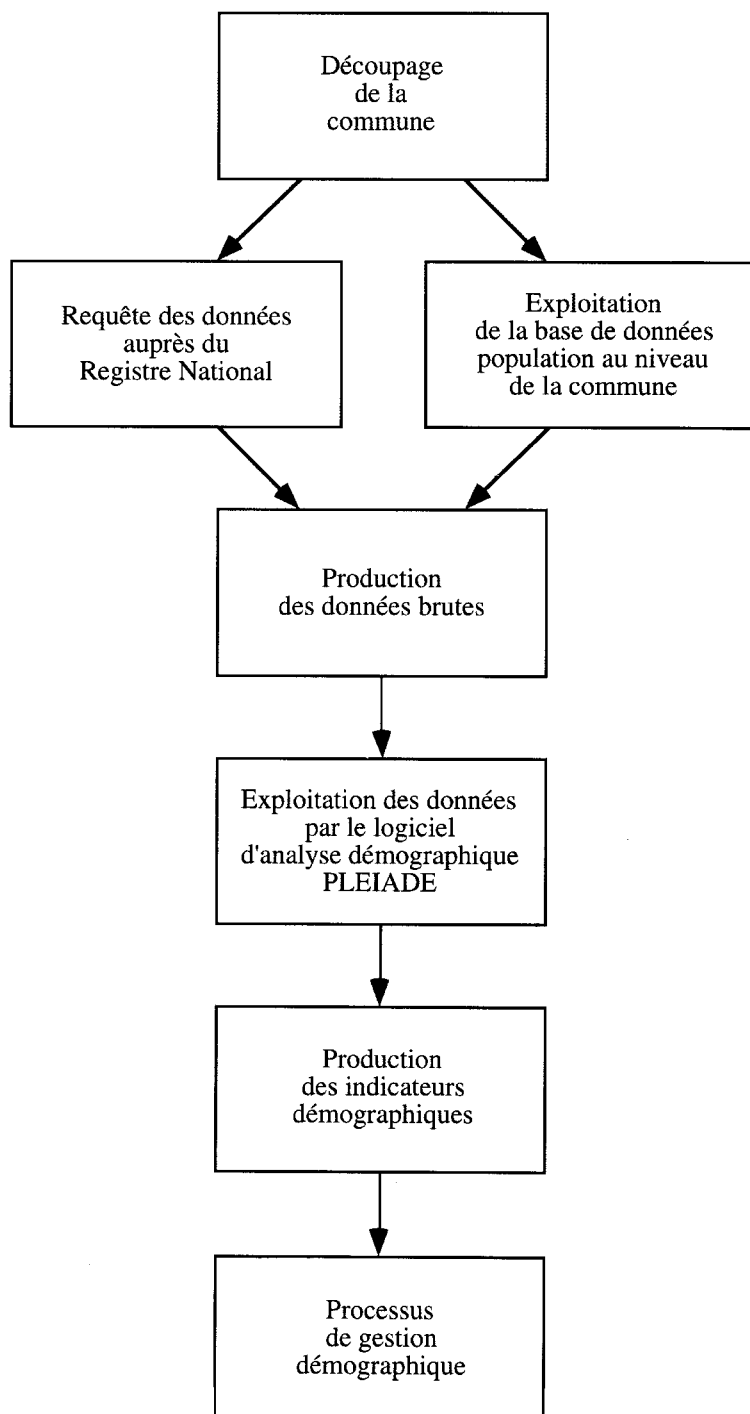
Pour dresser le portrait démographique complet de la commune et de ses quartiers, nous travaillons sur base des données fournies par les registres de population. La production des résultats suit le cheminement résumé par la figure 7.

3.1.1 Le découpage de la commune en quartiers

L'étape initiale consiste à subdiviser la commune en un certain nombre de parties du territoire que nous dénommons « quartiers ». La délimitation de ces quartiers est le fruit de concertations avec les responsables communaux - secrétaire communal, responsable de l'urbanisme, de l'aménagement du territoire, de la population, de l'aide sociale, de la police... - et, d'une manière générale, toutes les personnes ayant une bonne connaissance des spécificités démographiques, socio-économiques, culturelles, historiques et autres de leur commune. Ces quartiers, constitués par addition de rues, doivent néanmoins répondre à un certain nombre de critères de base :

1. Pour distinguer au mieux les différences démographiques entre les quartiers, il convient que ceux-ci constituent des univers sociaux les plus homogènes possibles. Généralement, ils se calqueront sur les centres de peuplement (anciennes communes, paroisses, hameaux, faubourgs...).

FIGURE 7. LES ÉTAPES SUIVIES DANS LA PRODUCTION DES DONNÉES DÉMOGRAPHIQUES COMMUNALES ET INFRA-COMMUNALES



2. Concentrer un chiffre minimum d'habitants afin que la plupart des indices démographiques calculés aient une signification statistique suffisante. Plus le quartier sera peuplé, plus cette exigence sera satisfaite. Le choix du seuil minimum exige une certaine souplesse. A priori, nous avons fixé un seuil minimum idéal à 1 000 habitants, mais il est évident que le calcul d'une table de mortalité ne peut se satisfaire de ce chiffre et exige le regroupement de quartiers. De même, la population de certains quartiers peut être largement inférieure à 1 000 habitants si le premier critère l'exige. Ici aussi, l'addition de plusieurs quartiers sera nécessaire pour le calcul des indicateurs démographiques les plus complexes.
3. Les quartiers ainsi définis, occupent une superficie largement supérieure à celle des secteurs statistiques, la plus petite unité spatiale pour laquelle le recensement de la population fournit des informations relatives aux caractéristiques des ménages et du logement, à l'activité professionnelle, à l'instruction... Afin de confronter ces informations aux données démographiques, et à la seule condition que les critères énoncés ci-dessus soient satisfaits, les quartiers doivent constituer une somme de secteurs statistiques. Certains écarts de délimitation sont acceptables pour autant que la population concernée soit nulle ou négligeable.
4. Lorsque les quartiers sont grossièrement délimités sur une carte, une vérification sur le terrain s'impose afin de préciser, au numéro de maison près, leurs limites. Cette opération s'impose chaque fois qu'une rue traverse plusieurs quartiers. Une trouée nette dans l'habitat, le passage d'un style d'habitation à un autre dans la même rue, un ruisseau, une voie de chemin de fer, une route... peuvent servir de frontière entre les quartiers concernés.
5. Au delà de ces critères, il faut enfin tenir compte des interactions entre les différents acteurs. En d'autres termes, les exigences scientifiques devront composer avec les « intérêts plus politiques » avancés par les uns et les autres.

3.1.2 La production des données brutes

L'étape suivante consiste à créer un « fichier tampon » dans lequel chaque rue ou portion de rue, identifiée sur base de la numérotation officielle, est affectée à un et un seul quartier. Ce fichier sera transmis au Registre National qui produira les données démographiques requises. Dans les communes qui disposent de leur propre infrastructure informatique, ce fichier servira de base au développement d'un logiciel de comptabilité démographique similaire à celui implanté au Registre National.

De quelles données disposons-nous ? Ce logiciel de comptabilité démographique produit, par quartier, par nationalité, par sexe et par groupe d'âges quinquennal, pour une période de une ou cinq années :

- l'état de la population en t_1 et en t_2 , le mouvement de la population (naissances, décès, entrées et sorties) et la répartition des naissances selon l'âge du père et de la mère entre ces deux dates ;
- le nombre d'émigrés et de décédés parmi la population présente en t_1 , le nombre d'émigrés et de décédés parmi la population immigrée entre t_1 et t_2 .

Dans le cas du logiciel de comptabilité démographique, plus complexe, installé en site propre dans les communes, nous pouvons ajouter à ces informations celles relatives à la mobilité intra-communale et aux lieux de destination et de provenance des migrants.

3.1.3 La production des résultats : le logiciel PLEIADE

Une fois encodées, les données brutes sont traitées par le logiciel PLEIADE. Il s'agit d'un logiciel d'analyse démographique adapté à l'étude des populations locales et opérationnel sous toute version d'EXCEL supérieure ou égale à 4.0, dans les environnements Macintosh et Windows. Dans sa version actuelle, PLEIADE permet

d'étudier le mouvement de la population entre t_1 et t_2 , la structure de la population par groupe d'âges et par sexe en t_1 et en t_2 , la fécondité (indice conjoncturel de fécondité et taux de fécondité selon l'âge de la mère) et la natalité, la mortalité (espérances de vie selon le sexe, quotients de mortalité par groupe d'âges quinquennal) et les migrations (répartition des entrées et des sorties par sexe et par groupe d'âges). De plus, il propose des projections de population à comportements démographiques constants et un tableau de synthèse récapitulant les principaux résultats démographiques. Il est à noter que les indicateurs de fécondité et de mortalité, les plus sensibles au problème des petits nombres, sont assortis des intervalles de confiance permettant d'estimer leur fiabilité statistique.

Ce logiciel PLÉIADE est installé sur micro-ordinateur au siège de l'administration communale. Des séances d'informations et un syllabus initient le personnel communal à la compréhension et à l'interprétation des résultats démographiques produits, ainsi qu'à la manipulation de PLÉIADE.

Au terme de cette première étape, nous sommes capables de dresser le portrait démographique complet - structures de population en t_1 et t_2 , mouvement de la population entre t_1 et t_2 , comportements démographiques entre t_1 et t_2 , - et récent de la commune et de chacun de ses quartiers. Parallèlement à cet objectif, se superpose un souci omniprésent de comparabilité :

- une comparabilité intercommunale afin de situer la commune parmi ses semblables, selon une série d'indicateurs démographiques ;
- une comparabilité infra-communale réalisée sur base du découpage en quartiers et illustrée notamment par une série de cartes.

Cette fonction comparative est, pour l'instant, indépendante du logiciel PLÉIADE. Dans sa seconde version, PLÉIADE proposera une banque de données que les utilisateurs pourront interroger afin de situer chaque indicateur démographique par rapport à celui d'autres communes, aux moyennes régionales et nationales. De même, pour les données infra-communales, une cartographie automatique et intégrée au logiciel sera proposée pour chaque indicateur.

3.2 La mise en évidence de situations anormales et la prévision

Après la radioscopie démographique de la commune et de ses quartiers, il s'agira de dresser un bilan d'ensemble et de sensibiliser les décideurs locaux à certaines anomalies. La cartographie des différents indicateurs démographiques permet déjà de mettre en évidence certaines inégalités de structure et de comportement. Citons, en guise d'exemples, une concentration excessive de telle ou telle nationalité dans certains quartiers handicapant les politiques d'intégration, une surreprésentation de personnes âgées dans d'autres quartiers posant des problèmes de sécurité ou incitant les responsables communaux à réajuster leur politique sociale en fonction de besoins mieux localisés, une surmortalité observée dans certaines zones de la commune...

Parallèlement à certains « problèmes », facilement repérables grâce aux outils statistiques et cartographiques, s'en cachent d'autres, indistincts, car occultés par d'autres réalités jugées positives, ou pas encore révélés. C'est dans ce cadre qu'intervient notamment l'outil prospectif que sont les projections de populations.

Pour illustrer notre propos, prenons le cas apparemment anodin de l'augmentation du chiffre de la population d'une commune. La croissance de la population est un argument fréquemment évoqué par les édiles locaux pour justifier la « bonne santé démographique » de leur commune ou plus globalement, l'efficacité de leur(s) politique(s). Certes, le chiffre de la population n'est pas une donnée « innocente » et intervient directement ou indirectement dans

la gestion des affaires communales¹. Or, comme nous l'avons observé précédemment, la croissance démographique d'une commune est largement conditionnée par l'ampleur du bilan migratoire. Si une commune est attractive, c'est qu'elle dispose d'atouts - disponibilité du parc immobilier, qualité de l'environnement, diversité et importance de l'emploi offert... - et l'appréciation de ceux-ci est variable selon les individus et notamment selon l'âge. Ainsi, un étroit parallèle a pu être observé entre la qualité de l'environnement des communes de la vallée de la Semois, dans le sud de la Belgique, et le bilan migratoire des personnes âgées² (Poulain, 1988-2). Cette attraction de personnes âgées se traduit par une augmentation du chiffre de la population ou compense les pertes démographiques consécutives à l'exode des jeunes, mais modifie radicalement la structure par âge de la population. Ceci n'est pas sans poser à terme des problèmes spécifiques aux collectivités locales. Il conviendra probablement d'entreprendre des aménagements en terme d'équipement et plus globalement de politique sociale en faveur des personnes âgées, de développer les réseaux de solidarité et d'entraide en investissant dans des domaines presque vierges d'emplois...

Les démographes doivent attirer l'attention des gestionnaires locaux sur les conséquences à moyen et long termes de décisions prises en fonction des avantages qu'elles peuvent procurer à court terme. C'est dans ce cadre qu'intervient aussi ce que nous appelons la « problématique des lotissements » (Poulain, 1988-2 ; Poulain, Eggerickx, 1993). De nombreuses communes belges ont connu récemment une expansion démographique rapide, grâce notamment à la création de lotissements. D'une part, les caractéristiques du lotissement et de ses logements déterminent dans une certaine mesure les caractéristiques socio-démographiques de la population qu'il accueille. D'autre part, un lotissement, c'est bien souvent la construction quasi simultanée d'une série d'habitations et donc l'arrivée en masse, dans un laps de temps relativement court, d'une population relativement semblable. Au sein du lotissement, c'est donc avant tout l'homogénéité qui prime : « [...] l'âge des adultes, le nombre et l'âge des enfants, le type de ménage varient bien souvent dans des limites fort étroites, au même titre d'ailleurs que les intérêts socio-culturels et l'appartenance à une classe socio-professionnelle ou idéologique » (Poulain, 1988-2).

Le vieillissement du lotissement est presque concomitant de celui de sa population. Après une période de forte croissance, liée à l'installation des ménages, le lotissement atteindra un plafond démographique déterminé par l'offre de logements. A ce moment, compte tenu d'une certaine stabilisation sociale de ses habitants et donc d'une mobilité en déclin, la population du lotissement vieillira en bloc, les générations les plus âgées n'étant pas remplacées par des générations plus jeunes. A titre d'exemple, nous avons pu analyser l'évolution des caractéristiques démographiques de 4 lotissements de l'agglomération namuroise, de 1980 à 1995. Le profil socio-économique de leur population est différent : Amée est une cité unifamiliale de standing moyen, Orjo est un immeuble de standing moyen à appartements multiples, Plomcot est un lotissement composé de logement sociaux abritant une forte proportion de population de nationalité étrangère et Suarlée est un nouveau lotissement résidentiel accueillant une population à revenu aisé. Pour chacun de ces lotissements, nous avons calculé l'évolution de l'âge moyen de leur population. Sachant que, de 1981 à 1991, l'âge moyen de la population de la Belgique a progressé d'un peu plus d'une année, on observe donc, dans chaque cas, un vieillissement nettement plus rapide (tableau 1).

¹ Il existe une relation, non directement proportionnelle, entre le nombre d'habitants d'une commune et le nombre de représentants communaux. Certains postes communaux ne sont ouverts que si le chiffre de la population dépasse un certain seuil. Le traitement des mandataires communaux est également lié à ces seuils de population. Toute personne supplémentaire représente un apport certain dans la caisse communale et cette somme sera d'autant plus élevée qu'il y a construction d'une nouvelle maison et que le ménage concerné bénéficie de revenus importants ...

² Entre 1982 et 1987, le groupe des plus de 60 ans a augmenté de plus de 6% et la migration est responsable à 63% de cette croissance. En 1991, à Vresse-sur-Semois, la part relative des personnes de plus de 60 ans s'élevait à près de 30 % (pour 22 % en Wallonie) pour seulement 14% de jeunes de moins de 15 ans.

Compte tenu de cette caractéristique, des équipements - écoles, centres sportifs, plaines de jeux... - perçus comme une nécessité impérieuse au départ, risquent à terme d'être largement sous-exploités avec, à la clé, un gaspillage des deniers de la commune et des citoyens.

TABLEAU 1. L'ÉVOLUTION DE L'ÂGE MOYEN DE LA POPULATION DE 4 LOTISSEMENTS DE L'AGGLOMÉRATION NAMUROISE (1980-1995)

	Amée	Orjo	Plomcot	Suarlée
1980	38,3 ans	36,6 ans	26,1 ans	27,5 ans
1985	39,9 ans	44,7 ans	27,9 ans	24,7 ans
1990	42,7 ans	48,4 ans	28,1 ans	25,8 ans
1995	43,2 ans	51,4 ans	30,7 ans	28,1 ans

La mise en évidence de situations qui peuvent engendrer, à moyen ou long termes, des problèmes nécessite :

- La localisation dans l'espace communal de zones à risque et la possibilité de les isoler statistiquement. Cette démarche est opérationnelle grâce au Registre national par l'addition de rues et/ou de portions de rues. Les données seront alors traitées par PLÉIADE.
- Le calcul de projections démographiques. Dans sa version actuelle, PLÉIADE propose des projections de population à comportements démographiques constants. Ces projections sont réalisées sur base des structures observées en début de période et en appliquant systématiquement à ces effectifs les comportements démographiques observés pendant la période quinquennale $t-t+5$. Il s'agit d'un scénario de référence qui prétend répondre à la question suivante : quelles seront la population et la structure par âge de telle commune ou de tel quartier en l'an $t+10$, $t+15$ et $t+20$ si, jusqu'à ces dates, les comportements démographiques, tels qu'ils viennent d'être observés, se maintiennent ? Cela suppose également qu'aucune décision politique ou qu'aucun événement « extérieur » (par exemple, une brusque détérioration ou un essor de la conjoncture démographique) ne viendra influencer, positivement ou négativement, ces comportements démographiques. Dans les versions ultérieures du logiciel PLÉIADE, nous comptons introduire d'autres scénarii tenant compte notamment de la construction de lotissement de tel type et de telle ampleur.

3.3 La réponse à une problématique précise

Dans bien des cas, on fera appel au démographe afin qu'il apporte une réponse à une problématique spécifique. Dans ce cas, l'approche quantitative et descriptive traditionnelle devra généralement se doubler d'une approche plus qualitative. Le passage de la description d'une situation à son explication, ou de la mise en évidence d'hétérogénéités ou d'inégalités démographiques à leur justification, pourra s'opérer de deux façons :

- par le biais d'une enquête bien ciblée sur une problématique précise. Par exemple, le mouvement des sortants est révélateur de besoins éventuellement insatisfaits par la commune. Il convient de les identifier afin de prendre, dans des conditions optimales, les décisions adéquates. Ce travail fut réalisé dans le cadre d'une commune de l'agglomération bruxelloise, dont les autorités s'inquiétaient de l'exode des jeunes ménages. Après avoir identifié les quartiers où se posait ce problème d'exode urbain et, en regard de l'âge et de la nationalité, les populations « soumises au risque de migrer », nous avons mené une enquête sur les motivations et le contexte dans lequel s'inscrivaient ces mouvements migratoires. Cette enquête a été réalisée, par voie postale, auprès de 151 ménages entrés ou sortis de la commune

au cours des années 1991 et 1992 et tirés aléatoirement dans la liste établie à partir des registres des entrées et des sorties (Eggerickx, Poulain, 1994). Avec ces éléments quantitatifs et qualitatifs, les responsables communaux disposent désormais d'informations récentes et précises leur permettant d'ajuster au mieux leur politique future au problème posé.

- par le biais du tableau de bord statistique de la commune. En recourant aux sources de données externes (recensement de la population et des logements, statistique du chômage...) et internes (Registre de population, données du Service d'Aide Sociale, du cadastre...) à la commune, il s'agit de construire une banque de données interactives relatives aux différents quartiers et à leurs citoyens. Outre les informations démographiques décrites ci-dessus, des données relatives au niveau d'éducation, à l'activité économique, à la profession, aux finances communales, à l'aide sociale, à l'occupation du sol, à la santé, au logement, à la sécurité... seraient intégrées dans le même découpage en quartiers. L'analyse des données et des interactions pourra être réalisée par :
 - une cartographie thématique
 - un diagramme en nuage de points pour tout couple de variables, auquel serait associé la droite d'ajustement linéaire par moindres carrés et le coefficient de corrélation correspondant.

Conclusion

En Belgique, une démographie locale est techniquement possible grâce à l'extraordinaire banque de données que constitue le Registre national. C'est sur cette base que nous avons développé un programme de gestion démographique destiné aux responsables locaux. Notre objectif est, d'une part, de sensibiliser ces derniers aux réalités démographiques de leur commune et de ses quartiers, et d'autre part, de favoriser l'intégration de la démographie parmi la panoplie des outils d'aide à la décision.

Pour la description des caractéristiques démographique de la commune et de ses quartiers, la mise en évidence de situations anormales et les perspectives démographiques à comportement constant, nous avons recours au logiciel de comptabilité démographique PLÉIADE. Très prochainement, ce logiciel bénéficiera d'améliorations sensibles. Citons notamment :

- l'intégration d'une banque de données comparatives, relatives à d'autres communes, aux situations régionale et nationale ;
- une cartographie automatique, par plage, des différents indicateurs démographiques, à l'échelle des quartiers ;
- de différents scénarii de projections démographiques, tenant compte notamment de la « problématique des lotissements » ;
- l'introduction des ménages comme unité de dénombrement.

En fonction des besoins, cet outil démographique peut être suppléé par une enquête qualitative, ciblée sur une problématique précise, et par le tableau de bord statistique de la commune. Celui-ci, sur base d'un même découpage en quartiers, rassemble, outre les données démographiques, des informations relatives aux professions, à l'activité économique, à l'aide sociale, à l'instruction...

BIBLIOGRAPHIE

- BACCAINI, Brigitte, (1994), Comportements migratoires et cycles de vie, *Espace, Populations, Sociétés*, 1, pp. 61-74.
- DEBUISSON Marc, EGGERICKX Thierry et POULAIN Michel, (1993), Démographie des grandes agglomérations urbaines en Wallonie, *Courrier hebdomadaire du CRISP*, Bruxelles, n° 1422 - 1423, 57 p.
- EGGERICKX Thierry, (1990), Les soldes migratoires et leurs conséquences à l'échelon local, *Revue Belge de Géographie*, fascicule 3, pp. 133-151.
- EGGERICKX Thierry, POULAIN Michel, (1994), Jette. Profil démographique et problématique de la mobilité spatiale, *Migrations et Espace*, n°8, Académia, Louvain-la-Neuve, 69 p.
- INSTITUT NATIONAL DE STATISTIQUE, (1990), *Statistiques démographiques. Mouvement de la population en 1988*, Bruxelles, n°1.
- POULAIN Michel, (1988-1), *Population et gestion communale. Huy, Hannut, Waremmes*, Fondation Roi Baudouin, De Bœck, Bruxelles, 117 p.
- POULAIN Michel, (1988-2), *Population et développement rural. L'arrondissement de Dinant*. Fondation Roi Baudouin, De Bœck, Bruxelles, 163 p.
- POULAIN Michel, (1994), La mobilité interne en Europe. Quelles données statistiques ?, *Espace, Populations, Sociétés*, 1, pp. 13-30.
- POULAIN Michel, EGGERICKX Thierry, (1993), Les conséquences indirectes de prises de décisions politiques sur la répartition spatiale de la population, *Croissance démographique et urbanisation. Politiques de peuplement et aménagement du territoire, Séminaire international de Rabat (15-17 mai 1990)*, AIDELF, n°5, Paris, pp. 411-428.
- POULAIN Michel, EGGERICKX Thierry, BEGEOT François, PAUL Benoît et TERRA Abrami Valério (1995), Fiabilité des comparaisons internationales des données démographiques de base en Europe, *Les comportements démographiques en Europe : facteurs de différentiation régionale*, Colloque organisé par le GEVERU, Bruxelles, février 1993, (à paraître).
- PUMAIN Denise et COURGEAU Daniel, (1993), Mobilité par temps de crise, *Population et Sociétés*, n°279.
- VAN CAUWENBERGHE Jean-Claude (1988), Aménagement du territoire et population : le point de vue d'un municipaliste, *Démographie et gestion locale, Chaire Quetelet 1988*, Institut de Démographie, U.C.L., non publiée.

Influences des familles sur le comportement du consommateur : le cas espagnol

Carlo M. GALLUCI, Ivana CASABURI et Michele QUINTANO

Escuela Superior de Administracion y Direccion de Empresas, ESADE

Département de management et de marketing

1. Segmentation du marché en tant qu'instrument du marketing

Lorsque l'on définit le marché et les aires d'affaire spécifiques où l'on veut opérer, l'identification du consommateur potentiel est une partie vitale et stratégique du processus. Un seul regard sur les stratégies appliquées par les compagnies les plus prospères montre qu'elles ont analysé soigneusement le segment de marché qui est leur cible.

Philip Kotler¹ définit la segmentation du marché comme « la subdivision du marché en sous-groupes de consommateurs, où chaque sous-groupe peut être choisi en tant qu'objectif de marketing à atteindre avec une combinaison spéciale d'un marketing mélangé ».

Il est assez évident que la segmentation mène à la création des sous-groupes de consommateurs potentiels qui partagent les mêmes caractéristiques (âge, niveau de revenu, habitudes concernant l'achat). Si les individus qui composent chaque segment peuvent partager les mêmes caractéristiques, les personnes qui appartiennent à des segments différents ont des caractéristiques différentes.

L'intensification de la compétition des prix sur les marchés pour les produits standardisés a obligé les compagnies à chercher de nouvelles sources de revenu en différenciant leurs produits. L'importance du processus de segmentation du marché est, par conséquent, enraciné dans la conscience, car la seule modalité d'augmentation des bénéfices et des actions est la prolifération des marques orientées vers les différents segments dont les besoins ne sont pas satisfaits par les produits disponibles, destinés au marché de masse.

La segmentation du marché peut être considérée comme l'incarnation du processus de différenciation des produits. Elle est centrée sur le consommateur, en tenant compte des besoins individuels et admettant le fait que plusieurs consommateurs peuvent faire des demandes différentes sur le même marché. Les compagnies peuvent alors essayer d'adapter, le mieux possible, leur produit à ces besoins et demandes tant qu'elles maintiennent toujours une position commerciale correcte.

La différenciation permet aux entreprises de séparer leurs produits de ceux de la concurrence sans s'engager dans la guerre des prix.

Généralement, il y a six conditions qui rendent les segments de marché attractifs. Pour que la segmentation soit efficace, les six conditions doivent être remplies :

- La première condition est que toutes les composantes de chaque segment doivent être **homogènes**, et il y doit y avoir des différences entre les divers segments.
- Si le segment réagit positivement aux actions de marketing des compagnies, cela veut dire qu'il a été **correctement** identifié.
- Une autre condition est que le segment soit suffisamment large afin d'être sûr qu'il existe un bon potentiel de ventes actuelles et futures.

¹ P. Kotler (1981).

- Les politiques de segmentation fonctionnent si vos politiques de communication et de marketing atteignent avec succès le segment qui vous intéresse. Les segments de marché qui ne sont pas **touchés** doivent être soigneusement étudiés pour ne pas gaspiller votre temps et votre argent.
- Votre segment doit aussi être **profitable et durable**.
- De plus, la segmentation de marché mettra votre entreprise dans une position avantageuse sur le plan de l'individualisation des opportunités du marché. Vous serez vraiment capable de découvrir combien vos consommateurs sont satisfaits et si la satisfaction est faible, vous pourrez profiter des opportunités disponibles du marché.

2. Les variables impliquées dans le processus de segmentation du marché

Sur le marché des biens de consommation, les consommateurs sont répartis en groupe de la manière suivante² :

- segmentation par avantages du produit
- segmentation par catégories socio-démographiques
- segmentation par catégories socio-culturelles
- segmentation par comportement

La segmentation par avantages du produit est la méthode la plus proche de l'idée fondamentale de la segmentation. Elle se réfère aux préférences du consommateur vis-à-vis des qualités que l'on exige d'un produit. Ce type de segmentation vous permet de découvrir comment le consommateur potentiel répond aux caractéristiques de votre produit particulier.

L'utilisation de ce modèle vous permet de tracer le plan des perceptions du consommateur en définissant les critères qui déterminent les décisions d'achat, en positionnant les marques concurrentes en fonction de ces critères, en identifiant les niches libres et en évaluant l'attraction des positions libres dans chacun des différents groupes de consommateurs.

La segmentation par catégories socio-démographiques est la méthode la plus utilisée ; elle traite de l'étroite relation entre le comportement du consommateur et ses caractéristiques socio-démographique. On prend en compte les variables géographiques, démographiques (âge, sexe, cycle de la vie de famille etc.) et économiques. Dans cet article, on va analyser ces variables, en insistant particulièrement sur l'analyse des changements dans le cycle de vie familial qui peuvent affecter la stratégie corporatiste de management.

L'idée que les personnes qui appartiennent à des milieux socio-économiques différents ont des habitudes de consommation similaires et inversement, les personnes appartenant à un milieu similaire peuvent avoir des habitudes différentes, représente le fondement de la **segmentation socio-culturelle**. L'attitude du consommateur est déterminée par le «style de vie», c'est à dire par l'ensemble individuel des valeurs, attitudes, intérêts et opinions.

La **segmentation par comportement** sépare les consommateurs selon leurs attitudes envers l'achat. Les indicateurs les plus utilisés sont la fréquence d'utilisation du produit par le consommateur, la fidélité du consommateur et sa sensibilité aux instruments du marketing.

La validité de chacune de ces méthodes dépend des diverses conditions existant sur le marché au moment où l'analyse est réalisée. En tout cas, la segmentation démographique est la plus populaire, parce qu'elle permet assez facilement de recueillir et d'utiliser des données. Bien que la segmentation par comportement et par milieu socio-culturel soient plus efficaces, elle est plus difficile à mesurer et à appliquer.

² J.J. Lambin, (1993).

3. L'importance des variables démographiques

Comme nous l'avons déjà mentionné, les variables les plus utilisées dans la segmentation du marché sont les variables démographiques. Par démographie on sous-entend les études empiriques, statistiques et mathématiques sur la population humaine.

La facilité de la collecte des données et la simplicité du processus d'application de ces données sont les raisons principales de la grande popularité de ce type de segmentation³.

La répartition de la population par groupes d'âge et l'examen des changements qui peuvent se produire dans ces groupes nous permettent de comprendre les schémas hypothétiques d'achat des consommateurs potentiels.

Les différents groupes d'âge peuvent être examinés en fonction de schémas d'achat fondés sur le niveau du revenu (qui augmente avec l'âge), des besoins et désirs, des changements naturels dans la famille nucléaire...

Le cycle de vie familial est utilisé en tant que référence dans la segmentation, car il combine différents éléments, par exemple, l'âge et le statut social, avec des informations sur le nombre des membres de la famille qui vivent dans la maison à un moment donné.

La corrélation entre une certaine étape du cycle de vie de la famille et la possession de différents biens (logement, nourriture, soins de santé) est un indicateur plus sérieux et efficace des habitudes d'achat que le seul âge.

Le comportement individuel dans le cadre du cycle de vie familial change constamment et les variations les plus significatives sont observées lorsqu'il s'agit du passage d'une étape du cycle à une autre.

L'utilisation des variables démographiques a toujours été critiquée, parce qu'elle ne reflète ni le processus interne aux différentes classes sociales à mesure qu'elles se rapprochent d'une moyenne standard, ni les similarités croissantes des revenus, qui sont deux caractéristiques clé de notre société contemporaine. Il peut arriver, par exemple, que deux personnes ayant le même âge, le même niveau du revenu et se trouvant dans la même étape de leur cycle de vie familial, aient des préférences et des attitudes très différentes.

Les variables démographiques sont, par conséquent, peu opérationnelles lorsqu'il s'agit de l'identification des perceptions des gens concernant les marques qui rivalisent sur le marché. Pourtant, ces variables révèlent d'autres indicateurs, plus fonctionnels, des dépenses du consommateur, par exemple la localisation géographique, la dimension de la famille, le niveau du revenu etc.

Malgré ces critiques, les variables démographiques sont généralement utilisées par les compagnies qui essayent de segmenter le marché. Ces variables fournissent des informations utiles pour la définition et la sélection du support de publicité le plus efficace. En fait, la plupart des estimations de « l'audience » utilisent les critères démographiques pour segmenter le public-cible.

4. La famille en tant que concept du marketing

L'analyse des décisions d'achat est l'un des objets principaux de l'étude de marché. Le but est de recueillir des informations qui nous permettront de concevoir une stratégie de marketing efficace qui influencera le processus de prise de décision des consommateurs.

La décision individuelle d'achat est influencée par le groupe social auquel l'individu appartient et parfois même c'est le groupe qui prend la décision. C'est fréquemment le cas des

³ E. Valdani, (1986).

familles et, pour cette raison, elle est considérée comme une unité séparée lorsqu'on analyse le processus de prise de décision.

Le but de l'analyse du consommateur « famille » est de découvrir comment ses membres réagissent et tombent d'accord sur les produits et services achetés à l'usage de la famille.

Le concept de famille tel qu'il est utilisé en marketing et dans les études démographiques des différents pays a des significations différentes⁴. Par conséquent, il est indispensable de clarifier précisément ce que l'on entend par « famille ».

La famille est un concept dynamique⁵, les changements sociaux, économiques, politiques et culturels des vingt dernières années ayant engendré l'apparition de nouveaux types de famille.

Le post-modernisme a causé le déclin des valeurs familiales traditionnelles et a engendré l'apparition de nouvelles attentes et demandes qui ne forment pas encore un ensemble de valeurs reconnues. Les nouveaux types de familles se trouvent, par conséquent, dans un processus de transition. Leurs liens ne sont plus aussi forts qu'avant, mais les familles elles-mêmes sont plus capables à s'adapter à un milieu en mutation.

Les statistiques officielles des divers pays fournissent actuellement des informations sur les structures familiales autres que la famille nucléaire traditionnelle, par exemple les familles monoparentales, les couples homosexuels, les couples hétérosexuels sans enfants etc. Pourtant, les informations disponibles ont tendance à agréger ces nouvelles unités familiales en catégories larges et sont très peu utilisées dans les recherches de marketing.

Sur le plan des études de marketing, la famille est une unité sociale qui prend des décisions de groupe sur les produits et les services achetés et utilisés⁶. La structure de la famille change à travers le temps et par conséquent le « cycle de vie familial » est un instrument important de la segmentation du marché.

5. Le cycle de vie de la famille

Avec le temps, les processus de prise de décision du consommateur changent. Cela est peut-être dû à l'âge (les jeunes n'ont pas les mêmes préférences que les personnes âgées), au montant du revenu disponible (qui généralement augmente jusqu'à la retraite, après quoi il décroît) ou à d'autres changements du statut social des consommateurs.

La littérature existante décrit les différents cycles de vie de la famille⁷ comme ayant généralement de six à neuf phases, mais la plupart des auteurs acceptent que les repères soient : la naissance et l'enfance, le célibat, le mariage, la naissance des enfants et le moment où les enfants quittent le foyer. Depuis quelque temps, le divorce et ses conséquences sont devenus aussi des jalons dans la vie d'un nombre de plus en plus grand de personnes.

Les chercheurs en marketing sont intéressés par la définition des différents stades du cycle de vie de la famille parce que les schémas des dépenses d'une famille changent, évidemment, en conséquence⁸.

Les neuf stades définis par W.L. Wilkie⁹ sont :

⁴ H.W. Berkman, C.C. Gilson, (1978).

⁵ R. Conde, (1983).

⁶ J.L. Leon, E. Olabarria, (1991).

⁷ Henry Assael, (1987).

⁸ H.W. Berkman, C.C. Gilson, op. cit., pp. 204-206

⁹ W.L. Wilkie, (1990), p. 482.

1. **Le stade du célibataire** : un jeune vivant seul. A ce moment-là, le revenu est relativement réduit et le consommateur vient juste de débiter sa carrière. Pourtant, les individus appartenant à ce groupe dépensent leurs revenus pour le divertissement et des produits à la dernière mode (voitures, voyages de loisirs, vêtements etc.). De plus, ils sont très réceptifs aux nouveaux produits.
2. **Les couples de nouveaux mariés**. Ces couples sont jeunes et n'ont pas encore d'enfants. Le mari et la femme ont de grandes chances de travailler tous les deux, par conséquent leur revenu total est plus élevé que celui d'un célibataire. Le schéma de leurs dépenses est similaire au schéma d'un célibataire, bien que les raisons pour lesquelles ils essaient de nouveaux produits sont plutôt de nature pratiques que liées à la mode. De plus, ils sont plutôt enclins à acheter des biens durables parce que leur situation familiale est plus stable et ils regardent plus vers l'avenir.
3. **Le « nid complet » I** : un jeune couple avec des enfants de moins de six ans. Dès l'arrivée du premier-né, de nouveaux produits viennent s'ajouter sur la liste d'achats : des meubles pour bébés, de la nourriture, des produits pharmaceutiques etc. Ce type de famille opère maintenant sur un marché qui est totalement nouveau pour elle. La publicité est particulièrement importante comme source d'informations qui guidera ces nouveaux consommateurs dans le choix des produits et des marques.
4. **Le « nid complet » II** : un jeune couple avec des enfants de plus de six ans. Les enfants ayant grandi, ils n'ont plus besoin de produits pour petits enfants, mais de produits pour enfants d'âge scolaire. Ces familles ont une expérience plus grande concernant les achats et elle sont moins influencées par la publicité. Elles commencent à acheter des paquets familiaux. Leurs enfants commencent à jouer un rôle actif dans la prise des décisions, en formulant des demandes et des préférences.
5. **Le « nid complet » III** : un couple vers la cinquantaine ayant des enfants qui vivent encore à la maison. A ce stade, la famille est assez solidement organisée. Généralement, ces familles dépensent une bonne partie de leurs revenus dans des biens durables (qui remplacent les produits obsolètes), quelques articles de luxe et pour l'éducation de leurs enfants.
6. **« Le nid vide » I** : des couples actifs (qui travaillent) et dont les enfants ont quitté le foyer. A ce moment-là, la situation financière et le style de vie de la famille sont assez confortables parce que la taille de la famille a diminué. Ce groupe n'est pas très influencé par l'apparition de nouveaux produits et dépense une grande partie de son revenu dans des articles de luxe et divertissement (voitures, voyages de vacances, etc.).
7. **« Le nid vide » II** : des couples à la retraite et dont les enfants ont quitté le foyer familial. Le revenu de la famille décroît alors brusquement. Le facteur « coût » devient maintenant extrêmement important et les décisions d'achat deviennent plus rationnelles. Les dépenses médicales et pharmaceutiques augmentent et si le couple n'est pas le propriétaire d'une maison, il doit déménager dans un appartement plus petit et confortable. A ce moment de leurs vies, les consommateurs sont plus réceptifs qu'avant aux messages des médias.
8. **Le solitaire encore actif** : une personne qui vit seule mais qui travaille encore. Généralement, cette personne gagne suffisamment pour avoir un style de vie confortable. Ses habitudes concernant l'achat sont similaires à celles des couples actifs.
9. **Le solitaire retraité** : une personne à la retraite qui vit seule. Comme chez les couples à la retraite, le revenu de cette personne a baissé brusquement et les décisions d'achat sont rationnelles et prudentes.

La structure de la famille à travers les différents stades de son cycle de vie est extrêmement importante sur le plan du marketing parce qu'elle influence directement la conception et le conditionnement du produit (la stratégie du produit), le choix de l'organe de

publicité, le message publicitaire, la qualité et le prix du produit, le choix des filières de distribution et beaucoup d’autres problèmes stratégiques.

6. Le processus de prise de décision dans la famille

Beaucoup de produits sont utilisés par plusieurs personnes dans la famille et toutes ont une influence sur la décision d’achat. Cela rend la prise de décision plus difficile que lorsqu’elle est individuelle parce que les divers membres ont des attitudes différentes et jouent des rôles différents dans la décision d’achat.

Les principaux rôles dans la prise de décision dans une famille sont :

Le veilleur : membre de la famille qui identifie le premier le besoin d’un certain produit ou service ;

L’informateur : personne qui recueille des informations sur les produits qui pourraient satisfaire les besoins de la famille ;

L’expert : personne qui étudie les informations sur les différentes marques, prix et points de vente ;

Le leader d’opinion : membre de la famille dont les opinions ont la plus grande influence, directe ou indirecte, sur la décision finale ;

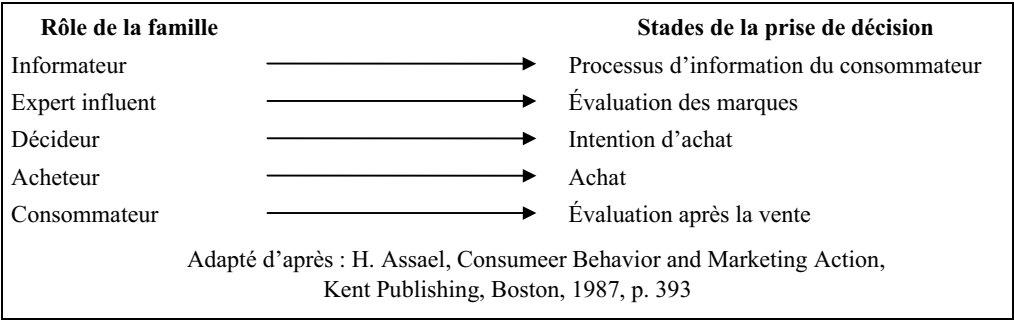
Le décideur : personne qui prend la décision d’achat. Cette personne peut être le veilleur ou l’expert et peuvent y être également associés d’autres membres de la famille ;

L’acheteur : personne qui met en pratique la décision et qui, généralement, achète exactement la marque et le type de produit qui fait l’objet de la prise de décision antérieure.

Le consommateur : la(les) personne(s) qui utilise(nt) le produit.

La Figure 1 montre les stades caractéristiques de prise de décision et révèle la distribution des rôles des différents membres de la famille pendant le processus de la prise de décision¹⁰.

FIGURE 1 : RÔLES DANS LES DIFFÉRENTS STADES DE LA PRISE DE DÉCISION



Les rôles du leader d’opinion et de l’expert dans le processus de la prise de décision sont les plus significatifs sur le plan des études de marché. Ce sont les personnes qui établissent les critères (coût, caractéristiques, style etc.) pour décider quel est le produit qui satisfait le mieux aux besoins de la famille.

Bien que les différents membres de la famille jouent des rôles différents dans le processus de la prise de décision, ces rôles peuvent avoir une importance plus ou moins grande.

¹⁰ Henry Assael, (1987).

Pourtant ces rôles ne sont pas répartis de la même manière dans les différents types de familles ou dans le même type de famille.

En fait, la distribution des rôles dans le processus de prise de décision reflète indiscutablement la structure du pouvoir de la famille. Cette structure a changé à travers le temps et est un facteur influent dans l'apparition de nouveaux types de familles.

Généralement, il y a deux types fondamentaux de rôles :

- le rôle instrumental qui définit les objectifs de la famille ;
- le rôle socio-émotionnel qui vise à maintenir la cohésion de la famille.

Dans les familles traditionnelles le rôle instrumental était joué par l'homme, qui prenait les décisions d'achat concernant certains produits, par exemple les automobiles et, d'une façon générale, les produits mécaniques. Le rôle socio-émotionnel était joué traditionnellement par la femme, dont les achats étaient basés surtout sur les qualités esthétiques (ou nutritionnelles dans le cas des produits alimentaires).

Durant les trente dernières années ces rôles ont été affectés par les changements sociaux qui ont altéré le concept même de famille. Le rôle instrumental est de plus en plus réparti entre les deux membres d'un couple et dans le cas de certains produits, les enfants ont aussi un rôle instrumental. Pendant ce temps, le rôle socio-émotionnel semble toujours être le territoire de la femme, peut-être parce qu'il implique la création des relations satisfaisantes dans le cadre de l'unité familiale et il a peu ou pas à voir avec la distribution du pouvoir.

Selon la manière dont les rôles sont répartis entre les membres de la famille, il y a quatre situations de dominance associées aux différents type de produits. Ici la dominance est jugée par rapport à la personne qui a la plus grande influence sur les autres membres de la famille, sur le plan de la prise de décision. Identifier quel est le rôle le plus important est essentiel pour les stratégies de marketing, puisque l'on identifie ainsi la personne qui sera la cible du message publicitaire.

La Figure 2 montre les quatre situations de dominance décrites ci-dessus¹¹ :

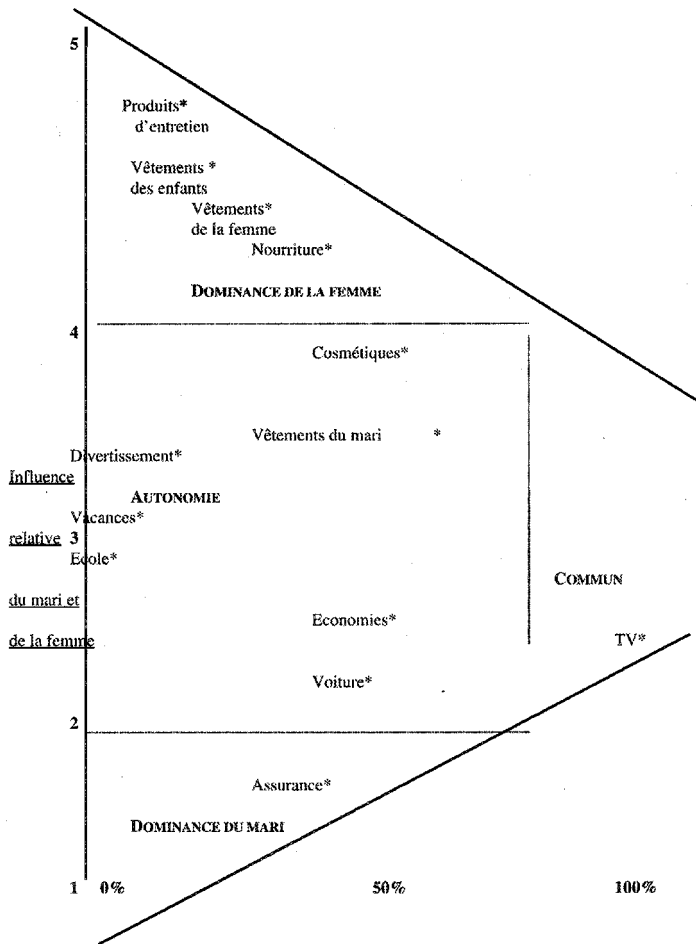
1. La femme domine la décision : c'est la cas pour des produits comme les vêtements, la nourriture et les produits ménagers en général¹² et ceci reflète le rôle socio-émotionnel traditionnel.
2. L'homme domine la décision : cette situation est de moins en moins fréquente puisque la répartition du pouvoir entre les membres d'un couple est plus également répartie. Beaucoup de décisions instrumentales sont maintenant partagées, ce qui engendre les situations suivantes :
3. Les décisions autonomes : la décision d'achat est répartie entre les deux membres du couple, en fonction de la personne qui est intéressée par le type de produit en cause ou qui est la plus compétente pour prendre la décision d'achat. C'est de plus en plus la situation du rôle instrumental dans le processus de décision de la famille.
4. Les décisions synchroniques : il s'agit des décisions prises en commun par le couple. Dans certains cas, les enfants aussi ont un rôle actif dans le processus de décision. Les décisions synchroniques sont fréquentes lorsqu'il s'agit de produits tels que les vacances, l'éducation, le divertissement etc.¹³

¹¹ H.W. Berkman, C.C. Gilson, op. cit. pp. 198 - 202

¹² Harry Davis and Benny Rigaux, (1974).

¹³ Jagdish Sheth, (1974).

FIGURE 2 : SPÉCIALISATION DES RÔLES DANS LA PRISE DE DÉCISION DANS LA FAMILLE
D'APRÈS H.L. DAVIS AND B.P. RIGAUX, « PERCEPTION OF MARITAL ROLES IN DECISION PROCESSES »
JOURNAL OF CONSUMER RESEARCH, VOL. 1 (JUN 1974)



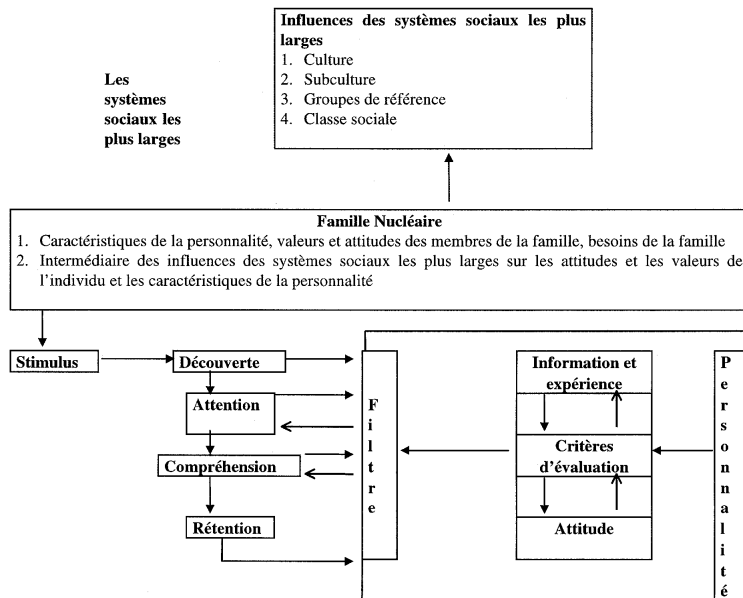
Les décisions synchroniques ou communes sont des décisions typiques de certains groupes et leurs caractéristiques sont les suivantes :

- plus l'achat comporte des risques et plus la probabilité que la décision soit commune est forte : par exemple, la décision d'acheter une maison est presque toujours commune, la responsabilité est alors répartie équitablement entre les époux.
- la prise des décisions communes est très probable lorsque les deux membres du couple ont des revenus, une éducation et une formation professionnelle similaires, étant donné qu'il s'agit de facteurs qui déterminent la distribution du pouvoir dans la famille.

Lorsqu'il s'agit des décisions communes, les spécialistes en marketing doivent adresser leur message publicitaire à la famille en entier, si l'achat implique un grand ou haut risque. Au contraire, quand l'implication de la famille est réduite et que la décision d'achat sera probablement individuelle, la publicité doit viser la personne qui prend effectivement la décision (leader d'opinion/expert).

Il est très important de souligner que la famille en tant que groupe social représente une partie du macro-groupe qui est la société dans laquelle les membres de la famille vivent et apprennent¹⁴. La Figure 3 montre comment les membres d'une famille sont influencés par les caractéristiques du système social, des groupes culturels et de référence et par la classe sociale à laquelle ils appartiennent ou veulent appartenir.

FIGURE 3 : INFLUENCES DE LA FAMILLE SUR LES CARACTÉRISTIQUES, LES ATTITUDES ET LES CRITÈRES D'ÉVALUATION DE LA PERSONNALITÉ DE L'INDIVIDU



Source : J.F. ENGEL, D.T. Kallant, R.D. Blackwell, « Consumer Behaviour », Driden Press, Illinois, 1973

De plus, cette influence peut être considérée comme un facteur de l'accumulation des connaissances, les habitudes de consommation de la famille étant le résultat des habitudes des anciennes générations. Autrement dit, les enfants observent comment agissent les adultes lorsqu'ils prennent une décision concernant l'achat. Cette attitude influence le comportement des enfants en tant que futurs consommateurs, dirigeant leurs décisions vers certaines marques préférées, modelant leur style d'achat et même les attitudes formées pendant leur enfance, dans le cadre de la famille.

On devrait avoir ces observations en tête lorsque l'on réalise des études de marché puisqu'elles offrent des renseignements sur la manière dont les préférences se constituent et sont importantes sur le plan de la définition du message publicitaire. C'est particulièrement valable pour les produits ménagers de nettoyage qui ont une certaine tradition. Il est important de rappeler aux utilisateurs potentiels que leurs mères utilisaient le même produit.

¹⁴ J.F. Engel, D.T. Kallant, R.D. Blackwell, op. cit., p. 192

7. Les typologies des familles et les habitudes concernant l'achat : le cas espagnol

7.1. Population et famille

L'enquête sur le Budget de la Famille Espagnole, réalisée en 1990 - 1991, définit la famille de la même façon que l'enquête réalisée dix ans plus tôt : «la personne ou les personnes qui occupent une part ou la totalité de la résidence de la famille et dont la nourriture et les autres biens sont payés par un seul budget».

Pourtant, les résultats de l'enquête reflètent une définition plus large de la « famille nucléaire » que celle appliquée à l'occasion du recensement national de 1970, qui comptait seulement les couples mariés ou une personne avec ou sans enfant(s) célibataire(s). Maintenant, la définition de la famille inclut des situations où les gens vivent ensemble pour d'autres raisons que le mariage ou la dépendance vis-à-vis de leurs parents. Ces cas sont le résultat des changements sociaux, économiques, politiques et culturels qui affectent aujourd'hui la société espagnole, de même que les autres pays développés.

Les données du recensement de 1970 et des enquêtes sur le budget des familles de 1980-81 et 1990-91 montrent l'évolution démographique suivante :

TABLEAU 1 : POPULATION ET FAMILLE EN 1970, 1981 ET 1991

	1970	1981	Δ%	1991	Δ%
Total de la population	34 040 641	37 682 355	+10,7	38 494 006	+2,1
Nombre des familles	8 853 657	10 111 514	+14,2	11 298 509	+11,7
Taille moyenne des familles	3,8	3,7		3,4	

Le Tableau 1 révèle une croissance plus importante du nombre des familles (+14,2%) que de l'effectif total de la population sur la période 1970-1981. Cette tendance est plus évidente encore sur la période 1981-1991 (population + 2,1%, nombre des familles +11,7%). La taille de la famille a diminué progressivement au fil des années. Même si, en 1950 et 1981 la famille moyenne comptait seulement 3,7 personnes, taille la plus réduite depuis le début du siècle (voir Tableau 2), certains spécialistes¹⁵ considèrent qu'il n'y a pas eu d'importants changements du nombre de personnes vivant sous le même toit en tant que famille.

En 1991, la famille moyenne comptait 3,4 personnes, ce qui confirme la tendance observée il y a dix ans.

TABLEAU 2 : TAILLE MOYENNE DE LA FAMILLE

Année du recensement	Nombre des personnes
1900	3,87
1910	3,98
1920	4,08
1930	4,09
1940	4,22
1950	3,74
1960	4,00
1970	3,81
1981	3,70
1991	3,41

¹⁵ M. Beltran, et al (1985).

La population espagnole a augmenté de seulement 2,1% durant la période 1981-1991, tandis que le nombre des familles a augmenté de 11.7%. Cette situation est due, au moins partiellement, au développement des familles « non-nucléaires », évolution qui peut très bien expliquer la diminution de la taille des familles (3,4 en 1991).

On ne peut pas confirmer cette hypothèse en analysant la variation du nombre des différents types de familles, puisque ces données ne sont pas disponibles pour 1981. Les données de 1991 donneront néanmoins une idée sur la manière dont la structure de la famille a changé.

7.2. Types de familles

Bien que l'Institut Espagnol de la Statistique (INE) utilise le concept « famille » plutôt que « famille nucléaire » dans l'enquête sur le budget familial de 1981, ni le nombre des types, ni les concepts qui se trouvent derrière, ne concordent avec ceux de la version de 1991 de la même étude.

Pour commencer, le nombre des types de famille a doublé (7 en 1981, 14 en 1991), permettant à l'analyse statistique d'être plus fidèlement adaptées à l'évolution actuelle de la structure familiale. Le Tableau 3 montre la tendance à définir avec plus de précision les types de famille, en faisant la distinction entre mariage et autres formes de vie en commun des adultes. Cette situation est due au fait que le nombre des familles non-traditionnelles progresse constamment puisque le nombre d'adultes non en union, avec ou sans enfants, et le nombre de couples non-mariés, avec ou sans enfants, augmentent.

TABLEAU 3 - TYPES DE FAMILLE EN ESPAGNES (1991)

Types de familles		Nombre
101	Adulte, 65 ou plus, sans enfants	720 058
102	Adulte, moins de 65, sans enfants	408 930
110	Adulte avec un ou plusieurs enfants	90 981
201	Couple, 65 ou plus, ayant un revenu et sans enfants	1 072 311
202	Couple, moins de 65, ayant un revenu et sans enfants	846 302
210	Couple avec un enfant	917 207
220	Couple avec deux enfants	1 340 179
230	Couple avec trois ou plusieurs enfants	549 637
291	Autres familles avec deux adultes, sans enfants	553 291
292	Autres familles avec deux adultes et enfants	86 664
301	Trois adultes sans enfants	1 348 601
302	Trois adultes avec enfants	1 044 005
401	Quatre ou plusieurs adultes sans enfants	1 320 131
402	Quatre ou plusieurs adultes avec enfants	1 000 204
Total		11 298 501
<i>Élaboré à partir des données de « Encuestas de Presupuestos Familiares 1990 – 1991 », INE, 1991</i>		

Cette tendance est confirmée par une étude réalisée par Institut d'Estadística de Catalunya (IEC) qui compare les données sur la structure des familles catalanes, données issues des recensements de 1981 et 1991 et qui montre l'augmentation du nombre de foyers d'une personne seule (+3,17%), des foyers composés d'une mère et d'enfants (+1,55%), d'un père et d'enfants (+0,23%), des couples sans enfants (+0.8%) et la diminution du nombre des familles

traditionnelles, par exemple les couples ayant des enfants (-4,9%) et des familles à plusieurs générations (-0,79%)¹⁶.

Les familles composées d'une seule personne ont enregistré la plus importante augmentation. En 1981, il y en avait 174 213, tandis qu'en 1991, il y en avait 262 719 soit 13,6% de la totalité des familles catalanes. Généralement, cette situation est interprétée comme le reflet d'une société plus individualiste où l'importance de la famille traditionnelle est en déclin. Beaucoup de personnes âgées (notamment veufs ou veuves) vivent seules, il en est de même des jeunes célibataires et des séparés ou divorcés. Bien des fois, il s'agit d'une situation temporaire, mais l'augmentation des personnes vivant seules est pourtant notable.

La réduction de la dimension de la famille moyenne espagnole a conduit à l'élimination de la catégorie « adulte avec quatre ou plus enfants mineurs » qui apparaissait dans l'enquête de 1981 et qui est maintenant incluse dans la catégorie « couple avec trois ou plusieurs enfants ».

L'étude de 1991 utilise le terme « enfants » plutôt que « mineurs ». Cela montre la tendance des jeunes d'aujourd'hui à vivre plus longtemps à la maison. Parmi les différentes raisons, citons la tendance à se marier plus tard.

7.3. Les schémas d'achat

Dans la recherche qui vise l'identification des segments du marché, il est intéressant d'analyser les relations entre les différents types de familles et leurs préférences de consommation.

Les spécialistes reconnaissent que le comportement de consommateur peut varier dans les différents types de familles, qui dépensent des sommes différentes de leurs revenus pour les divers produits et services à leur usage. Bien qu'il soit très important d'en savoir plus, l'INE ne fournit pas encore des informations quantitatives à ce sujet.

L'information incluse dans les pages qui suivent est strictement qualitative et n'est pas soutenue par les statistiques.

Les études INE sur les dépenses (Tableau 4) offrent des détails sur sept catégories de produits typiques ainsi que sur deux autres catégories restantes (Groupe 8 et 9).

L'étude de 1991 offre une image des dépenses totales de différents types de familles moyennes ainsi que la décomposition par catégorie de produit.

Nous allons analyser cette décomposition des dépenses et commenter la pertinence des dépenses dans chaque catégorie de produit dans les différents types de familles.

Pour réaliser cette analyse, on prend les chiffres (en pesetas) mentionnés par les différentes catégories de dépenses dans les familles moyennes de tous les types et on les a converti en pourcentages de la dépense moyenne.

L'information obtenue nous permet de voir comment les dépenses sont distribuées par type de famille et de comparer cette information avec les informations sur la famille moyenne et sur les différents types de familles.

Cette comparaison révèle les caractéristiques familiales qui conduisent à la distribution différente des dépenses par catégorie de produit. Pour rendre le Tableau 4 plus pertinent, on a ajouté des informations sur les dépenses moyennes par personne. On peut ainsi comparer les sommes dépensées pour diverses catégories de produits par les différents types de familles (Tableau 5). Par conséquent, on peut passer d'une analyse horizontale (le poids des différentes catégories de produits dans chaque type de famille) à une analyse verticale (dépenses par tête pour chaque catégorie de produit par chaque type de famille).

¹⁶ La Vanguardia, (1994).

TABLEAU 4 : RÉPARTITION DES DÉPENSES ANNUELLES SELON LE TYPE DE FAMILLE

Dépenses annuelles par famille	Dépense totale	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5	Groupe 6	Groupe 7	Groupe 8	Groupe 9
Dépense moyenne en pesetas - 1991	2527105	620566	239337	570573	138396	65086	309154	153922	325124	104947
Adulte, 65 ou plus, sans enfants	937961	236415	75614	375955	64687	32904	34979	16771	59917	40719
Adulte, moins de 65 ans, sans enfants	1482186	261457	129723	469753	87389	28263	147949	77847	216782	63022
Adulte avec un ou plusieurs enfants	1664927	445201	181706	423470	120596	48134	113367	134242	155686	42536
Couple, 65 ou plus, ayant un revenu et sans enfants	1482974	455546	122377	483881	83008	41304	76772	36502	113233	70351
Couple, moins de 65 ans, ayant un revenu et sans enfants	2173314	492851	203117	537315	136212	49253	293826	85175	285572	89993
Couple avec un enfant	2614599	568040	246955	584103	173142	82274	367593	170200	323090	99201
Couple avec deux enfants	2857878	672901	293900	602643	171874	77691	389742	207013	334319	107795
Couple avec trois ou plusieurs enfants	2726994	762334	283450	570426	152642	70266	313058	179862	294246	100709
Autres familles avec deux adultes, sans enfants	1762608	440400	151478	509332	97370	48937	161915	81352	213427	59407
Autres familles avec deux adultes et enfants	2210568	535930	211852	489681	147203	45704	252377	180347	272290	75182
Trois adultes sans enfants	2458013	621183	232110	560715	129878	66650	300557	128162	334103	111654
Trois adultes avec enfants	2964175	778311	301842	604881	151499	67080	358064	208618	373766	120113
Quatre ou plusieurs adultes sans enfants	3634721	825930	325251	728666	170434	86813	501508	264655	565278	166184
Quatre ou plusieurs adultes avec enfants	3551242	915476	341050	640774	168065	85986	481674	238626	529431	150159
En pourcentage	100	24,56	9,47	22,58	5,48	2,58	12,23	6,09	12,87	4,15
Adulte, 65 ou plus, sans enfants	100	25,21	8,06	40,08	6,9	3,51	3,73	1,79	6,39	4,34
Adulte, moins de 65 ans, sans enfants	100	17,64	8,75	31,69	5,9	1,91	9,98	5,25	14,63	4,25
Adulte avec un ou plusieurs enfants	100	26,74	10,91	25,43	7,24	2,89	6,81	8,06	9,35	2,55
Couple, 65 ans ou plus, ayant un revenu et sans enfants	100	30,72	8,25	32,63	5,6	2,79	5,18	2,46	7,64	4,74
Couple, moins de 65 ans, ayant un revenu et sans enfants	100	22,68	9,35	24,72	6,27	2,27	13,52	3,92	13,14	4,14
Couple avec un enfant	100	21,73	9,45	22,34	6,62	3,15	14,06	6,51	12,36	3,79
Couple avec deux enfants	100	23,55	10,28	21,09	6,01	2,72	13,64	7,24	11,7	3,77
Couple avec trois ou plusieurs enfants	100	27,96	10,39	20,92	5,6	2,58	11,48	6,6	10,79	3,69
Autres familles avec deux adultes, sans enfants	100	24,99	8,59	28,9	5,52	2,78	9,19	4,62	12,11	3,37
Autres familles avec deux adultes et enfants	100	24,24	9,58	22,15	6,66	2,07	11,42	8,16	12,32	3,4
Trois adultes sans enfants	100	25,27	9,44	22,81	5,28	2,71	12,23	5,21	13,59	4,54
Trois adultes avec enfants	100	26,26	10,18	20,41	5,11	2,26	12,08	7,04	12,61	4,05
Quatre ou plusieurs adultes sans enfants	100	22,72	8,95	20,05	4,69	2,39	13,8	7,28	15,55	4,57
Quatre ou plusieurs adultes avec enfants	100	25,78	9,6	18,04	4,73	2,42	13,56	6,72	14,91	4,23

GROUPE 1: Nourriture, boisson et tabac; GROUPE 2: Vêtements et chaussures; GROUPE 3: Logement, chauffage et électricité; GROUPE 4: Meubles, équipements et installations pour la maison et entretien; GROUPE 5: Dépenses médicales et de santé; GROUPE 6: Transport et communication; GROUPE 7: Récréation et divertissement; GROUPE 8: Autres biens et services; GROUPE 9: Autres dépenses.

Elaboré à partir des données de « Encuesta de Presupuestos Familiares, 1990 - 1991, INE, 1991

TABLEAU 5

Dépenses par personne - 1991 (en pesetas)	Nombre de personnes	Dépense moyenne par personne	Groupe 1	Comparé à la moyenne	Groupe 2	Comparé à la moyenne	Groupe 3	Comparé à la moyenne
Individus dans la famille moyenne (population/famille)	3,41	741087	181984	100	70187	100	167323	100
Adulte, 65 ou plus, sans enfants	1	937961	236415	130	75614	108	375955	225
Adulte, moins de 65, sans enfants	1	1482186	261457	144	129723	185	469753	281
Adulte avec un ou plusieurs enfants	2,83	588313	157315	86	64207	91	149636	89
Couple, 65 ou plus, ayant un revenu et sans enfants	2	741487	227773	125	61189	87	241941	145
Couple, moins de 65 ans, ayant un revenu et sans enfants	2	1086657	249426	135	101559	145	268658	161
Couple avec un enfant	3	871533	189347	104	82318	117	194701	116
Couple avec deux enfants	4	714470	168225	92	73475	105	150661	90
Couple avec trois ou plusieurs enfants	5,29	515500	144109	79	53582	76	107831	64
Autres familles avec deux adultes, sans enfants	2	881304	220200	121	75739	108	254666	152
Autres familles avec deux adultes et enfants	3,43	644480	156248	86	61764	88	142764	85
Trois adultes sans enfants	3	819338	207061	114	77370	110	186905	112
Trois adultes avec enfants	4,64	638831	167739	92	65052	93	130362	78
Quatre ou plusieurs adultes sans enfants	4,45	816791	185602	102	73090	104	163745	98
Quatre ou plusieurs adultes avec enfants	6,08	584086	150572	83	56094	80	105390	63

GROUPE 1: Nourriture , boisson et tabac; GROUPE 2: Vêtements et chaussures; GROUPE 3: Logement, chauffage et électricité;
GROUPE 6: Transport et communication; GROUPE 7: Récréation et divertissement; GROUPE 8: Autres biens et services;

Élaboré à partir des données de « Encuesta de Presupuestos Familiares, 1990 - 1991, INE, 1991

TABLEAU 5 (SUITE)

Groupe 4	Comparé à la moyenne	Groupe 5	Comparé à la moyenne	Groupe 6	Comparé à la moyenne	Groupe 7	Comparé à la moyenne	Groupe 8	Comparé à la moyenne	Groupe 9	Comparé à la moyenne
40585	100	19087	100	90661	100	45138	100	95344	100	30776	100
64687	159	32904	172	34979	39	16771	37	59917	63	40719	132
87389	215	28263	148	147979	163	77847	172	216782	227	63022	205
42613	105	17008	89	40059	44	47435	105	55013	58	15030	49
41504	102	20652	108	38386	42	18251	40	56617	59	35176	114
68106	168	24627	129	146913	162	42588	94	142786	150	44997	146
57714	142	27425	144	122531	135	56733	126	107697	113	33067	107
42969	106	19423	102	97436	107	51753	115	83580	88	26949	88
28855	71	13283	70	59179	65	34000	75	55623	58	19038	62
48685	120	24469	128	80958	89	40676	90	106714	112	29704	97
42916	106	13325	70	73579	81	52579	116	79385	83	21919	71
43293	107	22217	116	100186	111	42721	95	111368	117	37218	121
32651	80	14457	76	77169	85	44961	100	80553	84	25886	84
38300	94	19509	102	112698	124	59473	132	127029	133	37345	121
27642	68	14142	74	79223	87	39248	87	87077	91	24697	80

GROUPE 4: Meubles, équipements et installations pour la maison et entretien; GROUPE 5: Dépenses médicales et de santé;
GROUPE 9: Autres dépenses.

La variable la plus importante sur le plan de la distribution des dépenses semble être le nombre de personnes qui constitue un foyer. Cette variable est très importante dans les catégories de produits comme le logement (Groupe 3) et les dépenses connexes (Groupe 4) : lorsque la famille est plus restreinte, le taux de dépenses moyennes par personne est plus élevé. Si le logement représente entre 32% et 40% des dépenses totales d'une personne seule, les « couples avec deux enfants » n'y consacrent qu'environ 21% de leur budget global. Ceci est encore plus explicite en termes de dépenses par personne : les dépenses d'une personne seule s'élèvent à 281/100 tandis que les couples avec deux enfants dépensent 90/100 (Tableau 5).

La présence d'enfants dans le foyer influence aussi les schémas des dépenses. Les dépenses pour la santé (Groupe 5), l'éducation et la culture (Groupe 7) sont plus élevées dans les familles ayant des enfants, bien que les dépenses pour la santé soient aussi très fortes dans les familles « âgées » (65 ans et plus) où il n'y a pourtant pas d'enfants. Les dépenses du Groupe 7 sont très soutenues pour les personnes seules de moins de 65 ans. Cela est dû au fait que cette catégorie de produits inclut non seulement l'éducation et la culture, mais aussi « les loisirs », pour lesquels les jeunes célibataires dépensent une bonne partie de leur revenu.

Par conséquent, l'âge est une autre caractéristique importante qu'il ne faut pas oublier lorsque l'on étudie les schémas des dépenses familiales, pourtant, on ne lui accorde que peu d'attention dans les sources statistiques étudiées où seule est prise en compte la distinction entre les personnes de moins ou de plus de 65 ans.

8. Conclusions

Une analyse des concepts utilisés par les chercheurs en marketing dans les études sur le comportement du consommateur révèle peu de similarités avec les concepts utilisés dans les études démographiques.

Cela est dû aux approches utilisées par les deux disciplines : elles divergent car on analyse la famille selon des points de vue et pour des buts différents.

Les typologies des familles utilisées dans les études démographiques réalisées par INE en Espagne diffèrent de celles utilisées dans les études de marketing et leur façon d'aborder le cycle de vie de la famille révèle une divergence particulièrement forte. Par conséquent, les données démographiques ne sont pas une source adéquate d'informations pour les études efficaces sur les schémas des dépenses familiales.

Cela veut dire que lorsque les entreprises essayent d'établir des stratégies de segmentation du marché et ont besoin d'informations sérieuses sur la « consommation familiale », ils doivent recourir à des enquêtes spécifiques sur le terrain qui ne peuvent être aussi complètes qu'une analyse exhaustive.

Bien que, ces enquêtes sur le terrain soit le moyen le plus efficace d'étudier les processus de prise de décision dans la famille, nous croyons que les analyses démographiques peuvent fournir un point de départ utile. Elles fournissent des connaissances élémentaires sur les tendances passées et des prévisions sur les tendances futures concernant l'évolution du concept de « famille ». Ceci est important pour les entreprises et les aide à identifier les tendances qui, très probablement, affectent les dépenses en produits et services dont la famille a besoin.

Pourtant, les données démographiques ont un point faible : elles offrent une interprétation purement quantitative dans un domaine où l'information qualitative est plus importante. Dans les études de marché, les données ne peuvent pas fournir des informations sur certaines caractéristiques du consommateur telles que la motivation, la préférence pour une certaine marque, les goûts etc. Ce type d'analyse exige une approche plus sociologique et psychologique afin de cerner des aspects qui ne peuvent être saisis par des nombres. Malgré cet inconvénient, les démographes sont conscients que la relation entre les typologies familiales et

le processus de prise de décision est très importante. IEC a récemment commencé une étude sur cette relation, mais les résultats paraîtront ultérieurement.

Bien que, depuis les vingt dernières années, les études démographiques ont pris de plus en plus en considération les besoins collectifs, elles continuent à être en retard sur le plan des changements dans la structure familiale et commencent juste à saisir les nouveaux types de famille alors qu'elles sont déjà implantées dans la société. A l'inverse, le marketing est toujours attentif à l'apparition de nouveaux phénomènes avec lesquels les entreprises doivent composer. Mais cela est probablement dû aux différences inhérentes à la nature des ces deux champs.

Pour conclure, la relation entre démographie et marketing a certaines limites attribuables à la nature même de ces deux disciplines. La question qui se pose est de savoir s'il existe une possibilité de mettre leurs fonctions en liaison plus directe et si leurs buts différents justifient les problèmes liés à la transformation des études démographiques en instrument de marketing.

BIBLIOGRAPHIE

- ASSAEL Henry, 1987, *Consumer Behavior and Marketing Action*, Kent Publishing, Boston, p. 393.
- BELTRAN M., GARCIA FERRANDO M., LOPEZ PINTOR R., RODRIGUEZ CABRERO G., THIEBAUT C., TOHARIA J.J., 1987, *Estudios sobre la familia espagnola*, Ministerio de Trabajo y Seguridad Social, Madridi, p. 75.
- BERKMAN H.W., GILSON C.C., 1978, *Consumer Behavior : Concept and strategy*, Dickenson Publishing Company, Inc., New York, pp. 196-197.
- CONDE R., 1983, *Tendencias de cambio en la estructura familiar*, Revista Espagnola de Investigaciones Sociologicas, pp. 38-80.
- DAVIS Harry and RIGAUX Benny, June 1974, *Perception of Marital Roles in Decision Process*, Journal of Consumer Research, p. 54.
- KOTLER P., 1981, *Marketing Management*, ISEDI, Torino, pp. 155-159.
- LAMBIN J.J., 1993, *Strategic Marketing*, Mc GrawèHill, pp. 168-184.
- LA VANGUARDIA, Déc. 15, 1994, *Disminuye el numero de hogares habitados por parejas con hijos*.
- LÉON J.L., OLABARRIA E., 1991, *Conducta del consumidor y Marketing*, Ed. Deusto, Bilbao, pp. 212.
- SHETH Jagdish, 1974, *A Theory of Family Decisions*, in Jagdish Sheth (ed) Models of Buyer Behaviour, New York, Harper & Row, pp. 17 - 33.
- VALDANI E., 1986, *Marketing Strategico*, ETASLIBRI, Milano, pp. 209-220.
- WILKIE W.L., 1990, *Consumer Behavior*, John Wilkie & Sons Press, New York, p. 482
- ZALTMAN Gerald and WALLENDORF Melanie, 1983, *Consumer Behaviour, Basic Findings and Management Implications*, John Wiley & Sons, New York.

De l'utilité et des méthodes pour établir des perspectives démographiques relatives à de petites régions

Jacques MENTHONNEX

SCRIS, Lausanne

Introduction

Tant les décideurs politiques que certains milieux privés ont besoin de prévisions démographiques à un niveau régional fin, dans le cadre d'objectifs de planification. La difficulté alors rencontrée par le démographe « prévisionniste » est engendrée par le fait bien connu que, plus une région est petite, plus l'évolution de sa population est sensible à des événements particuliers, exceptionnels, touchant l'économie ou les infrastructures locales. Les caractéristiques des migrants, spécifiques à ces événements, peuvent même montrer des particularités surprenantes (âges, sexe, fécondité...). Dans ces conditions, l'élaboration de prévisions démographiques par la méthode des composantes ¹ peut s'avérer difficile à établir.

Une manière d'élucider ce problème, envisagée dans cette communication, serait d'utiliser une méthode empirique, simple et peu sensible aux fluctuations aléatoires, pour répartir dans les régions des résultats de prévisions obtenues au niveau suprarégional par la méthode classique des composantes. On justifie ce procédé par la simplicité de sa mise en oeuvre et par le fait que la perte d'informations locales devrait être compensée par un moindre risque en ce qui concerne l'influence des événements locaux récents et exceptionnels.

Utilisation des perspectives dans le contexte vaudois

Avant de décrire certains aspects méthodologiques propres aux perspectives de petites régions, une brève description du contexte qui a suscité cette communication précisera encore mieux la problématique.

Le fédéralisme suisse implique que les décisions se prennent à des niveaux géographiques très différents selon l'objet retenu. Ce phénomène renforce l'intérêt de l'information statistique régionalisée et notamment l'utilité de l'élaboration de perspectives démographiques à un niveau spatial fin ; par exemple, la planification d'infrastructures touchant aux domaines de l'université, de l'école, de la santé sont de la compétence des cantons.

Le Service cantonal de recherche et d'information statistiques du canton de Vaud (SCRIS) établit régulièrement depuis une dizaine d'années des projections au niveau du canton (600.000 habitants) et à celui de ses 19 districts par la méthode classique des composantes. On soulignera le fait que les districts constituent déjà de petites régions qui peuvent poser des problèmes engendrés par les faibles effectifs en présence : faible « inertie » démographique des phénomènes, rôle important des migrations. Sans trop entrer dans les détails techniques, on signalera que les hypothèses relatives à la fécondité, à la mortalité et aux migrations sont dans un premier temps étayées au niveau cantonal et que, dans un deuxième temps, on ajuste chaque phénomène aux caractéristiques propres à chaque district, tout en respectant le choix de notre scénario cantonal. Les migrations, par exemple, touchant un district dépendent donc :

¹ Avec la méthode des composantes, la fécondité, la mortalité et les migrations sont prises en compte séparément par le modèle de projections. A partir de la population répartie par âge et par sexe, les nombres de survivants, de naissances et de migrants sont calculés année après année conformément aux hypothèses formulées sur l'évolution présagée pour chacune des trois composantes.

- des hypothèses établies à l'échelon cantonal ;
- des arrivées et départs observés à celui du district dans le proche passé (par âge, sexe, nationalité suisse ou étrangère) ;
- éventuellement d'ajustements supplémentaires pour tenir compte d'hypothèses spécifiques à l'avenir du district ; une correction, sur les migrations touchant les autres districts, maintient effectives les hypothèses sur le total des régions.

Les perspectives, publiées sur un horizon plus court pour le niveau régional que pour le canton, nécessitent donc, pour chaque district, d'une part, d'effectuer des ajustements sur la base de données démographiques connues, d'autre part, une certaine connaissance pour pouvoir anticiper des phénomènes propres à une région (impact d'une amélioration conséquente du réseau de transport, par exemple). L'utilisation de la méthode des composantes sous-entend ainsi que l'on considère que le district est quand même suffisamment grand pour posséder certaines caractéristiques démographiques stables.

Alors que d'une manière générale, l'inertie démographique est d'autant plus faible qu'une région est petite, il existe des exceptions ; par exemple, le plus petit des districts vaudois comporte des caractéristiques stables depuis l'après-guerre : le Pays d'Enhaut (4.500 habitants) situé dans les Préalpes, a régulièrement un indice conjoncturel de fécondité au-dessus de la moyenne cantonale et comporte un solde migratoire qui se maintient à un niveau faible. Il est probable que cette situation perdurera encore assez longtemps.

En aval des perspectives de population par district, nous avons encore la possibilité d'établir des perspectives par catégorie de ménages, avec le calcul des conséquences vraisemblables sur la population active résidente et sur la demande de logements².

Des perspectives pour des régions « à la carte »

La difficulté, rencontrée par le démographe qui tente de répondre aux demandes en perspectives démographiques régionales, est que le découpage géographique idéal n'existe pas : cercles électoraux, zones sanitaires, secteurs psychiatriques, agglomérations, secteurs d'aide médico-sociale à domicile, secteurs d'aménagement du territoire, arrondissements scolaires, associations de communes constituées dans le but, par exemple, d'être soutenues par des aides fédérales (LIM) ou cantonales (LDER), etc. correspondent à autant de subdivisions du canton qui répondent à une problématique concrète mais qui sont différentes de celle des districts, même si certains découpages peuvent parfois partiellement coïncider.

La solution qui correspondrait à recalculer des perspectives pour chaque découpage n'est pas réaliste pour plusieurs raisons :

- l'analyse et le réajustement des hypothèses au niveau de tant de régions serait fastidieux et prendrait beaucoup trop de temps ;
- la cohérence spatiale entre les perspectives ne serait pas garantie ;
- certains résultats obtenus par la méthode des composantes seraient de qualité douteuse, vu la finesse de certains découpages (arrondissements scolaires, par exemple).

Dans ces conditions, nous avons préféré l'autre solution qui consiste à répartir les populations calculées (par âge et par sexe) pour chaque district dans les communes³. Cela permet alors, non pas de publier des perspectives par commune, mais de réagréger les perspectives par commune en n'importe quelle zone raisonnablement grande.

² MENTHONNEX J., 1995, «Un modèle de prévisions de ménages pour la Suisse», Ménages, familles, parentèles et solidarités dans les populations méditerranéennes. Actes du colloqued'Aranjuez, Paris, AIDELF/PUF

³ Il y a 385 communes dans le canton de Vaud comprises entre 33 et 117'000 habitants

Dans la suite de cette communication, nous allons aborder la problématique du choix du procédé utilisé pour répartir les perspectives, obtenues par la méthode des composantes, dans des sous-régions. Comme nous allons le voir, cette démarche est parfois utilisée par certains, pour élaborer des perspectives intra-régionales directement, (c'est-à-dire sans réaggrégation par zone comme nous le faisons). Elle constitue une alternative⁴ à la solution retenue généralement en France qui consiste à établir des projections au moyen d'un modèle pouvant travailler sur un zonage variable jusqu'au niveau communal par la méthode des composantes. Tant le modèle PRUDENT que OMPHALE⁵ sont caractérisés par le fait qu'ils s'appuient principalement sur les informations des recensements (notamment sur les migrants pour PRUDENT, et sur une estimation des migrations pour OMPHALE).

La prévision de sous-populations par la méthode des tendances partielles

Il s'agit du titre d'une communication de Marks Banens (IRHIS, Montpellier) récemment parue⁶. Par sous-population, il entend toute population qui fait partie d'une population plus vaste, mieux connue et mieux prévisible qu'elle.

Au lieu de décomposer les populations et les événements, comme par la méthode des composantes, Marks Banens prend le contre-pied en replaçant les populations dans les populations qui les enveloppent. Ce procédé peut s'appliquer à des niveaux successifs telle une poupée russe.

A partir de la population française totale P_0 (les perspectives étant à ce niveau établies par la méthode des composantes), on passe successivement à la région, au département, au bassin d'emploi, au canton et finalement à la commune. Chaque fois on utilisera un coefficient p_n correspondant à la part de la population P_n dans P_{n-1} .

On a donc pour la sous-population communale

$$P_5 = p_5 \cdot p_4 \cdot p_3 \cdot p_2 \cdot p_1 \cdot P_0$$

Les coefficients sont estimés par estimation linéaire calculée sur une longue période (plus de 100 ans). L'extrapolation linéaire est parfois corrigée « à la main ».

Marks Banens a effectué des tests avec la région Languedoc-Roussillon qui lui fait préférer la méthode des tendances partielles à celle des composantes ; il remarque en outre une divergence non négligeable entre les résultats obtenus par ces deux méthodes pour 2020.

La description de cette méthode pratique par sa simplicité, appelée aussi méthode des proportions⁷, convie à quelques remarques personnelles :

- pratiquement la demande de perspectives régionales pour un groupe d'âges spécifiques est courante. Nous ne pensons pas, contrairement à Marks Banens, que des calculs par âge peuvent être effectués de la même façon étant donné les fluctuations importantes apparaissant parfois au cours du temps pour la population d'un groupe d'âges donné ;
- même s'il est bien vrai que l'on est parfois trop influencé par les événements qui viennent de se passer, il me semble inadéquat d'utiliser de trop longues séries pour repérer les tendances : le dépeuplement d'une commune rurale au début du siècle n'a pas grand chose à

⁴ Voir aussi : WATTELAR C., «Les perspectives démographiques par sexe et par âge : simplisme ou sophistication ? », Perspectives de population, d'emploi et de croissance urbaine, Chaire Quetelet'80, Université Catholique de Louvain, Ordina Éditions

⁵ DESCOURS L., POINAT F., 1992, Le modèle de projection démographique OMPHALE, METHODES no19, INSEE

⁶ Population, 4-5, 1994, pp. 1130-1138

⁷ Méthodes de projection des populations urbaine et rurale, Manuel VIII Nations Unies, New York, 1975

voir, par exemple, avec la croissance que l'on pourrait observer pour cette même commune, due à sa proximité d'un centre urbain récemment développé ;

- l'utilisation d'un modèle linéaire pour chaque coefficient p_i induit une curiosité difficilement justifiable : la part de la population P_n dans la population « mère » P_0 suit en fonction du temps un polynôme de degré n ; elle est donc dépendante du nombre de régions intermédiaires.

Une méthode utilisée par le bureau de la statistique du Québec⁸

Ce bureau a publié des perspectives par groupe d'âges quinquennaux pour 102 territoires de références (94 municipalités régionales de comté, 3 communautés urbaines et 5 autres territoires résiduels). Il s'agit des résultats dérivés des perspectives des régions administratives du Québec. Cette méthode⁹, relativement simple, permet d'assurer une cohérence entre les perspectives des comtés et celles des régions.

Le calcul pour l'année $t+1$ et l'âge $a+1$ de la population du comté i ($P_{i,t+1}^{a+1}$) s'établit à partir de et des perspectives, élaborées par la méthode des composantes, pour la région administrative correspondante (Pr) :

$$P_{i,t+1}^{a+1} = P_{i,t}^a \cdot \left(\frac{P_{i,T+1}^{a+1}}{P_{i,T}^a} \right) / \left(\frac{P_{r,T+1}^{a+1}}{P_{r,T}^a} \right) \cdot \left(\frac{P_{r,t+1}^{a+1}}{P_{r,t}^a} \right) = P_{i,t}^a \cdot A_T \cdot \left(\frac{P_{r,t+1}^{a+1}}{P_{r,t}^a} \right)$$

où T correspond à la dernière année « statistiquement » connue.

En d'autres termes, l'évolution, en i , de la cohorte d'âge a dépend de celle déterminée pour la région r , corrigée d'un facteur observé à l'époque T ; cette correction A_T , constante au cours du temps, prend en compte les particularités des migrations et de la mortalité du comté i relativement à sa région r .

La population de moins d'un an a été calculée à l'aide des naissances déterminées au moyen de taux de fécondité.

L'observation de la période 1976-1981 a permis au BSQ de calculer les indices A qui ont été lissés et ramenés sur une base annuelle.

Pratiquement, les résultats obtenus par les Québécois semblent satisfaisants ; il s'avère toutefois que la méthode peut être quelque peu déficiente dans les sous-régions où l'évolution est rapide et subite.

Nous soulignerons que cette méthode sous-entend un modèle qui postule des rapports constants entre les taux par âge de migrations nettes de la sous-région et ceux de la région. Il y a aussi, comme principe sous-jacent, l'idée que les migrations ne dépendent pas de la population totale : par exemple, dans une sous-région peu dynamique avec une population vieille, lorsque la population totale diminuera, par l'effet de la mortalité, il n'y aura pas de « compensation » par des arrivées.

Première méthode utilisée au SCRIS

En prenant en compte la structure spécifique d'une sous-région observée à un moment donné, plutôt que les particularités des flux migratoires, on risque d'être moins tributaire d'un événement extraordinaire.

⁸ Perspectives démographiques infrarégionales, 1981 - 2001 Québec, BSQ 1984, 498 p

⁹ PITTENGER DONALD B., 1976, Projecting State and Local Populations, Ballinger, Cambridge, 247 p (p 183 et suivantes)

Cette méthode, que nous utilisons depuis plusieurs années, prend en compte l'évolution globale de la sous-région i (les communes dans notre cas) relativement à celle de la région mère r (le district correspondant). De plus, les caractéristiques de la pyramide i , par rapport à celle de son district, sont prises en compte en étant considérées comme stables. Cela peut se justifier lorsque ces caractéristiques sont le fait d'équipements spécifiques (type de logements, ménages collectifs) ou le fait d'une croissance globale nettement différente de celle du district.

Pratiquement le calcul est le suivant :

$$P_{i,t}^a = \left(\frac{P_{i,T}^a}{P_{i,T}} \right) \cdot \left(\frac{P_i}{P_r} \right)_t \cdot P_{r,t}^a \quad \text{où} \quad \sum_a P_{i,t}^a = P_{i,t}$$

La population de la commune i s'établit à partir des perspectives par district (par la méthode des composantes) qui sont ventilées par commune moyennant la proportion $(P_i / P_r)_t$ établie par extrapolation des tendances ; cette extrapolation satisfaisant la condition :

$$\sum_{i \in r} \left(\frac{P_i}{P_r} \right) = 1 \quad \forall t$$

Le résultat est encore corrigé par le premier terme qui correspond à la structure par âge de la commune i , observée en T , relative à celle de son district. Finalement un petit ajustement est encore effectué pour s'assurer que

On constatera que cette méthode, contrairement à la précédente, peut être opérationnelle en travaillant par groupe d'âges et sur une série temporelle pas forcément annuelle. Dans la littérature, on la retrouve décrite sous une forme assez différente, dans le manuel VIII des Nations Unies sous la dénomination « méthode d'élimination des différences » (voir note 6, pp. 60-68).

$$\sum_{i \in r} P_{i,t}^a = P_{r,t}^a \quad \forall a, t$$

En résumé, cette méthode maintient les particularités de la structure par âge des sous-régions tout en conservant les tendances observées dans l'évolution du poids de la sous-région et en restant compatible avec les perspectives calculées par âge et par sexe de la population « mère » (les particularités et les tendances étant définies par rapport à la région mère).

On pourrait reprocher à la méthode de ne pas reproduire des effets de génération : si, pour une raison historique, un groupe d'âges est sur (ou sous)-représenté dans une commune, n années plus tard c'est toujours au même endroit (au même âge) et non pas n ans plus haut, que l'on retrouvera cette particularité.

Une deuxième méthode, privilégiant les effets de générations

Un modèle simple permet d'avoir une alternative au modèle précédent ; il privilégie des effets de générations à la place d'effets d'âges et il présuppose des migrations réparties de façon homogène à l'intérieur de la région r à la place d'hypothèses de croissance spécifiques à chaque sous-région, pour $a \geq n$:

$$P_{i,T+n}^a = \left(\frac{P_{i,T}^{a-n}}{P_{r,T}^{a-n}} \right) \cdot P_{r,T+n}^a$$

Les résultats des perspectives pour la région r sont répartis dans les sous-régions âge par âge, en tenant compte du poids de chaque génération observée en T dans ces sous-régions.

Comme avec la méthode utilisée par le BSQ, il faut déterminer la base de chaque pyramide par un autre moyen ; pour notre part, nous avons réparti les perspectives en tenant compte d'un indicateur de la fécondité générale relative pour $a < n$:

$$P_{i,T+n}^a = \text{IFGR}_{i,r} \cdot \left(\frac{P_i^{20-39}}{P_r^{20-39}} \right)_{T+n-a} \cdot P_{r,T+n}^a \quad \text{où} \quad \text{IFGR}_{i,r} = \left(\frac{P_i^{0-4}}{P_i^{20-39}} \right) / \left(\frac{P_r^{0-4}}{P_r^{20-39}} \right)_T$$

Vraisemblablement, ce procédé n'est intéressant que si les taux de croissance des sous-régions sont semblables à l'intérieur d'une région et s'il n'y a pas de sous-régions avec des structures par âge spécifiques. Par exemple, pour une région comportant une partie typiquement urbaine, la méthode précédente devrait être nettement meilleure.

L'épreuve des faits

Pour tester empiriquement, dans notre contexte, les différents procédés décrits précédemment, nous avons réparti la population des districts observée dans les communes vaudoises sur la base des séries 1980-1986. Des perspectives ex-ante ont été calculées afin de pouvoir comparer les divergences entre résultats et observations pour l'année 1993.

Pour chaque méthode de répartition utilisée et pour chaque commune, des indicateurs d'écarts ont permis de comparer les « performances » des procédés. On soulignera en passant que le nombre de sous-régions ou le poids relatif d'une sous-région dans sa région ont une influence sur la qualité apparente des résultats...

C'est la méthode utilisée par le BSQ qui a posé le plus de problèmes : en calculant les coefficients A sur la base d'une période courte (2 ans), les coefficients sont très sensibles à des migrations extraordinaires (voire aussi aux imprécisions entachant notre connaissance de la population par âge des communes). Même après un lissage des coefficients en fonction de l'âge, on constate qu'il peut parfois y avoir de grandes divergences avec la réalité. Il faut dire que les coefficients A, fixes dans le temps, sont utilisés sur les résultats de l'année précédente ; autrement dit, les erreurs, ou les situations extrêmes, intégrées dans A, créent des processus exponentiels.

On en déduit que, d'une manière générale, même si les résultats sont plus satisfaisants lorsque la commune est grande, il ne faut pas perdre de vue deux choses :

- ne jamais utiliser cette méthode sans contrôle soigneux des résultats ;
- prendre un soin tout particulier pour choisir les coefficients de A.

Des sous-régions relativement grandes, un calcul de A sur la base d'une période de plusieurs années, des procédés de lissage sont autant d'éléments qui permettent à cette méthode d'être opérationnelle.

Les deux autres méthodes sont nettement mieux appropriées pour notre situation qui comporte des communes de très petite taille car, pratiquement, les résultats ne peuvent jamais être complètement aberrants.

Lorsque, pour juger des performances entre les deux méthodes, on compare les écarts entre pyramides, groupe d'âges par groupe d'âges, la méthode privilégiant les effets de générations vient en tête.

Par contre, si l'on souhaite donner l'avantage à la méthode qui limite les risques relativement à l'estimation de la population totale, la première méthode doit être préférée.

De fait, la méthode idéale n'existant pas, il faut choisir entre deux manières de travailler :

- soit on analyse finement les caractéristiques de chaque région afin de choisir le procédé technique le plus adéquat, voire celui qui permet d'introduire à la main des hypothèses spécifiques aux sous-régions;
- soit on utilise au niveau des sous-régions le procédé « automatique », le moins mauvais en général, qui est suffisamment robuste pour limiter les risques d'erreurs grossières en toutes circonstances.

Pour notre part, comme nous avons déjà intégré des hypothèses spécifiques aux régions à l'échelon des 19 districts, nous préférons suivre un procédé unique, systématique et robuste pour toutes les régions au niveau des 385 communes. Ce point de vue est renforcé par le fait que nous réagrégeons les résultats communaux pour obtenir des perspectives selon le découpage « à la carte » demandé.

Dans ces conditions, une solution apparemment complexe, mais pratiquement simple une fois programmée, a été utilisée : l'utilisation d'un compromis entre les deux méthodes.

La recherche d'un compromis

Dans la réalité, les mécanismes intervenant dans les deux procédés existent et permettent donc de les justifier. Empiriquement, on constate que l'un comme l'autre est susceptible de donner des résultats satisfaisants suivant les caractéristiques de la commune.

Néanmoins, l'analyse de performance montre que, en moyenne, les écarts ne sont pas du même ordre à tout âge. Cela nous a amené à déterminer, pour chaque groupe d'âges quinquennal, quel était le meilleur compromis entre les deux méthodes pour que les résultats soient le plus proche (au sens des moindres carrés) des observations. Cet exercice peut se faire tant pour les 385 communes que pour les communes d'un district sélectionné.

Les résultats obtenus, présentés à la figure ci-jointe, montrent que la 2^{ème} méthode (effet de générations) est souvent la meilleure pour définir la population âgée. De nombreuses communes sont par ailleurs marquées par une sur-, ou sous-représentation systématique de jeunes (15-24 ans) : c'est vers ces âges que la 1^{ère} méthode marche particulièrement bien. La recherche des coefficients de pondération optimaux pour les communes d'un seul district montre que l'on retrouve souvent le même profil ; des exceptions apparaissent toutefois : pour les âges élevés la 1^{ère} méthode semble bien adaptée pour les communes du district de Lavaux (qui a un nombre assez important d'établissements médico-sociaux) alors que la 2^{ème} méthode est tout à fait satisfaisante quel que soit l'âge pour certains districts ruraux (Aubonne, Avenches, Oron).

Le fait de finalement retenir un compromis, dépendant de l'âge, entre les résultats obtenus selon chacune de deux méthodes sous-entend des hypothèses assez restrictives : chaque sous-région est constituée de deux populations, l'une suivant le comportement décrit par la 1^{ère} méthode, l'autre suivant celui décrit par la 2^{ème}. La proportion, pour un groupe d'âges donné, entre les deux populations est indépendante du temps et de la sous-région.

De fait, des coefficients de pondération fixes au cours du temps ne peuvent théoriquement pas se justifier et leur indépendance à la sous-région ne semble pas réaliste. On a donc ici un procédé empirique simple qui assure néanmoins un résultat satisfaisant et non pas un modèle étayé par une démarche scientifique.

Un modèle plus satisfaisant, d'un point de vue théorique, nécessiterait plus de paramètres à ajuster aux caractéristiques de chaque sous-région.

Nous signalerons encore un modèle qui va dans cette direction tout en conservant certains aspects empiriques.

Un modèle de projection démographique par quartier

Sous ce titre¹⁰, nous avons présenté un modèle, testé sur six quartiers de la ville de Lausanne, qui correspond à une variante plus sophistiquée du choix du compromis. En effet, deux adjonctions ont été effectuées :

- la population obtenue par la 2^{ème} méthode (effet de génération) a été complétée par un terme additif $\cdot \alpha_{i,t} \cdot \omega_a$ pour pallier l'hypothèse trop contraignante de taux par âge de mortalité et de migrations nettes identiques pour les sous-régions de chaque région (les quartiers de la ville). Pour chaque quartier i , on rajoute (ou on soustrait) $d_{i,t}$ migrants que l'on répartit par âge selon le vecteur. Ces migrants correspondent en quelque sorte aux mouvements extraordinaires par rapport à la tendance générale déjà prise en compte dans le premier terme ; par définition, leur somme est nulle $\sum d_i = 0$. Ce modèle permet de simuler des cycles sur la structure d'âge de certains quartiers (les quartiers « vieux » deviennent « jeunes » puis revieillissent...). Des tests ont montré qu'il était avantageux d'utiliser deux vecteurs de répartition de ces mouvements spécifiques : l'un lorsque d_i est positif, l'autre, lorsqu'il est négatif ;
- la manière de pondérer les résultats correspondant aux deux méthodes a été adaptée. Outre les coefficients dépendant de l'âge, mais communs à toutes les sous-régions, nous avons introduit un coefficient propre à chaque sous-région (à chaque quartier).

Ces adjonctions améliorent l'intérêt et la qualité du modèle au détriment de sa simplicité. L'estimation des paramètres ne peut plus se faire directement. Il a en effet été nécessaire d'utiliser un procédé itératif (note 10). Expérience faite, le processus de calcul converge vers une solution sans être sensible aux valeurs initiales choisies. L'estimation des paramètres a aussi permis de mieux déceler certaines spécificités des quartiers et offre un intérêt descriptif évident.

La richesse de ce procédé nécessiterait néanmoins plus de temps pour le rendre opérationnel au niveau des 385 communes vaudoises regroupées en 19 districts.

Conclusion

L'utilité de disposer de prévisions démographiques pour des régions adaptées aux objectifs de planifications paraît clairement établie. Le territoire adéquat, qui ne correspond alors pas forcément à une entité administrative, peut ne regrouper qu'un faible effectif de population. Dans ces conditions, les démographes doivent poursuivre leurs réflexions sur les méthodes à utiliser lorsque la région est particulièrement petite. La méthode idéale n'existe pas ; il faut trouver un compromis entre deux voies :

- la méthode qui prend fidèlement en compte les spécificités locales mais qui nécessite une bonne connaissance de la région et une analyse démographique fine, au risque d'être « piégé » par des événements exceptionnels lorsque les hypothèses ont été définies trop mécaniquement sur la base de l'évolution passée,
- la méthode simple et robuste, plus fiable sur le long terme (moins sensible à des événements extraordinaires récents), mais qui risque de moins bien prendre en compte certains phénomènes propres à la région et identifiables comme pouvant probablement influencer l'avenir.

¹⁰ MENTHONNEX J., 1993, "Un modèle de projection démographique par quartiers", Croissance démographique et urbanisation, séminaire de Rabat, 1990, PARIS, AIDELF/PUF

L'utilisation des méthodes démographiques en démographie appliquée¹

Louis G. POL²

College of Business Administration
University of Nebraska at Omaha

Introduction

En général, la sous-discipline de la démographie étiquetée *démographie appliquée* est considérée comme faisant partie des directions relativement nouvelles de la recherche. L'utilisation peu fréquente de ce terme dans les années 70 et le fait qu'avant 1980, rares étaient les démographes qui se considéraient comme étant des démographes appliqués, vient soutenir cette affirmation. Les démographes de l'Association de la Population de l'Amérique (PAA) qui se qualifiaient eux-mêmes de « démographes appliqués » n'ont eu la possibilité de l'afficher dans les enquêtes sur les membres de la PAA qu'au milieu des années 80. L'association des chercheurs intéressés par la démographie appliquée, au sein de la PAA, a été créée au début des années 80.

La revue *American Demographics*, considérée par beaucoup comme la première publication centrée à peu près exclusivement sur les applications démographiques, est apparue en 1979. Le premier livre qui incluait la démographie appliquée dans son titre a été publié en 1984 (Rives and Serow, 1984). *Applied Demography*, le journal officiel de l'association dont le centre d'intérêt est la démographie d'entreprises (locales ou nationales) n'est paru qu'au milieu des années 80.

Pourtant, un examen plus approfondi de la littérature démographique conduit à des conclusions quelque peu différentes en ce qui concerne les origines de la démographie appliquée. Bien que la plupart des recherches semblent avoir été initiées à partir de 1980, les démographes se sont engagés dans la démographie appliquée longtemps auparavant. Les sources démographiques ou non contiennent des articles sur ce sujet qui furent publiés dans les années 60 (par exemple, Casparis, 1969 ; Starsinic et Zitter, 1968 ; Goldstein, 1968 ; et Schmitt, 1968) et quelques ouvrages trouvent leurs origines dans les années 50 (par exemple, Frisen, 1958). Vers la fin des années 50, Bogue introduit le concept de la micro-démographie et un modèle de démographie appliquée pour des champs restreints. Le modèle présentait des applications de planification générale, considérant le transport et les équipements, le renouvellement urbain et l'analyse de marché, comme trois axes d'impact. L'analyse du marché, comprenait l'évaluation des emplacements, nouveaux et existants, des centres commerciaux et des petits commerces. Il est clair que Bogue avait en tête la même direction de recherche que celle suivie aujourd'hui par les « démographes appliqués ».

Il est fort possible qu'au début du développement de la démographie en tant que discipline, il y ait eu des démographes engagés dans des travaux de démographie appliquée, mais ces ouvrages n'ont jamais été publiés parce que :

1) la discipline de la démographie était plus strictement ciblée qu'aujourd'hui ; par conséquent, une approche appliquée était considérée comme ayant peu de signification ;

¹ Je veux remercier Jakie Lynch et Irina Gotcu pour l'assistance accordée à la préparation de cet article.

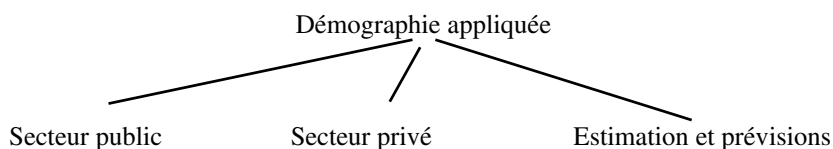
² LPOL@CBFACULTY.UNOMAHA.EDU

2) certains des rapports issus de ces efforts étaient du domaine de la propriété privée et n'ont donc jamais été destinés à être diffusés. Aujourd'hui encore, quelques « démographes appliqués » essayent de ne pas se faire remarquer parce que leurs employeurs ne veulent pas faire connaître la nature de leur travail. De plus, certaines distinctions appliqué/non-appliqué sont d'ordre sémantique et si l'on utilise une définition très large, on peut déclarer que les démographes étaient engagés dans les recherches de la démographie appliquée depuis longtemps.

Au delà de l'histoire, le travail contemporain dans le champ de la démographie appliquée est impressionnant en ce qui concerne la rigueur méthodologique et les domaines d'activités où les théories, les méthodes et les données démographiques ont été appliquées. Le nombre de démographes travaillant dans ce domaine est plus grand que jamais. Un quart environ des membres de la PAA a identifié la démographie appliquée parmi leurs trois domaines majeurs d'intérêt et la même proportion reçoit la publication *Applied Demography*. Pendant la réunion annuelle de la PAA, les sessions sur la démographie appliquée sont parmi celles qui ont l'assistance la plus nombreuse. Depuis 1986, la *Conférence Internationale sur la Démographie Appliquée* a été organisée tous les deux ans à Bowling Green State University. Le premier rapport d'études de cas de démographie appliquée a été publié récemment (Kinter, Merrick, Morrison and Voss, 1994) et il a été très bien accueilli. Une publication spéciale « Revue des politiques et des recherches sur la population » centrée uniquement sur la démographie appliquée a été publiée en 1996.

De plus, il faut reconnaître que l'utilisation du terme « démographie appliquée » incluant le champ entier des applications du secteur public et privé ne reflète pas suffisamment les types de recherches qui sont faites. En fait, nous proposons de diviser la démographie appliquée selon les trois seules composantes présentées dans la Figure 3. Comme l'on peut le voir, il y a une distinction entre les applications du secteur public et celles du secteur privé. Si les méthodes et le matériau utilisés sont, le plus souvent, les mêmes, l'orientation spécifique de la démographie d'entreprise (tournée vers le profit et, relevant le plus souvent du secteur privé), justifie de la considérer séparément. Une bonne partie du travail appliqué au secteur public semble avoir aujourd'hui une histoire plus longue et, plus récemment, aurait pu être classé comme une partie de la rubrique plus générale de la démographie sociale. Ford et DeJong (1970 :4-14) considèrent que la démographie sociale inclut les relations entre démographie et étude des phénomènes sociaux. D'ailleurs, la démographie sociale implique l'intersection de trois systèmes majeurs d'analyse : démographie, société et action sociale. Les systèmes d'actions sociales incluent les systèmes politiques, économiques et de santé, qui sont les centres d'intérêts majeurs de la démographie appliquée actuelle, bien qu'il est difficile de savoir si Ford et DeJong ont désiré ou non inclure le secteur privé et la démographie au service des entreprises comme une partie de leur vision des systèmes économiques.

Figure 1 : Composantes de la démographie appliquée



L'estimation et la prévision des données, définies antérieurement plus strictement comme l'estimation et la projection de la population, représentent la troisième catégorie distincte. Les applications au secteur public et privé se concentrent sur la prise de décision (Burch, Swanson and Trendow, 1996). Cette catégorie est particulière parce qu'elle a ses propres méthodes bien établies ainsi que ses ensembles de données et l'histoire de son développement est relativement longue. Depuis le milieu des années '80, certains démographes considèrent la démographie appliquée et les estimations et les prévisions comme un ensemble unique (ex., Rives et Serow, 1984). Les mettre dans la même catégorie n'a rien de surprenant étant donné qu'il s'agit de la composante la plus facile à distinguer du reste de la démographie.

Cet article se propose d'étudier l'utilisation des méthodes démographiques dans ce que l'on appelle la démographie appliquée. Le but de cette étude est de :

1. lister l'ensemble des méthodes utilisées pour résoudre un large éventail de problèmes ;
2. comprendre les tendances de l'utilisation, la sophistication des méthodes ;
3. identifier les champs dans lesquels les méthodes démographiques, peuvent et doivent être utilisées.

1. Un examen de la littérature

Les procédures utilisées pour identifier les études sont simples et directes. Nous avons utilisé une série de mots clé pour identifier les différentes recherches dans les revues américaines, bien que nous désirons inclure des travaux des auteurs du monde entier. Une fois les sources et les résumés produits, on a obtenu des copies de ces articles. De plus, nous avons utilisé nos contacts et connaissances dans la discipline pour avoir accès aux articles et aux rapports qui n'ont pas été publiés ou à paraître. Bien que l'on considère que le domaine a été très largement couvert, on reconnaît aussi que, selon toute probabilité, la recherche ne couvre pas toutes les contributions méthodologiques "importantes" de la démographie appliquée. Aussi, les références utilisées dans cet article constituent seulement un sous-ensemble des tous les articles trouvés.

Avant d'évaluer les méthodologies, dressons le tableau final qui ressort de cette étude de la littérature. Une série d'études de démographie appliquée ont apporté une contribution significative à la recherche, mais la méthode démographique utilisée n'est pas identifiable. Une grande partie de ces études utilisent la statistique descriptive et introduisent notamment les variables démographiques dans des analyses multivariées. Ces travaux représentent une contribution importante à la littérature de la démographie appliquée.

Dans ce sens, on a accordé une attention plus particulière à un domaine spécifique, celui de l'impact de la population vieillissante sur les secteurs du marché. Bien qu'elles soient descriptives, quelques études se sont centrées sur l'importance et la croissance soutenue du marché des seniors (ex. Goldstein, 1968 ; Reineke, 1964 ; Morrison, 1990). Bartos (1980) fut précurseur en insistant sur la nécessité de segmenter le marché des seniors et des travaux plus récents ont mis l'accent sur la segmentation du marché seniors en domaine de produits et de services spécifiques (ex. Pol, May et Hartranft, 1992). D'autres études, plus liées à l'âge en général, ont mis en avant les différences, en fonction d'âge, dans la consommation de nombreux produits et services (ex., Exter, 1991).

D'autres études encore, la plupart descriptives, mettent l'accent sur des politiques et/ou des actions particulières, par exemple la demande de logements (Morrison, 1977), la transformation de l'activité de la population noire (Boyd, 1990), l'impact des lois publiques sur le recrutement du personnel (Leck et Summers, 1992), l'emplacement des sites (Casparis, 1969 ; Tayman, Parnot et Carnevale, 1994), l'impact des lois sur la vente de détail (McNeil and Yu, 1989), et l'impact des changements de résidences sur la composition de la population scolaire (Clark, 1987). Du point de vue analytique, les indicateurs démographiques prédictifs

ont été utilisés pour sélectionner des marchés de produits et de services (ex., Johson, 1994) et pour produire des prévisions dans le cadre de programmes de planification (ex., Dunton, 1994).

2. Les méthodes utilisées

2.1 Estimation et prévision des données

L'estimation et la prévision des données, composante de la démographie appliquée, est la seule catégorie qui a son propre ensemble de méthodologies. Ces méthodes ont été passées en revue ailleurs (ex. Murdock et Ellis, 1991) et le but de cet article n'est pas de faire le même type d'analyse. Nous mettons néanmoins l'accent sur ce qui distingue les applications les plus récentes du travail traditionnel de prévision et d'estimation :

- 1) les unités géographiques pour lesquelles les estimations et les prévisions sont désormais réalisées sont beaucoup plus petites ;
- 2) il y a une augmentation très nette du nombre de variables démographiques projetées ;
- 3) les prévisions démographiques sont désormais liées à d'autres facteurs (ex. les taux d'incidence, de prévalence, de consommation) pour réaliser des prévisions/pronostics d'autres phénomènes et
- 4) les « marchands de données » privés ont contribué à la création d'un environnement où les estimations et les prévisions sont désormais plus exactes et plus faciles à réaliser. Le souci de l'exactitude et de la rigueur méthodologique a, malheureusement, été occulté par l'attrait de la disponibilité et peu d'utilisateurs de ces données semblent préoccupés par leur qualité.

En ce qui concerne les petites unités géographiques, par exemple, si la méthode démographique de l'unité-logement a d'abord été appliquée dans de larges aires géographiques comme les régions et les villes (Starsinic et Zitter, 1969 ; Smith et Lewis, 1980 ; Smith et Lewis, 1983), les perfectionnements récents ont visé des aires plus réduites - les zones de recensement (Smith, 1994). Il faut reconnaître que les efforts de perfectionnement de cette méthode en vue de réaliser des estimations de bonne qualité sur de petites unités géographiques ont été initiés à la demande du secteur public et du secteur privé concernant des données nécessaires au processus de prise de décision qui peut être à un micro échelon. Les estimations des petites aires sont nécessaires pour des décisions qui vont de l'emplacement des casernes de pompiers (Tayman, 1994) jusqu'au lieu de construction des commerces de détail (ex. Morrison, 1996).

L'augmentation rapide de la demande de données démographiques s'est également traduite par une croissance de l'ampleur des informations démographiques à prévoir. Les estimations et les prévisions sur la population par sexe, âge et race accompagnées d'estimations et de prévisions sur le revenu, l'unité familiale et les structures de l'emploi sont maintenant courantes. Néanmoins, le niveau de la rigueur méthodologique dont on a besoin pour utiliser ces données, n'a pas reçu l'attention nécessaire.

Au delà des efforts d'amélioration et de perfectionnement de la méthode d'estimation de la population par la méthode des logements mentionnée ci-dessus, la plupart des travaux d'estimation et de prévision ont utilisé des méthodes de régression ou de corrélation (ex., Schmitt et Crosetti, 1954 ; Schmitt et Grier, 1966 ; Namboodori, 1972 ; O'Hare, 1976, 1980 ; Martin et Serow, 1978 ; Swanson, 1980 ; Schroeder et Pittenger, 1983). On a constaté des éléments spécifiques d'amélioration dans l'introduction des variables simulées (Pursell, 1970) et des efforts supplémentaires ont été élaborés pour « s'attaquer » à l'incompatibilité de la relation temporelle entre la structure modélisée et l'application (Mandell et Tayman, 1982 ; Swanson et Tedrow, 1984). En plus, en ce qui concerne les méthodologies d'estimation et de prévision, on a fait des progrès au niveau de l'introduction d'autres indicateurs de changement

(ex. données sur les soins médicaux ; Smith, 1986) et par l'amélioration des composantes des estimations/prévisions, comme les profils de migration (Kennedy, DeJong et Lichter, 1986).

Malgré les progrès importants des méthodes d'estimation et de prévisions et l'augmentation logique de l'utilisation de ces données, il faut reconnaître l'existence de quelques problèmes de production/utilisation. Les estimations et les prévisions se basent sur des hypothèses et par conséquent sont susceptibles de contenir des erreurs. Bien qu'une partie des meilleurs travaux d'estimation soient orientés vers le renforcement de la qualité et le calcul des erreurs (Swanson et Tayman, 1996 ; Tayman et Swanson, 1995 ; Schaefer, Tayman et Carter, 1995) et que quelques chercheurs demandent que des standards soient établis, la plupart des « non académiques » accordent surtout de l'importance à l'existence de données plutôt qu'à leur fiabilité. Pour de simples utilisateurs de données, l'utilisation des estimations et des prévisions pour la prise de décision peut être un sérieux problème. Les données sont souvent considérées comme une réalité, par opposition à des évaluations, et cette difficulté est renforcée par la prépondérance de la « boîte noire » des modèles utilisés par toute une série de marchands de données pour la réalisation de prévisions spécifiques sur le revenu, groupe d'âge et race/ethnie pour les petites aires géographiques. Ces données contiennent probablement beaucoup d'erreurs, mais nous ne sommes pas écouté lorsque nous mettons en garde sur le niveau de fiabilité des données.

A l'avenir, certains des meilleurs travaux d'estimation et de prévision seront orientés vers le renforcement et l'analyse de ces modèles, vers la production des critères utilisables pour déterminer quels schémas peuvent être considérés comme fiables dans un but de planification.

2.2 Tables de survie

Les méthodes de la table de survie arrivent en second quant à leur fréquence d'utilisation. D'un certain point de vue, l'application des techniques de la table de mortalité illustrent très bien la valeur de l'utilisation des méthodes démographiques en vue de mieux comprendre les problèmes et les perspectives d'avenir du secteur public et du secteur privé. En fait, les premiers usages de ces techniques n'étaient pas la stricte étude de la mortalité, mais analysait aussi l'incidence des probabilités de décès sur la tarification et la valeur des rentes (Halley, 1963). Et, évidemment, les mêmes formules sont nécessaires dans le domaine des assurances de vie. Le rapport ${}_nL_a/l_x$ est combiné avec l'information sur la prime d'assurance pour estimer l'accumulation des dollars pendant la vie des souscripteurs (Smith, 1992 : 198 - 205). On fait des ajustements opportuns compte tenu de la durée de vie à partir du terme de la police d'assurance. Actuellement, les techniques de la table de mortalité sont utilisées par les vendeurs d'investissements financiers pour fixer le prix des produits et calculer les bénéfices disponibles de ceux qui achètent les divers produits.

Pourtant, les applications dépassent le domaine des marchés financiers. Elles ont été utilisées dans le domaine de la santé ou des soins de la santé dans le cas des maladies chroniques (Manton and Stallard, 1988) et dans le domaine de l'épidémiologie pour étudier les effets du traitement sur les malades leucémiques (Cox, 1972). Les applications industrielles se sont orientées vers « les tests de survie », par exemple pour calculer le temps moyen avant qu'un produit ou une partie du produit tombe en panne, le taux de panne, une fonction de survie ou une fonction de vie efficace (par ex. Bain, 1978 ; Nelson, 1982). Les principes de l'extinction multiple ont été utilisés pour construire des tables de cursus scolaire et les calculs d'acrémentation / incrémentation ont été utilisés pour produire des tables de vie active. On a produit d'autres tables de survie de type extinction multiple pour les handicaps.

Les applications les plus récentes ont utilisé le principe des tables dans quelques autres voies. Garcia (1994) a construit une table d'inscription, avec l_x , représentant une nouvelle cohorte d'étudiants, q_x la proportion d'étudiants inscrits l'année x mais non inscrits l'année $x+1$ et e_0 le nombre moyen d'années passées avant la fin des études ou abandonnées sans avoir une

licence. Ces calculs peuvent être utilisés pour établir le taux de l'obtention d'une licence, pour réaliser une estimation du temps moyen pour l'obtention d'un diplôme, et pour aider à prévoir les inscriptions. Kintner et Swanson (1994) ont utilisé la banque de donnée de General Motors (GM) pour calculer l'espérance de vie des ouvriers de GM. Les données ont ensuite été utilisées pour prévoir l'hospitalisation d'un groupe de retraités, projeter l'évolution des cohortes des retraités actuels, estimer les dépenses hospitalières et prévoir l'utilisation des hôpitaux par cohortes.

2.3 Analyse par cohorte

Le terme cohorte est utilisé en tant que concept (Ryder 1965) et comme modalité pour décrire un groupe particulier d'instruments analytiques destinés à l'étude des cohortes (ex. Glenn 1997, Palmore 1978). Une cohorte est « ...un ensemble d'individus ayant vécu un événement semblable pendant la même période de temps. » (Ryder 1965 : 845). Le plus souvent, les démographes utilisent les cohortes de naissance, de mariage et/ou d'éducation dans des buts analytiques, mais le concept intéresse ceux qui étudient les changements de la main d'œuvre dans une organisation (c.-à-d. ceux qui ont été embauchés en même temps), le choix d'un produit et sa dispersion (c.-à-d. ceux qui achètent un produit ou un service en même temps) et les chercheurs qui s'occupent des prévisions sur les ventes.

L'analyse des cohortes commence d'abord par l'identification des cohortes et puis étudie les comportements, les opinions et les caractéristiques à travers le temps. La clé de cette analyse est la capacité de discerner les effets différentiels du temps et de l'appartenance à une cohorte sur les facteurs liés à l'achat des produits ou autres comportements concernant une entreprise spécifique. De plus, l'analyse inclut souvent l'âge comme variable et les études les plus traditionnelles essayent de faire des distinctions entre les influences de la cohorte, de la période (temps) et de l'âge sur les facteurs mentionnés ci-dessus (ex : Palmore 1978 et Rentz et Reynolds 1980).

L'analyse par cohorte est considérée comme un instrument de la planification stratégique (Reynolds et Rentz 1981). Parce qu'ils voulaient mieux comprendre l'effet du vieillissement de la population sur la consommation des produits et des services, ces auteurs ont créé un diagramme qui reflète la consommation hypothétique spécifique à chaque âge d'un produit par six intervalles de temps (1930 - 1980). En utilisant la méthode en triade de Palmore (1978), longitudinale (effets d'âge et de la période), transversale (effets d'âge et de la cohorte) et diachronique (effets de période et de cohorte) des effets différenciés furent identifiés et quantifiés. Autrement dit les effets de période (les différences de consommation dans la même cohorte à deux périodes de temps) et les différences entre les cohortes furent mises en avant en ce qui concerne les variations de la consommation du produit. Cette application implique le calcul simple de différences absolues et relatives, bien que d'autres méthodologies plus sophistiquées soient disponibles pour cette tâche (Hallie et Rao, 1992 : 49 - 61). Rentz, Reynolds et Stout (1983) ont utilisé la régression multiple pour étudier l'effet des cohortes, de l'âge et de la période sur la consommation des boissons non-alcoolisées de 1950 à 1979. L'une des raisons de la réalisation de cette étude fut la constatation que la consommation de boissons non-alcoolisées décroît au fur et à mesure que l'âge augmente et ils formulèrent l'hypothèse selon laquelle une population vieillissante peut connaître une réduction de la consommation globale des boissons non-alcoolisées. Les auteurs ont constaté que les effets de la cohorte sont plus grands que les effets de période, ce qui, affirmèrent-ils, étaient un résultat du bon sens vu les efforts croissants de marketing déployés et constatés dans cette période de temps. Du point de vue stratégique, les résultats soutiennent l'idée selon laquelle le total des ventes ne devrait pas décroître à cause du vieillissement de la population.

Bien que ces études introduisent clairement les méthodologies de l'analyse par cohorte chez les non-démographes et offrent quelques exemples sur la manière dont les études peuvent

être réalisées, leur impact global sur les recherches en démographie appliquée a été assez réduit. Autrement dit, malgré l'intérêt évident de l'analyse par cohortes démontré par ces études, la méthodologie n'a pas fait de progrès par rapport au travail réalisé dans les années 80 et on ne trouve pas d'applications plus récentes dans la littérature. Les chercheurs en marketing, par exemple, ne considèrent pas l'analyse par cohortes comme étant l'une de *leurs* options méthodologiques. Cette situation est, malheureusement due au fait que les chercheurs en marketing, ainsi que les autres chercheurs en démographie appliquée, sont généralement intéressés par la consommation globale et les autres tendances comportementales qui apparaissent comme résultant du changement culturel (effets de période) et des comportements constants ou en évolution à l'intérieur des cohortes (effets de cohortes). Les prévisions sur les ventes, par exemple, sont fondées souvent sur l'hypothèse que les comportements peuvent ou non changer et sur la croissance du nombre de la population des consommateurs. Les taux d'utilisation, actuels ou prévus, sont multipliés par les populations projetées pour évaluer les prévisions de ventes. Si nous pouvions mieux comprendre les effets du changement du comportement à travers le temps, notamment les variations entre les cohortes, alors la précision des prévisions sur les ventes pourrait être considérablement améliorée.

2.4 Standardisation

L'étude classique de Kitagawa (1964) sur la standardisation des taux dans la recherche comparative a engendré une nouvelle et précieuse série d'instruments méthodologiques dans les études démographiques. Grâce aux taux comparatifs les effets de la structure de l'âge, par exemple, ont été clarifiés, et la décomposition des indices facilitée (Hallie and Rao, 1992 : 8 - 15). L'introduction des tests statistiques pour l'analyse de la différence des taux a perfectionné les comparaisons (Smith, 1992 : 62 - 70). L'élargissement du concept de comparaisons standardisées a avancé en appliquant les modèles log-linéaires (Clogg and Eliason, 1988).

Malgré les avantages évidents des procédures de standardisation pour la comparaison des marchés, des segments de marché et d'autres secteurs public et privé concernés, on ne trouve que quelques exemples d'utilisation. Billings et Pol (1994) ont appliqué les procédures de standardisation pour comparer le marché potentiel pour l'introduction de la téléphonie mobile dans deux régions de Floride, avec comme objectif final de déterminer la valeur de chaque marché. Les données initiales montraient que l'achat des téléphones variait considérablement selon l'âge et les régions avaient généralement des structures d'âge différentes. Les indices de pénétration de la téléphonie mobile ont été calculés et ajustés selon les différences de structure par âge ce qui finalement conduisit au changement en valeur relative entre les deux marchés. On peut trouver une analyse similaire chez Pol et Tymkiw (1991) qui comparent deux marchés relatifs à la consommation de la bière. Les deux marchés différaient en termes de dimension et d'âge or, la consommation de la bière variait considérablement en fonction de l'âge. On a constaté que la puissance relative des deux marchés changeait lorsque les indices standardisés de la consommation de bière par l'âge furent calculés.

Deux articles de Pol et Pak (1996a, 1996b) ont utilisé la standardisation log-linéaire pour faciliter les comparaisons entre les segments de marché. Dans l'étude précédente (1996a) les effets de l'âge et de la race/ethnie ont été éliminés en ce qui concerne les comparaisons dans trois segments de marché ayant trait à la nourriture consommée à l'extérieur du domicile (non dépendants, faibles et forts). Les comparaisons des taux comparatifs entre les segments de marché et des taux bruts à l'intérieur et entre segments conduisent à l'identification des sous-marchés qui semblent avoir du potentiel pour l'augmentation des ventes. Dans la dernière étude (1996b), la standardisation a été utilisée pour faciliter les comparaisons entre les segments d'âge du marché des seniors (50 ans et plus) en considérant leur état de santé subjectif. Les implications de marketing de ces comparaisons « à l'état pur » ont été discutées.

3. Discussion

Il est clair qu'il y a un certain éventail de méthodes démographiques utilisées en démographie appliquée, même s'il est vrai que beaucoup d'études sur la démographie appliquée n'utilisent pas de méthodes particulières. Bien qu'il faille reconnaître la contribution à la littérature des dernières études, l'absence de rigueur méthodologique peut déconcerter nombre d'adeptes de la démographie appliquée. De plus, il faut noter que les méthodes dont on a discuté ici sont probablement un sous-ensemble de toutes les méthodes utilisées. Les méthodes de lissage des données, d'ajustements de distribution de revenus tronqués, par exemple, sont vraisemblablement utilisées dans certaines études ; elles ne sont pas encore particulièrement discutées parce que leur utilisation est habituelle.

En général, pourtant, en dehors du domaine bien défini des estimations et des prévisions sur la population, les méthodes démographiques utilisées sont souvent peu développées ou le sont de façon inadéquate. Selon toute probabilité, la nature du travail sur des marques déposées (du domaine de la propriété privée) écarte l'introduction publique/professionnelle de certaines des meilleures méthodes. Par conséquent, il y a beaucoup de travail à faire en ce qui concerne 1) l'introduction plus générale de ces méthodes pour une plus large audience des utilisateurs potentiels et 2) l'utilisation de ces méthodes ou d'autres pour les applications qui interviennent significativement dans la prise de décision.

Le but à long terme de ces applications doit être l'intégration de ces techniques dans un ensemble plus large de méthodes utilisées dans une série de disciplines. Par exemple, la comparaison standardisée des zones de marché devrait se faire couramment étant donné les ressources exigées et les risques auxquels il faut faire face lorsqu'on aborde un nouveau marché ou que l'on reconsidère les anciens marchés.

De plus, d'autres méthodes démographiques doivent être mises en avant. Par exemple, beaucoup de démographes ont utilisé et perfectionné des techniques d'estimation des phénomènes et structures démographiques sur la base de données incomplètes ou peu fiables.

En même temps, l'utilisation des méthodes qui existent déjà dans la littérature doit être élargie. Par exemple, Keyfitz (1977 : 357 - 359) offre un court exemple sur la manière dont l'analyse par les tables peut être utilisée pour étudier le processus de promotion dans le cadre des organisations. Étant donné les intérêts en jeu dans la stagnation d'une organisation et l'existence d'un plafond, il semble logique que l'exemple devrait être développé pour inclure les données d'une ou d'un groupe d'organisations.

BIBLIOGRAPHIE

- BAIN, Lee J., (1978), *Statistical Analysis of Reliability and Life Testing Models*, New York, Dekker.
- BARTOS, Rena, (1980). « Over 49 : The Invisible Consumer Market », *Harvard Business Review* 58 (January) : 140-148.
- BEALE, Calvin L., (1964). Rural Depopulation in the United States : Some Demographic Consequences of Agricultural Adjustments, *Demography* 1 (1), pp. 264-272.
- BILLINGS, George H. and POL, Louis G., (1994), Improving Cellular Market Area Evaluation With Demographic Data, pp. 93-108, in KINTNER Hallie J. et al., *Demographics : A Casebook for Business and Government*. Boulder, CO, Westview Press.

- BOGUE, Donald J., (1957), Micro-Demography, pp. 46-52 in BOGUE Donald J. (ed.), *Applications of Demography. The Population Situation in the U.S.*, Chicago, University of Chicago.
- BOYD, Robert L., (1990), « Black business transformation, Black Well-Being and Public Policy. », *Population Research and Policy Review* 9 (May), pp. 117-132.
- CASPARIS, John, (1969), Shopping center location and retail store mix in metropolitan areas, *Demography* 6 (May), pp. 125-131.
- CLARK, William A. V., (1987), Demographic change, attendance area adjustment and school system impacts, *Population Research and Policy Review* 6 (3), pp. 199-222.
- CLOGG, Clifford and ELIASON, Scott (1988), A flexible procedure for adjusting rates and proportions, including statistical methods for group comparisons, *American Sociological Review* 53 (April), pp. 267-283.
- COX, David R., (1972), Regression models and life tables, *Journal of the Royal Statistical Society*, Series B34, pp. 187-202.
- DUNTON, Nancy, (1994), Planning for children's residential care, pp. 327-342 in KINTNER, Hallie J. et al. (eds.), *Demographics : A Casebook for Business and Government*, Boulder, CO, Westview Press.
- EXTER, Thomas, (1991), Boozing Boomers, *American Demographics*, (December), pp. 6.
- FORD, Thomas R. and DEJONG Gordon F., (1970), *Social Demography*, Englewood Cliffs, N.J., Prentice-Hall Inc.
- GARCIA, Phillip, (1994), Predicting college enrollment results from a variant of the life table, pp. 307-326 in KINTNER Hallie J. et al. (eds.), *Demographics : A Casebook for Business and Government*, Boulder, CO, Westview Press.
- GOLDSTEIN, Sidney, (1968), The aged segment of the market, 1950 and : 1960, *Journal of Marketing* 32 (April), pp. 62-68.
- HALLEY, Edmund (1693), An estimate of the degrees of the mortality of mankind, *Philosophical Transactions*, 17, pp. 596-610.
- HALLI, Shiva S. and RAO, K. Vaninadha, (1992), *Advanced Techniques of Population Analysis*, New York, Plenum Publishing Company.
- HYRENIUS, Hannes and QUIST Jan, (1970), Life table technique for the working ages, *Demography* 7 (November), pp. 393-399.
- KENNEDY, John M., DEJONG, Gordon F. and LICHTER Daniel T., (1987), Updating local area population projections with current migration estimates, *Journal of Economic and Social Measurement* 14 (2), pp. 107-120.
- KEYFITZ, Nathan, (1977), *Applied Mathematical Demography*, New York, Wiley.
- KINTNER, Hallie J. and SWANSON David A., (1994), Estimating vital rates from corporate databases : How long will GM's salaried retirees live? pp. 265-297 in KINTNER Hallie J. et al. (eds.), *Demographics : A Casebook for Business and Government*. Boulder, CO, Westview Press.
- KINTNER, Hallie J., MERRICK Thomas W., MORRISON, Peter A. and VOSS Paul R., (1994), *Demographics : A Casebook for Business and Government*, Boulder, CO, Westview Press.
- LECK, Joanne D. and SAUNERS David M., (1992), Canada's employment equity act : effects on employee selection, *Population Research and Policy Review* 11 (1), pp. 21-49.
- LEGARE, Jacques (1972), Methods for measuring school performance through cohort analysis, *Demography* 9 (November), pp. 617-624.
- MANDELL, Marylou and TAYMAN Jeffrey, (1982), Measuring temporal stability in regression models of population estimation, *Demography* 19 (February), pp. 135-146.

- MARTIN, Julia H. and SEROW, William J., (1978), « Estimating demographic characteristics using the Ratio-Correlation Method. », *Demography* 15 (May), pp. 223-233.
- McNEIL, Douglas W. and YU Shirley S., (1989), Blue laws : Impact on regional retail activity, *Population Research and Policy Review* 8 (September), pp. 267-278.
- MORRISON, Peter A. (1977), Demographic trends that will shape future housing demand, *Policy Sciences* 8 (June), pp. 203-215.
- MORRISON, Peter A. and ABRAHAMESE Allan F., (1996), Applying demographic analysis to store site selection, *Population Research and Policy Review* (forthcoming).
- MORRISON, Peter A., (1990), Demographic factors reshaping ties to family and practice, *Research on Aging* 12 (December), pp. 399-408.
- NAMBOODIRI, N. Krishnan (1972), « On the ratio-correlation and related methods of subnational population estimation. », *Demography* 9 (August), pp. 443-453.
- NELSON, Wayne B. (1982), *Applied Life Data Analysis*, New York, Wiley.
- O'HARE, William (1976), « Report on a multiple Regression Method for Making Population Estimates. », *Demography* 13 (August), pp. 369-379.
- O'HARE, William P., (1980), A note on the use of regression methods in population estimates, *Demography* 17 (August), pp. 541-543.
- PITTENGER, Donald B., (1977), « Population forecasting standards : some consideration concerning their necessity and content. », *Demography* 14 (August), pp. 363-368.
- POL, Louis and TYMKIW Douglas, (1991), A technique to compare demographically different markets, *Marketing Research* 3 (March), pp. 29-34.
- POL, Louis G. and PAK Sukgoo, (1996a), Consumer unit types and expenditures on food away from home, *Journal of Consumer Affairs* (forthcoming).
- POL, Louis G. and PAK Sukgoo, (1996b), Segmenting the senior health care market, *Health Marketing Quarterly* (forthcoming).
- PURSELL, Donald E., (1970), Improving population estimates with the use of dummy variables, *Demography* 7 (February), pp. 87-91.
- REINECKE, John A., (1966), The older market - Fact or fiction? *Journal of Marketing* 28 (January), pp. 60-64.
- RENTZ, Joseph O. and REYNOLDS Fred D., (1980), Separating age, cohort and period effects in consumer behavior, *Advances in Consumer Research* 8, pp. 596-603.
- RENTZ, Joseph O., and REYNOLDS Fred D. and STOUT Roy G., (1983), Analyzing changing consumption patterns with cohort analysis, *Journal of Marketing Research* 20 (February), pp. 12-12.
- REYNOLDS, Fred D. and RENTZ, Joseph O., (1981), Cohort analysis : An aid to strategic planning, *Journal of Marketing* 45 (Summer), pp. 62-70.
- RYDER, Norman B., (1965), The cohort as a concept in the study of social change, *American Sociological Review* 30 (6), pp. 843-461.
- SCHAEFER, Edward and TAYMAN Jeff, (1995), Forecasting errors and confidence intervals for small area population forecasts : The relationship among population size, rate of growth, and precision and bias, Paper presented at the *Population Association of America Annual Meeting*, San Francisco, CA.
- SCHMITT, Robert C., (1968), « Travel, tourism and migration. », *Demography* 5 (1), pp. 306-310.
- SCHMITT, R. C. and CROSETTI, A. H., (1954), Accuracy of the ratio-correlation method for estimating postcensal population, *Land Economics* 30 (August), pp. 279-281.

- SCHMITT, R. C. and GRIER, J. M., (1966), A method of estimating the population of minor civil divisions, *Rural Sociology* 31 (September), pp. 355-361.
- SCHROEDER, Esther C. and PITTENGER Donald, (1983), « Improving the accuracy of migration age detail in multiple-area population forecasts. », *Demography* 20 (May), pp. 235-248.
- SMITH, David P., (1992), *Formal Demography*, New York, Plenum Publishing Company.
- SMITH, Stanley K. (1986), Using medicare data for short-run projections of the elderly population, *Journal of Social and Economic Measurement* 14 (April), pp. 37-49.
- SMITH, Stanley K., (1994), Population estimates, projections, and expert testimony in adversarial legal proceedings : A case study in automobile dealerships, pp. 180-202 in KINTNER Hallie J. et al., *Demographics : A Casebook for Government and Business*, Boulder, CO, Westview Press.
- SMITH, Stanley K. and LEWIS Bart B., (1980), Some new techniques for applying the housing unit method of local population estimation, *Demography* 17 (August), pp. 323-339.
- SMITH, Stanley K. and LEWIS Bart B., (1983), Some new techniques for applying the housing unit method of local population estimation, Further Evidence, *Demography* 20 (August), pp. 407-413.
- STARSINIC, Donald E. and ZITTER Meyer, (1968), Accuracy of the housing unit method in preparing population estimates for cities, *Demography* 5 (1), pp. 475-484.
- SWANSON, David A., (1980), Improving accuracy in multiple regression estimates of population using principles from causal modelling, *Demography* 17 (November), pp. 413-427.
- SWANSON, David A. and TEDROW Lucky M., (1984), Improving the measurement of temporal change in regression models used for county population estimates, *Demography* 21 (August), pp. 373-381.
- SWANSON, David A. and TAYMAN Jeffrey, (1996), Between a rock and a hard place : The evaluation of demographic forecasts, *Population Research and Policy Review*.
- TAYMAN, Jeff, PARROTT Bob and CARNEVALE Sue, (1994), Locating fire station sites : The response time component, pp. 203-217 in KINTNER Hallie J. et al. (eds.), *Demographic s : A Casebook for Business and Government*, Boulder, CO, Westview Press.
- TAYMAN, Jeff and SWANSON, David A., (1995), Alternative measures for evaluating population forecasts : A comparison of state, county, and sub-county geographic areas, Presented at the *Population Association of America Annual Meeting*, San Francisco, CA.
- TAYMAN, Jeff and POL Louis G., (1996), Retail site selection and geographic information systems, *Journal of Applied Business Research*, (forthcoming).

Quel avenir pour le clergé namurois ?

Le point de vue du démographe

Michel POULAIN

Université Catholique de Louvain, Belgique

Introduction

A l'aide des outils qu'il manie couramment, le démographe peut proposer une analyse de l'évolution de toute une diversité de sous-populations et tenter de fournir une vue prospective. Ces deux tâches, souvent indissociables, sont sans conteste des atouts essentiels dans l'aide à la prise de décision. L'exemple que nous envisagerons concerne le clergé namurois, c'est-à-dire l'ensemble des prêtres séculiers rattachés au diocèse de Namur. Ce dernier s'étend sur les provinces de Namur et de Luxembourg, et compte plus de 650.000 habitants. Administrativement parlant, ce diocèse compte 742 paroisses à la tête desquelles se trouve un curé ou desservant. En outre, 250 postes de vicaires paroissiaux sont prévus « au cadre civil » et servent d'appui pour les plus grosses paroisses. Au total, si l'on ajoute l'évêque de Namur, chef spirituel du diocèse, il y a donc 973 postes distincts de « pasteurs d'âmes » pouvant être occupés dans le diocèse. Par ailleurs, une part variable du clergé namurois occupent d'autres fonctions. Ils sont :

- chargés de l'administration du diocèse ;
- enseignants au sein de l'enseignement diocésain ou dans l'enseignement officiel ;
- prêtres au travail engagés directement dans la vie active ;
- prêtres « pro deo » occupant des charges dans les terres de mission, notamment en Amérique Latine et en Afrique francophone ;
- prêtres pensionnés, statut que l'on peut obtenir à partir de l'âge de 65 ans.

En période de pleine occupation de tous ces postes, au lendemain de la seconde guerre mondiale, le diocèse comptait environ 1300 prêtres dont environ 300 étaient chargés de tâches administratives ou d'enseignement. En 1974, on en dénombre 1019 et en 1994, selon l'annuaire officiel de diocèse de Namur, ils ne sont plus que 635. Pour pallier à cette diminution importante du clergé séculier, un appel a été fait aux prêtres réguliers appartenant aux différentes congrégations religieuses du diocèse. En 1974, 113 prêtres réguliers administraient une paroisse, sans toutefois y résider dans la plupart des cas. En 1994, ce chiffre s'est réduit à 72 prêtres réguliers.

L'avenir du clergé namurois pose donc une inquiétude de plus en plus nette au sein du monde catholique. Les estimations et les prévisions du démographe devraient jeter des points de repère afin de mieux asseoir les décisions qui influenceront sur l'avenir du diocèse de Namur.

L'évolution du clergé de 1974 à 1994

Nous avons choisi les vingt dernières années afin de suivre l'évolution détaillée du clergé namurois. Pour ce faire, nous avons dépouillé les annuaires du diocèse de Namur pour les années 1974, 1979, 1984, 1989 et 1994. En outre, nous avons examiné, année après année, la nécrologie relative au clergé namurois. Chaque prêtre a été ainsi identifié avec sa date de naissance et sa date d'ordination. Pour les 5 dates retenues, on vérifie la présence du prêtre parmi la liste du clergé du diocèse et on retient son statut : curé, vicaire, prêtre administratif ou enseignant, prêtre « pro deo », prêtre au travail ou prêtre pensionné. L'absence d'un prêtre dans un des relevés peut s'expliquer par le fait que l'intéressé n'est pas encore ordonné ou entré dans

le diocèse en provenance d'un autre ou, par contre, parce qu'il est sorti vers un autre diocèse, qu'il a renoncé au sacerdoce ou qu'il est décédé. Pour le démographe, le mouvement de la sous-population du clergé comporte donc des « naissances » par ordination des nouveaux prêtres et des « décès » par la mort des prêtres. Les « entrées » se font en provenance d'un autre diocèse à une date postérieure à celle de l'ordination tandis que les « sorties » se font vers un autre diocèse ou par défection. Nous n'avons pu considérer séparément ces deux possibilités de sortie. Le tableau 1 détaille le mouvement de la population du clergé namurois au cours des quatre périodes quinquennales.

TABLEAU 1. LE MOUVEMENT DE LA POPULATION DU CLERGÉ NAMUROIS PAR PÉRIODES QUINQUENNALES ENTRE 1974 ET 1994.

Période	Population initiale	Nouveaux prêtres	Décès	Entrées	Sorties	Population finale
1974-1978	1019	33	100	8	21	939
1979-1983	939	21	126	5	18	821
1984-1988	821	13	123	4	9	706
1989-1993	706	11	81	6	7	635
TOTAL		78	430	23	55	

En l'espace de 20 années, la diminution de l'effectif du clergé approche 40%. Ceci est principalement dû au faible nombre de nouveaux prêtres par comparaison à celui des décès. Dans une mesure moindre, on constate un déséquilibre entre les entrées et les sorties du diocèse et ce, de façon plus marquée, pendant la première décennie 1974-1983 au cours de laquelle on a enregistré une trentaine de défections au sein du clergé. Il convient de remarquer également que le nombre d'ordinations est en forte diminution puisqu'il se réduit d'un facteur deux entre la première décennie et la seconde. Les décès, quant à eux, diminuent en nombre absolu, mais en chiffres relatifs, le taux brut de mortalité des prêtres du namurois est en hausse (24,5% pendant la première décennie et 28,0% pendant la seconde). Le vieillissement d'ensemble du clergé est évidemment responsable de cet accroissement apparent de la mortalité.

Le tableau 2 compare la structure par âge du clergé en 1974 et 1994, il précise le mouvement de la population par groupe de générations pendant les 20 années considérées. Le vieillissement de la structure par âge du clergé namurois est manifeste comme le montre la figure 1. Deux indices témoignent de cet évolution : l'âge moyen des prêtres passe de 53 à 63 ans en l'espace de 20 ans tandis que la part des 65 ans et plus s'accroît de 21 à 44%.

La figure 2 fournit un aperçu de la distribution des âges à l'ordination. Elle permet de constater un âge plus avancé des nouveaux prêtres pendant les cinq dernières années.

La mortalité des prêtres namurois peut être estimée en faisant l'hypothèse qu'ils ne peuvent sortir et mourir pendant la même période quinquennale. Dans ce cas, on peut rapporter les décès enregistrés pendant une période à la population de chaque groupe d'âges en début de période en faisant abstraction du faible nombre d'entrées et de sorties pendant la période d'observation. Ce faisant, on pourra regrouper les mêmes tranches d'âges et calculer au tableau 4 les quotients perspectifs quinquennaux. Ces derniers peuvent être directement comparés avec les quotients identiques calculés pour les hommes de la province de Namur pendant les années **1991-1993**. Malgré la ténuité de certains chiffres, on peut observer qu'à une exception près, la mortalité des prêtres est partout moins élevée. Il ne pourrait être question de généraliser ce constat, mais il renforce tout l'intérêt que revêt une étude de la mortalité pour des groupes bien déterminés de personnes.

TABLEAU 2. LE MOUVEMENT DE LA POPULATION DU CLERGÉ NAMUROIS PAR GROUPE D'ÂGES ENTRE 1974 ET 1994 (LES 65 ANS ET PLUS SONT EN GRISÉ).

Groupes de générations	Population initiale	Nouveaux prêtres	Décès	Entrées	Sorties	Population finale
1879-83	4	0	4	0	0	0
1884-88	10	0	10	0	0	0
1889-93	14	0	14	0	0	0
1894-98	50	0	50	1	0	1
1899-1903	60	0	57	0	0	3
1904-08	80	0	68	0	0	12
1909-13	131	0	79	0	2	50
1914-18	97	0	50	3	3	47
1919-23	128	0	56	1	2	71
1924-28	119	0	22	4	3	98
1929-33	120	0	9	1	4	108
1934-38	83	2	5	2	6	76
1939-43	73	0	5	3	19	52
1944-48	50	13	1	2	12	52
1949-53	0	30	0	3	1	32
1954-58	0	17	0	2	2	17
1959-63	0	12	0	1	1	12
1964-68	0	4	0	0	0	4
TOTAL	1019	78	430	23	55	635

TABLEAU 3. LA MORTALITÉ DU CLERGÉ NAMUROIS COMPARÉE À CELLE DES HOMMES DE LA PROVINCE DE NAMUR (INS, TABLE RELATIVE AUX ANNÉES 1991-1993) - QUOTIENTS PERSPECTIFS SUR CINQ ANNÉES (1974-1994).

Groupes d'âges		Décès	Population au début	Quotient perspectif	Quotient Province. Namur	Rapport
Début	Fin					
90-95	95-100	8	11	0,727	0,834	0,872
85-90	90-95	33	43	0,767	0,701	1,095
80-85	85-90	54	99	0,545	0,579	0,942
75-80	80-85	73	208	0,351	0,428	0,821
70-75	75-80	73	278	0,263	0,289	0,908
65-70	70-75	64	353	0,181	0,201	0,901
60-65	65-70	50	420	0,119	0,124	0,959
55-60	60-65	34	434	0,078	0,082	0,958
50-55	55-60	16	428	0,037	0,049	0,771
45-50	50-55	10	365	0,027	0,033	0,823
40-45	45-50	5	310	0,016	0,022	0,740

FIGURE 1. LA PYRAMIDE DES ÂGES DU CLERGÉ NAMUROIS EN 1974 ET 1994

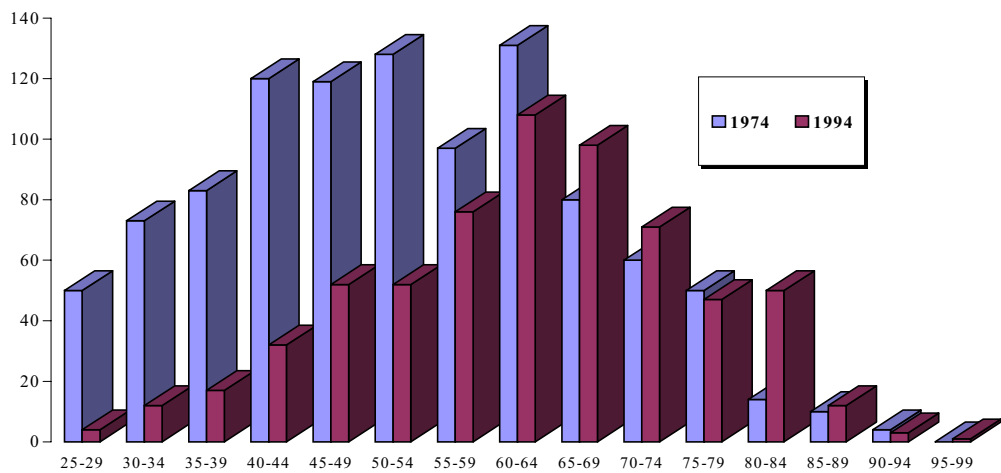
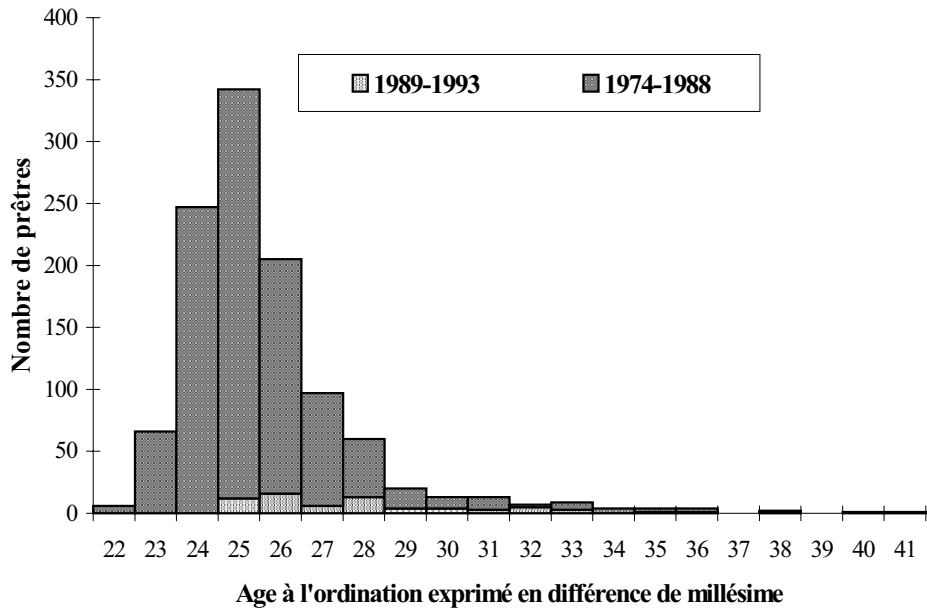


FIGURE 2. DISTRIBUTION PAR ÂGE DES NOUVEAUX PRÊTRES À LEUR ORDINATION



Le statut des prêtres permet également une analyse intéressante et, tout particulièrement, la distinction entre les prêtres actifs et les pensionnés. Le tableau 4 détaille l'évolution du nombre de prêtres de 65 ans et plus selon leur statut. Il présente en dernière colonne, à titre

d'indicateur qui s'avérera utile dans la phase prospective, le rapport entre le nombre de prêtres pensionnés (y compris ceux de moins de 65 ans) et l'effectif des prêtres de 65 ans et plus. L'extrapolation de cet indicateur permettra de proposer une estimation de l'évolution du nombre de prêtres actifs ou pensionnés.

TABLEAU 4. ÉVOLUTION DU NOMBRE DE PRÊTRES DE PLUS DE 65 ANS SELON LEUR DIFFÉRENTS STATUTS. (la dernière colonne présente le rapport entre le nombre total de pensionnés et celui des 65 ans et plus)

Dates	Administratifs	Curés	Pensionnés	Total 65+	% pensionnés	Autres pensionnés (de moins de 65 ans)
1974	13	73	132	218	70,2	21
1979	21	102	135	258	60,9	22
1984	15	102	144	261	60,5	14
1989	23	106	128	257	54,1	11
1994	34	115	133	282	51,4	12

Enfin, le tableau 5 propose une répartition de tout le clergé selon son statut d'activité. Comme on peut le constater, au statu quo du nombre de pensionnés, on peut opposer la chute du nombre de curés et celle, plus vertigineuse encore, du nombre de vicaires.

TABLEAU 5. RÉPARTITION DU CLERGÉ NAMUROIS SELON LEUR STATUT DE 1974 À 1994.

Dates	Administ.	Curés	Pro Deo	Prêtres au travail	Vicaires	Total Actifs	Pensionnés	TOTAL
1974	238	505	20	3	100	866	153	1019
1979	207	465	19	8	83	782	157	939
1984	180	399	20	6	58	663	158	821
1989	149	351	19	7	41	567	139	706
1994	136	304	18	7	25	490	145	635

Quelle prospective pour le clergé namurois ?

Afin de prévoir l'évolution numérique du clergé namurois pendant les 25 prochaines années, jusqu'à l'horizon 2019, nous appliquerons tout d'abord les quotients perspectifs de mortalité observés entre 1974 et 1994 à la structure par âge de 1994. Ces quotients ont fait l'objet d'un lissage et d'une légère réduction pour tenir compte de la régression progressive escomptée de la mortalité. Les entrées et les sorties seront supposées se neutraliser tandis que pour le nombre de nouveaux prêtres, trois scénarios distincts sont retenus :

- **scénario 1** : 10 ordinations par période quinquennale, soit la situation des cinq dernières années;
- **scénario 2** : 20 ordinations par période quinquennale ;
- **scénario 3** : 30 ordinations par période quinquennale, soit la situation des années '70.

FIGURE 3. ÉVOLUTION DE L'EFFECTIF TOTAL DU CLERGÉ NAMUROIS SELON LES TROIS SCÉNARIOS RETENUS (1994-2019).

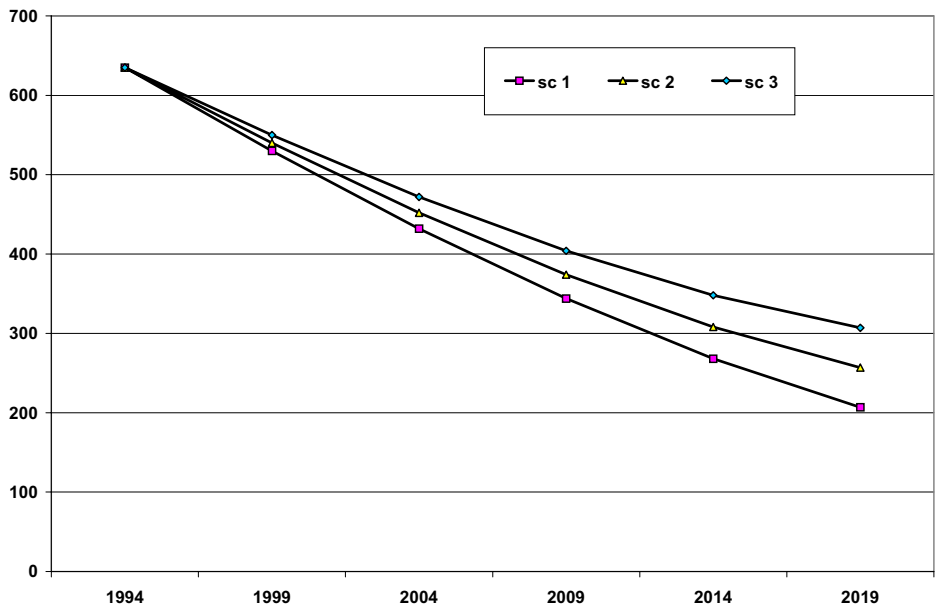
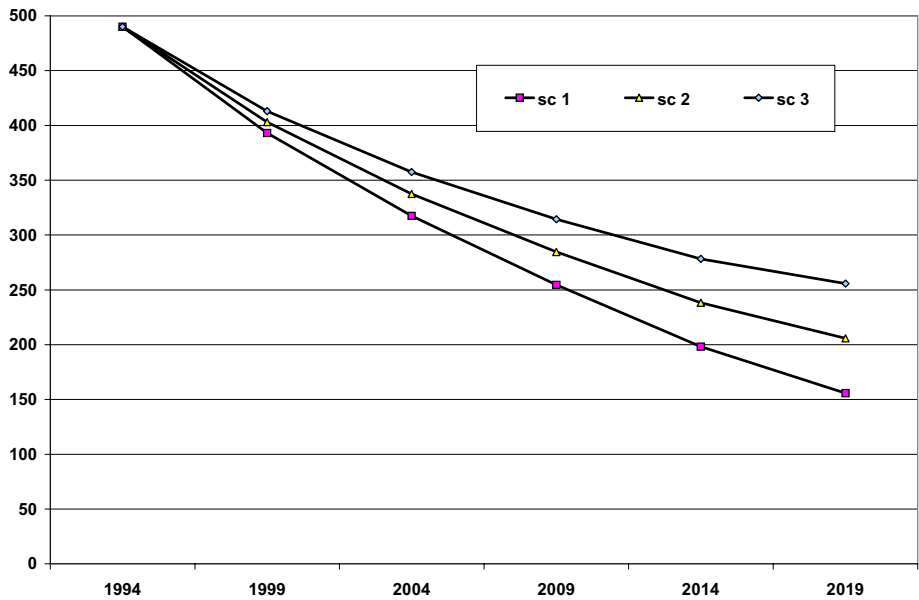


FIGURE 4. ÉVOLUTION DE L'EFFECTIF DU CLERGÉ ACTIF (EN EXCLUANT LES PENSIONNÉS) 1974 - 2019.

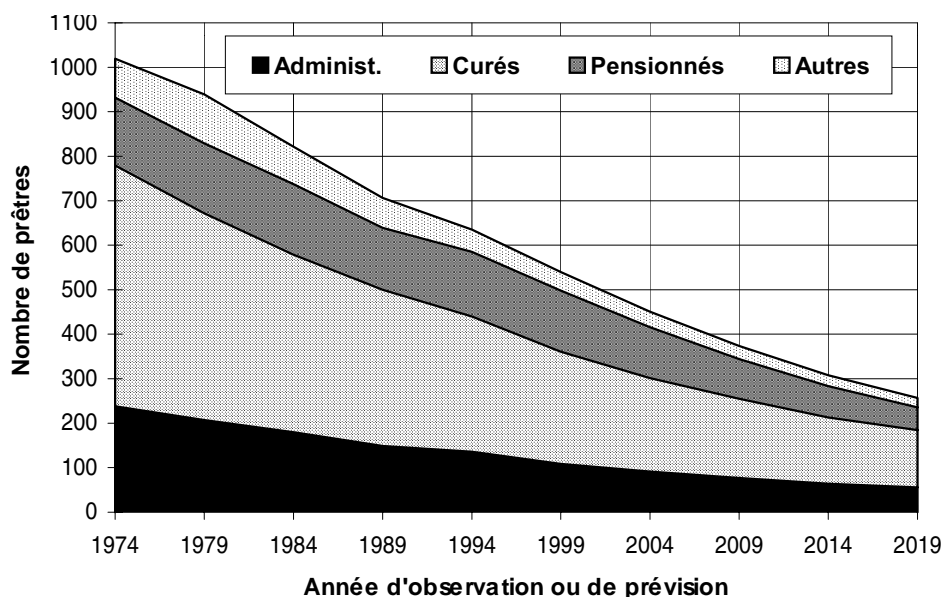


Comme le montre la figure 3, à l'horizon 2019, ces projections donnent respectivement 207, 257 et 307 prêtres au sein du clergé namurois. Commentaires sur cette évolution globale.

En estimant que le nombre de pensionnés non actifs soient successivement dans un rapport de 0,48, 0,45, 0,43, 0,41 et 0,40 avec la population âgée de plus de 65 ans, estimations ne dépendant pas du scénario choisi, on obtient, par soustraction, l'évolution du nombre de prêtres occupant un poste actif qui est proposée à la figure 4, selon les trois mêmes scénarios.

En supposant une répartition stable de ces actifs selon leur statut, c'est à dire 62,5% de curés, 27% d'administratifs, 5% de vicaires, 4% de prêtres « pro deo », et 1,5% de prêtres au travail, on peut proposer, à la figure 5, une évolution du clergé selon les quatre grandes familles de statuts, la catégorie « autre » reprenant les vicaires, prêtres « pro deo » et prêtres ouvriers.

FIGURE 5. ÉVOLUTION DE L'EFFECTIF DU CLERGÉ NAMUROIS SELON QUATRE GRANDES FAMILLES DE STATUT POUR LE CLERGÉ DE 1974 À 2019.



Les deux derniers graphiques mettent l'accent sur deux points sensibles qui résultent de ces évolutions. Tout d'abord, la figure 6 laisse apparaître l'évolution du nombre de paroisses ayant un prêtre résidant : cette proportion varierait, selon les scénarios, entre 13 et 22 % en 2019.

Enfin, la figure 7 permet de prévoir l'évolution du nombre de paroisses à charge de chaque prêtre ayant une tâche pastorale. Parmi ces derniers, nous excluons les administratifs, les prêtres « pro deo », les prêtres au travail et les vicaires qui sont attachés à un curé et à une seule paroisse. Par contre, nous avons inclus les religieux agissant au titre d'administrateur de paroisse en supposant que leur nombre variera comme suit : 60 en 1999, 50 en 2004 et 40, par la suite. Selon le scénario retenu, le prêtre de 2019 devrait prendre en charge entre 3 et 5 paroisses actuelles.

FIGURE 6. PROPORTION DU NOMBRE TOTAL DE 742 PAROISSES DISPOSANT D'UN CURÉ RÉSIDANT (EN EXCLUANT LES AUXILIAIRES) (1974-1994).

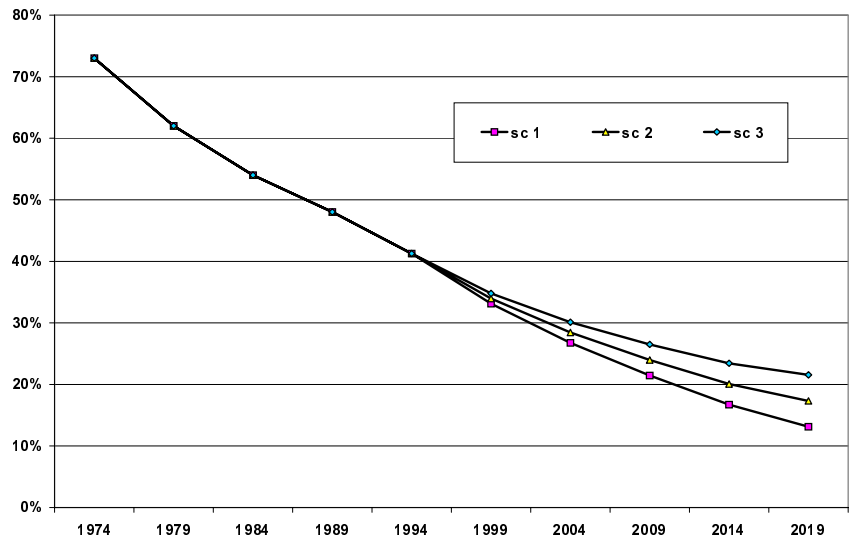
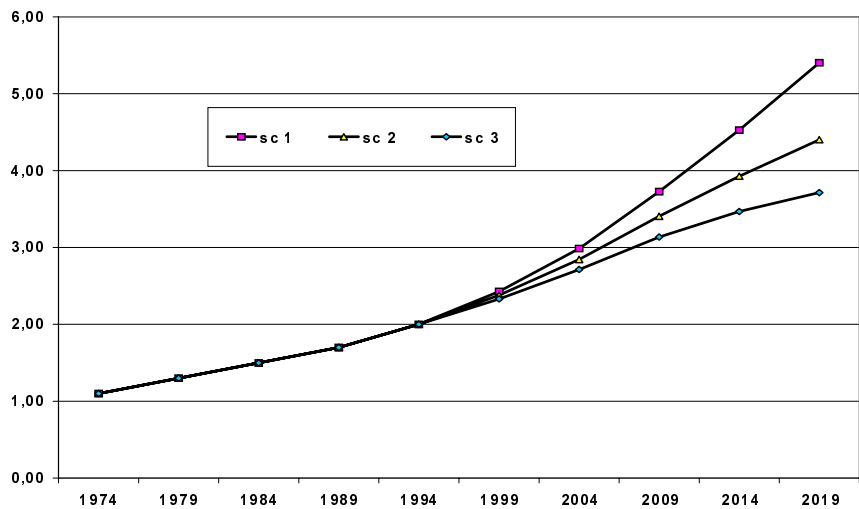


FIGURE 7. NOMBRE MOYEN DE PAROISSES PAR MEMBRE DU CLERGÉ NAMUROIS EXERÇANT DES FONCTIONS PASTORALES (CURÉ, DESSERVANT OU ADMINISTRATEUR DE PAROISSE, Y COMPRIS LES AUXILIAIRES)



En guise de conclusion

Ces chiffres alarmants sont bien entendu le fruit de simulations effectuées sur la base d'une observation précise de l'évolution du clergé namurois entre 1974 et 1994. La réalité sera vraisemblablement toute autre car face à une situation de crise, les comportements évoluent bien souvent. Néanmoins, ils ont le mérite de prévenir le risque potentiel d'une évolution à moyen terme qui serait non désirable et de servir de référence pour mettre en évidence l'effet de toute option qui serait prise afin de remédier à cette situation.

Démographie professionnelle : Opportunités et occasions de formation

Filomena RACIOPPI

Introduction

Récemment, au sein des branches en expansion de la démographie, est apparu un grand intérêt pour l'aspect « professionnel » de la démographie. Les données et les méthodes démographiques jouent un rôle important dans l'organisation et la planification de l'administration publique et dans les stratégies de marketing des compagnies.

Dans cet article, les discussions sur les « nouvelles frontières » de la démographie ne porteront pas tant sur le contenu de ces nouveaux secteurs et directions, mais sur les occasions, dans la mesure où l'Italie est intéressée, d'introduire de nouveaux aspects dans la formation démographique.

En 1985, à l'occasion de l'une des réunions les plus connues (*L'insegnamento della demografia e la formazione dei demografi in Italia* - Omaggio a Nora Federici : l'enseignement de la démographie et la formation des démographes en Italie - Hommage à Nora Federici) la démographie a été traitée en tant que science, sujet d'enseignement et profession (Sonnino et al., 1987), le but étant la définition des rôles spécifiques des différentes fonctions de cette discipline. Donc, ce ne serait pas la première fois que l'on discute d'une « démographie professionnelle » qui peut contribuer à l'expansion des opportunités d'emploi sur le marché du travail dans le secteur public ou privé.

Pour expliciter les raisons de cet article, il faudrait peut être mentionner les nombreuses positions adoptées par les démographes italiens ces dernières années, influencées notamment par les échos qui nous sont parvenus des États-Unis à la fin des années 80, sur la présence et l'utilisation des données, des techniques et des modèles démographiques dans les stratégies de l'administration publique et des affaires (Casacchia et Racioppi, 1992).

Il y a une crainte très répandue selon laquelle la démographie, en tant que discipline, peut être corrompue et réduite à un simple exercice de compétences techniques adaptées aux besoins d'un client. L'hésitation d'une partie des démographes italiens vis-à-vis de l'acceptation et l'adoption de ce qu'en Amérique on appelle « Démographie des affaires » (ou Démographie appliquée : version plus large du terme qui inclut la démographie de l'organisation et la formulation des programmes pour l'administration publique et les politiques du gouvernement) confirme et souligne la perplexité qui, jusqu'à présent, bien qu'il n'ait pas fait d'obstacle, a rendu plus difficile le plein développement de la profession de démographe, même si l'on a tracé les lignes de l'extension de la formation démographique vers les sphères professionnelles.

Depuis quelque temps déjà, tout le monde connaît l'intensité et les caractéristiques spécifiques des changements démographiques en Italie et l'on sait aussi que cette dynamique continue et ces changements démographiques structurels déterminent les politiques sociales et économiques, locales et nationales.

Les différenciations très prononcées entre les régions rendent plus cruciaux encore les aspects les plus problématiques de l'administration et de la planification locale des stratégies de production et d'investissement : des régions caractérisées par un manque de dynamisme démographique, dépeuplées ou âgées - où, d'une part, les conditions ne permettent pas d'investir et où, d'autre part, il est nécessaire d'intervenir immédiatement pour prévenir la disparition des communautés entières - peuvent côtoyer des régions très peuplées, caractérisées par des

conditions difficiles, où la planification et l'organisation sociale, économique et de santé ne sont pas adaptées aux besoins de la population (Golini et Mussino, 1987 ; Soliani, 1992).

En effet, les changements démographiques engendrent des changements dans la consommation, les économies, les investissements individuels ou collectifs, les rôles sociaux etc. et, conséquence directe de cette situation, il devrait y avoir une demande importante d'informations de la part des dirigeants dans n'importe quel domaine directement ou indirectement apparenté.

En même temps, la démographie académique doit développer un secteur consacré à la formation et à la recherche qui permettra aux structures locales et aux entrepreneurs locaux de manier les données et d'interpréter et analyser ces données, pour la planification, l'investissement etc. Par conséquent, la démographie académique doit créer une certaine culture et sensibilité démographiques qu'il faut prendre en compte durant la planification stratégique, sociale et territoriale ou l'élaboration des politiques d'affaires. En fait, il ne suffit pas de fournir les outils et les compétences : le concept « point de vue démographique » doit être encouragé et propagé. Cela doit être le point de vue à partir duquel la prise de décision et le processus organisationnel relatif à l'administration publique ou au secteur privé doivent commencer.

1. L'expérience américaine et l'Italie : affinités et différences

Cet article n'a pas l'intention de reconstruire les étapes qui, durant les vingt dernières années aux États-Unis, ont conduit au développement de nouveaux secteurs de l'étude démographique : la Démographie locale et la Démographie des affaires. Pour les références bibliographiques, il faut se reporter à celles mentionnées dans l'ouvrage de Casacchia et Racioppi.

Le marché et le monde des affaires étant devenus de plus en plus complexes et compétitifs en Amérique, la demande considérable de compétences démographiques de la part des entreprises et des dirigeants publics a amené les spécialistes (en principe, des sociologues et des consultants de marketing) à examiner de nouveaux domaines d'étude avec, systématiquement, la participation des structures académiques et administratives et des entreprises.

En général, la plupart des hommes d'affaires qui utilisent les données démographiques n'ont pas une formation universitaire en démographie, mais ils ont étudié l'économie, les affaires etc. De plus, dans les années 80, la présence des démographes étaient peu commune dans le secteur privé.

Aux États-Unis, en tout cas, malgré le fait que les trois contextes de travail des démographes - public, privé et universitaire - sont tous assez bien représentés (Kintner et Swanson, 1987), beaucoup doit être fait afin que chacun à sa place accepte ses collègues travaillant dans les autres secteurs et surtout, que l'on reconnaisse que « le travail démographique réalisé par le secteur privé n'est pas nécessairement inférieur à celui réalisé par les universités et les agences gouvernementales » (Russel, 1984). Cette situation, d'après les Américains, est le résultat de l'absence de formation démographique spécialisée qui s'adresse à ceux qui réalisent ces activités dans le domaine des affaires, puisque seule la formation académique donne aux étudiants cette sensibilité démographique qui, autrement, exigerait des années d'expérience de travail.

L'idée que l'on puisse gagner sa vie sans avoir besoin de la démographie a constitué le sujet d'un important débat parmi les universitaires et a été souvent mise au pilori pendant la longue période qui a précédé l'élimination du conflit entre les intérêts professionnels et la recherche académique (les universitaires étant depuis longtemps, contre la tendance vers une démographie appliquée).

Il est pourtant important que les universités reconnaissent le droit aux étudiants de prendre en considération les postes des deux secteurs : public et privé (De Jong, 1987). Il doit y avoir une reconnaissance réciproque entre ces catégories professionnelles.

En Amérique, en 1984, sur 88 collèges ayant des programmes démographiques, 40 offrent des cours de Démographie Appliquée et des Affaires et sur ces 40, 16 ont des cursus universitaires sur ce sujet. La différence entre ces cursus et les cursus traditionnels réside dans le fait que l'on insiste sur des applications, des domaines d'activité spécifiques qui peuvent au départ, constituer un défi pour les démographes.

En général, aux États-Unis, il existe une demande de démographes de très haut niveau, capables de travailler en équipe et ayant des compétences dans les domaines des études économiques, géographiques et d'investissement et qui montrent des aptitudes pour le traitement informatique des données, l'organisation, le management et la communication.

Aux États-Unis, l'orientation vers l'étude de cas est devenue écrasante. A partir du slogan « aucun manuel ne peut se substituer à l'expérience concrète », l'Amérique a produit beaucoup de matériel de référence, qui peut aussi être utilisé pour la formation. Un exemple parmi d'autres est le guide publié par Kintern Merrik, Morrison et Voss (1994) qui inclut nombre d'études de cas dans plusieurs domaines (politiques et loi, marchés, emplacements, planification et politique générale).

Selon ces auteurs, pareil exposé était nécessaire, puisque, généralement, l'expérience des démographes académiques en démographie appliquée est limitée et ceux qui travaillent dans ce domaine n'ont ni le temps ni la motivation nécessaires pour faire connaître leur propre expérience. Le livre est un guide dans le vrai sens du mot et un supplément d'un manuel qui peut proposer de nouvelles méthodes d'enseignement. Quelques domaines d'application des études démographiques apparaissent dans le sommaire :

- évaluation de la rentabilité de quelques affaires proposées ;
- évaluation des aires de marché avec des données démographiques (par exemple le marché des cellulaires) ;
- planification pour la garde des enfants ou les services de santé ;
- planification financière stratégique pour les hôpitaux ;
- prévision démographique pour les sous-groupes, par exemple les chômeurs et les étudiants.

Tous ces matériaux peuvent être objet d'enseignement, pour faire de l'étudiant un des acteurs dans la résolution des problèmes du gouvernement ou du monde des affaires (ce qui est le cas, par exemple, dans certains cours de Démographie des affaires des universités américaines)

En comparaison avec les États-Unis, la situation en Italie, a des inconvénients. Aux États-Unis, surtout parmi les sociologues et, plus tard, parmi les gens de marketing, les managers etc., a commencé une campagne qui souligne l'importance des indicateurs démographiques. Traditionnellement, aux États-Unis, un démographe n'est pas seulement un chercheur mais plutôt quelqu'un qui peut se trouver dans les compagnies, les agences gouvernementales, l'administration locale et les universités. Ce qui est sûr, c'est que la croisade n'a pas été impulsée par les démographes. On peut conclure que ce n'est pas l'offre des démographes mais la demande qui a pris l'initiative de l'orientation des compétences vers des buts pratiques.

En Italie, ni les sociologues ni les utilisateurs des indicateurs démographiques n'ont fait de pareils pas, l'une des raisons étant les différences évidentes entre les sociologues italiens et anglo-américains. Au lieu de cela, un certain nombre de démographes ont anticipé les problèmes concernant ce nouveau secteur de la démographie. Pourtant, les positions opposées ont empêché la reconnaissance complète de ces nouvelles approches, ou l'acceptation de la formation démographique offerte est insuffisante pour rendre professionnelles ces activités.

Que peut-on faire en Italie pour élargir les perspectives des étudiants, pour qu'ils puissent pénétrer le marché du travail en tant que techniciens et professionnels dans le cadre de leur propre domaine et non seulement en tant que chercheurs ou programmeurs ? Peut-être les utilisateurs doivent-ils être sensibilisés et la demande stimulée, peut-être la formation professionnelle doit-elle être offerte à des non-démographes, en tout cas, le rôle « professionnel » de la démographie doit être créé et, beaucoup pensent que, presque naturellement, il doit être réservé aux démographes ; ce rôle doit être défendu avant que d'autres ne le prennent.

Il faut souligner qu'en Italie, les termes de cette question sont différents de ceux qui, aux États-Unis, ont favorisé le développement d'une nouvelle profession pour les démographes : il ne s'agit pas de former des non-démographes en vue qu'ils acquièrent la soi-disant « perspective démographique » (Pol, 1986) - c'est-à-dire de munir du « point de vue/angle démographique », les opérateurs de marketing, les sociologues les entrepreneurs et les économistes des compagnies. Il s'agit plutôt de réviser la formation et le curriculum démographique de sorte que les démographes puissent opérer professionnellement dans les domaines économiques, du management et de la planification des services et, pour qu'ils puissent, par voie de conséquence, dialoguer réellement avec les entrepreneurs et les chefs de l'administration publique et participer activement à l'organisation et à la planification.

Autrement dit, bien que les économistes, les dirigeants du secteur public et les entrepreneurs américains aient commencé, il y a déjà 20 ans, à exprimer le besoin d'indicateurs démographiques, en Italie, après un retard considérable, les démographes universitaires ont exprimé, d'une part, le besoin et, d'autre part, les avantages de la formation directe en vue d'obtenir de nouvelles opportunités professionnelles.

De plus, en Italie aussi, le secteur privé de l'information a connu récemment un développement important : les compagnies qui fournissent des bases de données à d'autres compagnies (et à la recherche) créent sur commande selon l'expérience américaine et en compétition évidente avec les institutions publiques et les agences du domaine. Par exemple, depuis quelques temps, Sarin a vendu un certain nombre de produits très faciles à utiliser pour les institutions de recherche, les opérateurs du secteur privé, l'administration locale ou les services publics, étant donné les détails territoriaux considérables fournis par ce genre d'informations. Ce processus de marketing et de distribution des données stimule considérablement la propagation de la démographie professionnelle, puisque les compagnies privées qui produisent de l'information sont généralement plus flexibles vis-à-vis des exigences des clients et peuvent satisfaire plus rapidement ces besoins. C'est l'une des opportunités du développement de la « conscience démographique » du côté des compagnies et des structures publiques. Par conséquent, on peut conclure qu'un marché des données statistico-démographiques se développe, même si les problèmes sous-jacents diffèrent :

- les utilisateurs ne fournissent pas toujours des explications correctes sur leur demande d'information et, par conséquent, ceux qui la produisent fondent leur travail sur les besoins perçus, besoins de plus en plus variés et qui créent, en tout cas, le besoin de flexibilité (Biggeri, 1994) ;
- souvent, les utilisateurs des données démographiques et, plus généralement, des données statistiques ne sont pas sensibilisés au point de vue démographique, ce qui veut dire qu'ils sont, non seulement, incapables de réaliser des analyses et interprétations indépendantes des tendances de la population à moyen et à long terme (de sorte que pareilles informations puissent être exploitées aux moyens de la politique de la compagnie), mais aussi d'examiner les événements par des stratégies adéquates.

2. La démographie universitaire en Italie

Puisqu'il y a certains signes d'intérêts vis-à-vis de ce problème dans les universités italiennes, il nous semble approprié de faire une courte description de la démographie universitaire. Au début des années 90, ce type d'activités n'est pas très bien distinct, étant donné qu'il existe qu'un seul Département de démographie à l'Université de Rome « La Sapienza » (l'ancien Institut de Démographie) et que cette discipline est pratiquement gérée en exclusivité par le Département de Statistique (par opposition aux États-Unis, où la démographie est diffusée notamment dans les Départements de Sociologie).

Dans les universités d'Italie, sur plus de 80 cours de démographie, 67% (55 sur 82) concernent la démographie institutionnelle (Maffioli, 1994 ; année universitaire de référence 1990 - 1991). Les autres s'occupent de la méthodologie et du contenu, conformément aux besoins actuels d'informations : l'analyse démographique, les théories de la population et modèles démographiques, les enquêtes de démographie, la démographie sociale, la démographie économique et historique sont les bases culturelles utilisables pour créer un curriculum concernant les études de la population. A part les cursus (et diplôme) de sciences statistiques (démographique, actuariel et économique), il y a quelques cours de démographie en sciences politiques et économiques et rarement en droit, médecine et sciences humaines.

Enfin de compte, seuls les cursus de statistiques et démographie (les cursus institutionnels étant complétés par au moins deux cursus de démographie supérieure) délivrent un programme de formation à la démographie bien défini.

Nous n'avons pas l'intention de fournir ici une évaluation partielle ou globale de l'enseignement de la démographie dans les universités (voir, par exemple, Pinnelli, 1987), de l'efficacité et des résultats finaux de cet enseignement ou de la cohérence avec le contexte de la formation dans lequel ce type de cours a lieu ou l'adéquation de l'enseignement aux compétences visées. Pourtant, il faut mentionner certains défauts structuraux :

- beaucoup signalent l'absence d'« l'harmonisation » des cours de démographie universitaires et des autres caractéristiques de ces cours malgré la tendance à développer l'approche interdisciplinaire de la démographie : « une science dont le contenu est dispersé entre différentes disciplines » (Livi Bacci, 1994) ; évidemment, il s'agit d'un problème moins présent dans les facultés de statistiques que dans les autres ;
- l'habitude de donner à des cours différents le même nom peut aussi cacher des différences considérables en ce qui concerne les contenus de ces cours, d'un lieu à un autre. Un exemple est « la démographie économique » où les uns donnent une approche économique et les autres une approche démographique aux problèmes et qui englobe l'étude du marché du travail, des ménages et de la consommation, du développement et des ressources et des modèles pour les phénomènes démo-économiques (Maffioli, 1994), mais, sans jamais fournir une analyse complète du problème complexe des relations entre l'économie et la démographie,
- le traitement non-systématique des aspects concernant l'information est apparu dans les cours de démographie ou de statistique : le recueil des données, l'exploitation et la qualité des données démographiques, les aptitudes concernant le traitement des informations en vue de construire des indicateurs sont considérés comme des problèmes très simples, d'importance secondaire ou qui, de par leur nature même, sont considérés comme faisant partie des cours traitant des sujets de l'information (et, par conséquent, sont supposés comme allant de soi). Comme Maffioli le note dans l'ouvrage cité ci-dessus, cette situation peut avoir des conséquences sérieuses d'un point de vue professionnel, sur la création de nouvelles opportunités de travail. A lieu de cela, un nombre de plus en plus grand de personnes considèrent que la tâche du recueil et de la sélection des informations peut être déléguée à d'autres ou que les statistiques officielles sont toujours suffisantes.

Dans les deux ouvrages cités ci-dessus concernant la situation de la démographie universitaire (Pinnelli 1987, Maffioli 1994), il faut noter que le côté professionnel de la formation démographique ne peut pas être facilement identifié ni exploité en conséquence. De plus, malgré l'utilisation de plus en plus fréquente des compétences démographiques dans la planification et le marketing, le standing professionnel des démographes doit être soutenu comme moyen de répondre à la demande de développement de considérations démographiques dans les politiques, la formation et la production de biens et de services. Beaucoup de démographes ont signalé que ce type de demandes existe depuis quelque temps.

3. Une recherche préliminaire

Comment les démographes italiens voient-ils le problème et quelle est leur expérience professionnelle ?

Nous avons réalisé une recherche en utilisant un questionnaire auprès des professeurs de sujets démographiques sur l'éventuelle expérience professionnelle acquise et le type de cours qu'ils donnent. Le but était aussi de voir ce qu'ils pensaient de l'élaboration de plans de formation orientés vers la démographie appliquée. On leur a également demandé leur avis concernant l'efficacité professionnelle de la formation démographique universitaire et sur les thèmes à étudier pendant les cours de démographie qui devraient être inclus dans le cursus pour les démographes professionnels.

Le questionnaire présenté en annexe, est un premier moyen pour essayer de déterminer quelle est l'attention accordée par le monde académique aux profils professionnels. En fait, il s'agit d'un questionnaire très court, envoyé à 63 professeurs de démographie du premier et du second niveau ; 86% des personnes interrogées ont répondu (54 questionnaires ont été remplis et rendus).

Étant donné le faible nombre de cas, nous allons essayer de fournir une description aussi complète que possible concernant les données obtenues sans fournir pour autant, une interprétation statistique sophistiquée des informations disponibles.

En ce qui concerne l'incidence des activités professionnelles parmi les démographes italiens, on peut noter que la majorité des démographes ont déclaré qu'ils fournissent ou ont fourni occasionnellement des services de consultants dans le secteur public. Pour 15% des personnes interrogées, les activités en question concernent à la fois le secteur public et privé et pour 45% le seul secteur public. Dans 68% des cas, les activités professionnelles ont été occasionnelles.

Douze personnes interrogées ont déclaré qu'il ne s'agissait pas d'un travail de consultant : ils parlent de journalisme et de collaboration avec la RAI - le poste national italien de radio et de télévision - pour des activités de vulgarisation, de recherche pour la coopération internationale, de participation aux comités ISTAT de recensement et d'enquêtes spéciales, de collaboration avec les autorités municipales, régionales et de province pour la planification des services sociaux, l'enregistrements des émigrants, des plans qui résultent du dépeuplement ou du vieillissement de certaines municipalités des régions montagneuses, des tendances démographiques locales (tableau 1).

TABLEAU 1 : RÉPONSES CONCERNANT LES ACTIVITÉS PROFESSIONNELLES (POURCENTAGES ET EFFECTIFS)

Distribution des enseignants selon les activités professionnelles, secteur et fréquence					
Activité professionnelle		Secteur d'activité		Fréquence de l'activité	
Oui	79 (43)	Public	54 (29)	Occasionnel	68 (37)
		Privé	7 (4)	Fréquemment	6 (3)
Non	20 (11)	Les deux	15 (8)		
		Sans réponse	4 (2)	Sans réponse	6 (3)
		Sans objet	20 (11)	Sans objet	20 (11)
Total	100 (54)	Total	100 (54)	Total	100 (54)
Types d'activités					
Activité de consultant			Autres activités		
Oui	60 (32)		Oui	22 (12)	
Non	11 (6)		Non	52 (28)	
Sans réponse	9 (5)		Sans réponse	6 (3)	
Sans objet	20 (11)		Sans objet	20 (11)	
Total	100 (54)		Total	100 (54)	

54% des enseignants ont répondu que certaines des leçons de leurs cours comportaient une orientation de démographie appliquée, 44% ont répondu par la négative (tableau 2). Quelques-uns ont noté qu'ils utilisent parfois des prévisions dérivées ou des prévisions pour les petites aires, et dans un seul cas nous avons découvert que les cours étaient complètement basés sur l'utilisation de sources et le traitement de cas concrets, dans le but explicite d'offrir une qualification professionnelle aux étudiants. Comme on s'y attendait, à l'exception de trois enseignants¹ - personne n'était informé sur les cours démographiques pour le secteur public ou privé en Italie (Tableau 2)

TABLEAU 2 : SUJETS DIDACTIQUES (POURCENTAGES ET EFFECTIFS)

Buts professionnels des cours		Connaissances sur les affaires ou les cours de démographie appliquée	
Oui	54 (29)	Oui	6 (3)
Non	44 (24)	Non	94 (51)
Sans réponse	2 (1)		
Total	100 (54)	Total	100 (54)

La formation démographiques obtenue à la fin des études est considérée comme, au moins partiellement, adéquate pour accomplir des activités professionnelles extra-universitaires par 29% des personnes interrogées (16 sur 54). Seuls 4 enseignants ont répondu nettement par la négative à la questions 6 concernant l'adéquation des cursus universitaires à une formation

¹ Il s'agit de :

- a) cours pour les opérateurs statistiques régionaux réalisés par le Département d'étude des sociétés méditerranéennes à la Faculté de droit et de sciences politiques, Bari.
- b) Cours de démographie à l'Université LUISS de Rome.

en démographie appliquée. Ils affirment qu'aucun des cursus examinés n'accomplit cette tâche ; mais plusieurs personnes reconnaissent que cela dépend de l'expérience didactique personnelle. D'autre part, la démographie et les statistiques ont été indiquées par 83% des personnes interrogées, les statistiques et l'économie par 52%, les sciences statistiques actuarielles par 41% et les statistiques et l'économie par respectivement 41% et 39% (tableau 3). Dans un seul cas, les nouveaux diplômés universitaires ont été inclus dans la rubrique « autres », et les cours de droit dans une autre.

TABLEAU 3 : FORMATION PROFESSIONNELLE (POURCENTAGES ET EFFECTIFS)

Utilité professionnelle de la formation démographique reçue		Cursus qui prévoient une formation professionnelle adéquate (nombre des indications).	
Suffisante	20 (11)	Statistiques	41 (22)
Presque suffisante	9 (5)	Statistiques et démographie	83 (45)
Insuffisante	67 (36)	Statistiques et économie	52 (28)
Je ne sais pas	4 (2)	Sciences statistiques actuarielles	41 (22)
		Économies	39 (21)
Total	100 (54)	Sciences politiques	17 (9)
		Autres	4 (2)

Enfin, en ce qui concerne les sujets possibles à traiter en profondeur en démographie en vue de créer un curriculum professionnel, les démographes indiquent (tableau 4) : les prévisions (pour les petites aires, les ménages, les prévisions dérivées) ; l'analyse des territoires et des petites aires (avec plus d'attention à porter à la démographie urbaine et régionale et à la démographie liée à l'organisation du territoire) ; les problèmes concernant le recueil et l'exploitation des informations (avec mention faite des pays où les statistiques sont insatisfaisantes, du contrôle de la qualité des données, de l'analyse critique des données, attention portée à l'exploitation des données administratives, entièrement consacrée à la création d'un système d'informations démo-socio-économiques, à la stimulation de la capacité de reconstruire l'information et de créer des indicateurs qui utilisent des ensembles de données) ; les modèles et les méthodes (technique de standardisation et de reconstruction et analyse par cohortes à appliquer aux problèmes du monde des affaires), l'analyse des problèmes (vieillesse, systèmes de santé, politiques familiales pour les structures administratives locales, l'impact de la croissance de la population sur l'environnement, les conséquences des changements démographiques sur les champs économique, social et de la santé). Certains ont indiqué également la démographie économique et les pays en voie de développement. Parfois, on a souligné l'importance de l'analyse des études de cas dans les cours - des cas qui peuvent fournir des exemples d'application dans les secteurs macro-économiques et des entreprises (un exemple de la convergence entre les sphères universitaires et du travail).

Il faut aussi noter clairement que toutes les réponses ont été mises en rapport avec l'expérience personnelle des personnes interrogées.

TABLEAU 4 : SUJETS À APPROFONDIR VISANT UN CURRICULUM
POUR LES « DÉMOGRAPHES PROFESSIONNELS »

Sujets	% et nombres d'indications	
Prévisions	55	(30)
Recueil de données et management	33	(18)
Modèles et méthodes	30	(16)
Problèmes démographiques	24	(13)
Pays en voie de développement	6	(3)
Rapports avec les autres disciplines	7	(4)
Software	4	(2)
Démographie économique et études de cas	17	(9)
Analyse des territoires et des petites surfaces	28	(15)

Dans le tableau 5, on trouve la distribution des enseignants interrogés selon les deux variables incluses en vue d'exercer un contrôle minimal : la licence obtenue et l'année de l'obtention. Contrairement à ce que l'on aurait pu attendre, on ne peut pas considérer que ceux qui ont obtenu leur licence il y a quelques dizaines d'années font preuve d'un esprit moins large face à ces problèmes : il n'y a pas de différences significatives entre les réponses selon la période d'études des personnes interrogées (tableau 6).

En effet - en dehors du fait, plus ou moins attendu, que l'expérience professionnelle était plus présente chez ceux qui avaient fait leurs études universitaires il y a plusieurs années - les efforts pour intégrer les problèmes opérationnels dans les cours ne sont pas le fait ni des plus âgés ni des plus jeunes enseignants. Par rapport à leurs homologues plus âgés, les membres jeunes de l'équipe déclarent plus fréquemment que la formation démographique qu'ils ont reçue pendant leurs cours universitaires n'était pas adaptée à une activité de démographie appliquée (tableau 7).

TABLEAU 5 : LICENCE/DIPLÔME ET ANNÉE DE L'OBTENTION

Licence obtenue			Année	
Statistiques et démographie	31	(17)	avant 1959	9 (5)
Sciences statistiques actuarielles	7	(4)	1960 - 1965	37 (20)
Statistiques et économie	4	(2)	1966 - 1970	28 (15)
Économies	48	(26)	1971 - 1975	20 (11)
Sciences naturelles	2	(1)	1976 et après	6 (3)
Sciences politiques	6	(3)		
Droit	2	(1)		
Total	100	(54)	Total	100 (54)

TABLEAU 6 : ACTIVITÉ PROFESSIONNELLE ET BUTS DES LEÇONS, PAR DIPLÔME ET ANNÉE DE L'OBTENTION (EFFECTIFS)

Année de l'obtention du diplôme	Activité professionnelle		Buts professionnels des cours	
	Oui	Non	Oui	Non
avant 1959	5	-	4	1
1960 - 1965	17	3	10	10
1966 - 1970	11	4	7	8
1971 - 1975	8	3	6	5
1976 et après	2	1	2	1
Diplôme obtenu				
Statistiques et démographie	13	4	9	8
Économies	20	6	13	13
Sciences statistiques actuarielles	5	1	4	2
Sciences politiques	3	-	1	2
Sciences naturelles	1	-	1	-
Droit	1	-	1	-
Total	43	11	29	25

TABLEAU 7 : UTILITÉ DE LA FORMATION DÉMOGRAPHIQUE ACQUISE, PAR DIPLÔME ET ANNÉE DE L'OBTENTION

Année de l'obtention du diplôme	Utilité professionnelle de la formation démographique acquise			
	Suffisant	Presque suffisant	Insuffisant	Je ne sais pas
avant 1959	1	-	4	-
1960 - 1965	4	2	13	1
1966 - 1970	4	1	5	1
1971 - 1975	2	1	8	-
1976 et après	-	1	2	-
Diplôme obtenu				
Statistiques et démographie	5	3	8	1
Économies	4	-	21	1
Sciences statistiques actuarielles	2	1	3	-
Sciences politiques	-	-	3	-
Sciences naturelles	-	1	-	-
Droit	-	-	1	-
Total	11	5	36	2

Il peut paraître curieux d'apprendre que les membres des générations de titulaires de diplômes où l'enseignement de la démographie ne prenait pas en considération les aspects professionnels, déclarent qu'ils ont reçu une formation adéquate de ce point de vue professionnel. Il est vrai qu'avec le temps, lorsque l'expérience se substitue à la formation élémentaire suivie, certains jugements sur les limites de l'enseignement reçu deviennent moins précis. En dehors de ce motif, il ne faut pas oublier l'importance de la perception individuelle concernant les aptitudes personnelles.

4. Derrière le questionnaire

Cette recherche n'est certainement pas exhaustive et cela nous autorise, dans un certain sens, à aller au-delà d'une simple interprétation des réponses et à exploiter l'échange instructif d'opinions qui suit parfois le questionnaire afin d'avoir une idée plus claire des implications des affirmations avancées.

Le premier point important est que, au moins du point de vue universitaire, il n'existe pas une définition reconnue et acceptée de tous du terme « démographie professionnelle ».

Qu'est-ce qu'un démographe professionnel ? Après tout, le démographe universitaire, en tant que chercheur et professeur, lors de l'application des compétences de démographie appliquée qu'il a acquises, ne peut-il pas se considérer comme un professionnel dans le domaine de la démographie ?

Évidemment, il ne s'agit pas de cela. Il faut plutôt reconnaître que depuis la réunion de 1985, dédiée à Nora Federici, on a évolué vers l'opinion que le côté professionnel concerne seulement la démographie appliquée par les aspects techniques et opérationnels pouvant être mieux orientés vers les utilisateurs, dans les champs du management, de l'organisation et de la planification de l'administration publique, ou de l'investissement, de la planification et des activités financières du secteur privé. De plus, certains ont l'impression que le type de démographie dont on s'occupe ici a été relégué à une position secondaire, situation qui ressemble au débat initié aux États-Unis où plusieurs, influencés par les circonstances, ont affirmé que le travail démographique a la même valeur quel que soit le domaine d'application, tel que nous l'avons souligné dans la section 2.

Il paraît qu'il existe une opinion très répandue selon laquelle la démographie professionnelle n'existe pas, et qu'il n'y a aucune chance pour que la démographie des affaires puisse exister en tant que discipline. Il en résulte que les utilisateurs qui font la demande d'indicateurs démographiques peuvent se trouver facilement dans une situation où eux-mêmes se familiarisent avec les techniques et les problèmes démographiques, ou bien ils résolvent tout seuls les problèmes ou demandent seulement aux démographes des consultations occasionnelles (et cela se passe déjà) sans faire appel aux services d'une équipe de démographes. Comme notre enquête l'a révélé, il s'agit de ceux qui croient que les aspects opérationnels sont une partie de cette discipline et que, normalement, un cours de démographie, tel quel, vise déjà la formation professionnelle des étudiants.

Donc, du point de vue de la formation acquise, quand est-on capable d'agir en tant que consultant ? A la fin des études ? C'est un point discutable, étant donné que l'expérience nous dit que la fréquence des demandes « d'intervention démographique » est en proportion directe avec le dossier de recherche (c'est-à-dire que seuls les professeurs reçoivent pareilles demandes).

Donc, les étudiants qui acquièrent des compétences démographiques et statistiques, doivent-ils attendre de devenir professeurs pour pouvoir exercer ces compétences dans un champ non-universitaire ?

Si la carrière académique représente la seule possibilité de pratiquer le métier de démographe au service des institutions, ce type de recherche et d'enseignement n'est pas une fin en soi, il s'agirait plutôt d'une contraction et non pas d'une expansion du marché du travail. De toute façon, il est vrai que les consultations sollicitées ne demandent jamais de compétences « purement démographiques ».

Un autre problème que l'on peut mentionner est, qu'en Italie, les statisticiens non plus ne bénéficient pas d'un statut professionnel spécifique dans l'univers de ce type de travail et, d'une certaine façon, cela conditionne l'avenir des démographes. Aujourd'hui même, il n'existe pas un accord général concernant la création d'un registre ou ordre professionnel qui

permettrait aux intervenants d'acquérir des compétences spécifiques et de recevoir des rémunérations conformément à ces compétences. Pourtant, il y a des personnes qui croient que cette situation peut entraîner une marginalisation ou une classification trop simplifiée des services destinés au marché.

Si la démographie professionnelle n'existe pas, il faut rappeler qu'il y a des gens qui reconnaissent, avec réserve, qu'il faut la créer. Comment ? Il faut accorder beaucoup d'attention au contenu. Cette profession ne doit pas être rétrogradée au rôle de démographie professionnelle, de sorte qu'elle ne finisse par être incorporée dans les statistiques appliquées et que les aspects méthodologiques puissent être protégés.

La démographie professionnelle peut se développer à partir de la convergence de la demande, de la part des utilisateurs, et de l'offre, de la part des démographes (Racioppi, 1994). Mais ce serait une erreur d'accorder une attention trop grande à la demande et trop peu à l'offre : ceux qui sont du côté de la demande, souvent, ne savent pas comment articuler leurs propres besoins ou peuvent croire qu'ils sont capables de diriger seuls leurs affaires. Le problème est que la définition claire de l'offre démographique actuelle doit être formulée ensemble, avec la perception de l'étendue du domaine du point de vue du contenu, et du potentiel et des moyens disponibles, y compris les bénéfices dérivés de l'expérience acquise par les étudiants qui travaillent comme « internes » dans les agences et les bureaux municipaux (des projets pareils ont déjà été mis en place dans quelques localités). L'attention de plus en plus grande que l'on accorde à l'offre doit dépasser, bien sûr, l'aire des bureaux municipaux ou des petites entreprises et doit concerner les aires micro- et macro-économique pour l'administration publique, l'autorité locale, l'intervention gouvernementale : par conséquent, les démographes doivent aussi connaître les problèmes auxquels le monde des affaires se confronte, ainsi que ceux, à une échelle plus réduite, qui apparaissent au niveau de l'autorité locale.

Les relations entre la démographie et les politiques ont toujours été plutôt délicates (cet aspect a souvent été discuté) et ceux qui font partie du gouvernement sont intéressés par les aspects dont les résultats se verront pendant leur mandat (et, actuellement, ces mandats peuvent être très courts). La démographie, en tant que variable, en ce qui la concerne, implique plutôt un laps de temps trop grand entre investissements et résultats et le recueil des effets tarde à venir.

Par conséquent, la tâche principale peut être la définition d'un statut professionnel clair, pour des compétences spécifiques, qui ne doit pas empêcher la promotion de ces intervenants dont les domaines d'activité, conformément à la pratique en Italie, incluent aussi celui des démographes (statisticiens, analystes des marchés, les gens du marketing).

Quelles sont les options pour la démographie italienne ? Peut-être, cette discipline peut-elle agir comme un catalyseur pour faire admettre que les ressources professionnelles formées de différentes compétences démographiques sont vitales pour le management, comme un stimulus pour les utilisateurs potentiels - et par conséquent, pour la demande - capables de percevoir la démographie comme un investissement. Comment ? Si ce développement ne se réalise pas parmi les sociologues et les géographes, comme aux États-Unis, alors les démographes doivent eux-mêmes propager et illustrer ce que la perspective démographique impose. Ils doivent aborder les utilisateurs en accordant, principalement, une attention spéciale à l'offre démographique (il faut donc s'orienter vers la formation, non pas en augmentant le nombre de cours mais en examinant de près ces aspects qui impliquent plus concrètement des activités de nature professionnelle ; peut-être, par la création des modules complémentaires, on peut mettre en pratique des cours qui ont déjà été testés).

Les observations que nous avons faites dans ce court article ne sont pas exhaustives et, bien sûr, il faut retenir que nous n'avons pas essayé de fournir une image définitive des perspectives professionnelles réelles des démographes ; il n'est pas non plus dans notre intention de fournir des indications techniques au sujet des nouveaux curriculums statistiques et

démographiques. Des objectifs pareils seraient trop ambitieux et prématurés. Notre seul objectif a été de réaliser une analyse des perspectives et des opinions qui dans le passé n'avaient pas été exprimées, de stimuler un débat qui, nous l'espérons bien, élargira le contexte au sein duquel pourront se créer de nouveaux débouchés pour les démographes en Italie.

Annexes

Quelles sont les débouchés disponibles pour la démographie professionnelle en Italie ?

1) En dehors de vos activités universitaires, avez-vous jamais acquis de l'expérience professionnelle en tant que démographe ou travaillé dans des domaines ayant un contenu démographique ?

OUI

NON

Si oui, veuillez mentionner :

- a) secteur public _____ secteur privé
- b) consultant _____ autre _____
- c) occasionnel _____ fréquemment

2) Pendant vos cours de démographie, vous êtes-vous déjà concentré sur les aspects techniques et opérationnels de cette discipline du point de vue de la formation professionnelle des étudiants ?

OUI

NON

3) Est-ce que vous connaissez des cours de Démographie des affaires en Italie, ou des cours formulés explicitement pour le secteur privé ou l'administration publique ?

OUI

NON

Si oui, veuillez fournir plusieurs informations sur ces cours (professeur, localité etc.).

4) Quelle est votre formation professionnelle et quand avez-vous terminé vos études ?

Formation :

Année :

5) Pensez-vous que la formation démographique que vous avez acquise a été adéquate du point de vue des futures activités professionnelles en dehors des domaines universitaires et de la recherche générale ?

OUI

NON

6) Selon vous, quels types de cours sont suffisants, concernant la formation démographique fournie, actuellement, aux étudiants en vue de leur offrir des débouchés professionnels pour les structures publiques et privées, compagnies etc. ? (Vous pouvez mentionner plusieurs cours)

Statistiques, Statistique et économie, Économie et affaires

Statistiques et démographie, Sciences actuarielles et statistiques, Sciences politiques

Autres, Aucun

7) A votre avis, quel est le sujet de la Démographie Institutionnelle qui devraient être développé pour réaliser un curriculum spécifique pour les « démographes professionnels » ayant des perspectives de travail dans le secteur privé, l'autorité locale ou le secteur public ?

L'approche de localisation optimale

Djilali SARI

En général, toute localisation obéit à des critères d'implantation bien établis et devant satisfaire des objectifs clairement définis et à atteindre dans les meilleures conditions possibles. Mais, bien souvent, des contraintes diverses perturbent toute localisation adéquate. Des contraintes pas toujours liées à la rentabilité des ressources financières, comme le montrent parfois certaines localisations par suite de la proximité de sources de pollution, ou de fortes concentrations d'activité ...

En effet, à la lumière de faits et observations puisés dans l'actualité de ces dernières années, de graves défaillances sont relevées en la matière et ne cessent de nous interpellier, à plus forte raison en ce qui concerne les autorités et les services techniques directement concernés à tous les niveaux, particulièrement quand des catastrophes entraînent de lourds bilans sur les plans matériel et humain. Tels sont les cas de sinistres aussi bien d'origine naturelle (inondations, séismes, mouvements de terrain spectaculaires...) que chimique (émanation de gaz, pollution multiforme) et thermonucléaire (Tchernobyl). De plus en plus, bien des « certitudes » tombent les unes après les autres.

En conséquence, les problèmes de localisation se posent plus que jamais avec acuité. Il convient non seulement de tirer les enseignements les plus en vue, fussent-elles d'origine naturelle mais aussi et de plus en plus en ne tenant pas compte de certaines lois bien établies par les observations et relevant d'une connaissance intime du milieu naturel, géomorphologique et historique, soit tout l'environnement humanisé à travers des siècles interrompus d'occupation...

Aussi proposons-nous de fixer l'attention sur les trois points suivants :

- les enseignements de certaines catastrophes
- les objectifs de localisation optimale
- l'approche générale

Toutefois, compte tenu des axes retenus dans la thématique générale, les aspects ayant trait à la population proprement dite ne sont pas abordés, à l'exclusion de reflux et vieillesse, soit ce qui a trait à la désertification « humaine ». En fait, il s'agit de simples allusions et il en va de même des effets de la double explosion : démographique et urbaine.

Les enseignements de certaines catastrophes

Plus que jamais, des catastrophes bien déterminées offrent un champ privilégié d'observations et d'enseignements, principalement quand la part de responsabilité de l'homme est nettement lisible dans l'aggravation des bilans en pertes humaines et autres dommages. Pour cela les exemples ne manquent pas.

Les catastrophes liées à certaines intempéries

Dans les régions méditerranéennes, ce sont bien les inondations qui induisent une part notable - en termes financiers - des dégâts, jusqu'à 80% dans le territoire français, et sont bien souvent et de plus en plus accompagnées de victimes. Or comme elles résultent de précipitations torrentielles pouvant revêtir la forme de tornades cataclysmiques, le phénomène ne saurait surprendre, comme le souligne tout examen approfondi de longues séries de relevés. Cependant, dans les conditions présentes, ce qui importe le plus, c'est de différencier ces précipitations exceptionnelles de leurs effets et conséquences. D'autant plus que le bilan tend à

s'alourdir sur le double plan, matériel et humain, en frappant l'opinion publique par suite de la « surmédiation ».

Il en est ainsi en particulier des inondations catastrophiques survenues en automne 1993 et en 1994 dans le Midi de la France (Vaucluse) avec le lourd bilan relevé à Vaison-la-Romaine (une vingtaine de morts) et à Nîmes. On peut citer aussi les effets dévastateurs enregistrés en Italie du Nord lors la dernière saison hivernale. Incontestablement l'aggravation doit être recherchée dans l'intensité particulièrement forte des précipitations dans le contexte des implantations économiques et des infrastructures dévastées. Pour cela limitons-nous au cas précité de Vaison-la-Romaine. C'est bel et bien la localisation du site endommagé et à l'origine du nombre élevé des victimes qui a été dénoncé. *Les constructions se sont « aventurées » jusqu'au ... mineur du cours*, mettant en évidence d'autres défaillances, d'ordre humain avant tout. De plus, des investigations plus poussées, tant en aval qu'en amont, et partout à travers le bassin versant mettent en évidence d'autres défaillances, toutes à l'origine de l'aggravation du bilan.

En définitive, cet exemple et bien d'autres rendent compte dans des cas précis de la tournure de l'évolution en cours. En conséquence, c'est la remise en *cause d'équilibres plus ou moins précaires* ... Dans de telles conditions aussi, la vigilance doit être accrue et constante. A plus forte raison dans les pays du Sud où bien souvent les conditions sont encore plus aggravantes !

En effet, là incontestablement il faut s'attendre à bien des catastrophes plus grandes comme le montre l'exemple suivant emprunté à l'épisode fluvial du 24 au 27 janvier 1992 et limité au centre de l'Algérie et intervenu après une longue sécheresse récurrente sévissant pratiquement depuis les débuts des années 70. C'est ainsi qu'à la faveur d'une infiltration d'une goutte froide, on a relevé des hauteurs oscillant de 150 à 200 mm sur le piémont de l'Atlas tellien et jusqu'à 300 mm en altitude, contre 120 à 150 mm en plaine (fig.1). Or ces hauteurs, sans être trop excessives, n'ont donné lieu qu'à une intensité assez faible (1,5 à 3 mm/h) et sans pointe significative. En revanche, compte tenu de la sécheresse, les cours d'eau n'ont commencé à monter qu'à partir de la fin de la deuxième journée pour atteindre le maximum le 28 janvier avec des crues allant de 58,2 hm³ aux Issers à 105,5 hm³ à Oued El Harrach. Seul ce dernier représente 20% du total annuel. Quant aux dégâts, ils doivent être appréciés à leur juste valeur, sur le double plan matériel et humain. Dans un tout autre contexte, précédant celui des trois dernières décennies... les conséquences auraient été fort réduites. L'aggravation résulte avant tout du manque d'entretien des drains, par suite de l'obstruction de nombreuses sections du lit ... mineur de l'oued (décharges publiques et privées illégales) et de l'occupation des méandres par l'habitat spontané, soit le développement spectaculaire de bidonvilles... submergés après la montée des eaux.

Or c'est bien au cours de ces trois dernières décennies que l'on assiste à une croissance démographique sans précédent, et par voie de conséquence aussi et surtout à une véritable explosion urbaine (Sari, 1992 : 371-75) qui a fini par envahir à la fois les lits majeurs et mineurs, voire des vallées entières, et indépendamment des zones mal drainées et de surcroît insalubres, d'où les conséquences très graves avec des bilans particulièrement élevés dans le cas de retour et d'exacerbation des précipitations torrentielles. Bien plus, avec des intensités moins accusées que naguère, les dégâts seront plus accablants sur les divers plans. Il en sera ainsi des zones d'extension urbaine récente ainsi que de certaines infrastructures localisées sur les piémonts et montagnes.

Les catastrophes liées à la sismicité et à une certaine géomorphologie

Une sismicité et une géomorphologie bien déterminée s'impliquent de plus en plus dans l'aggravation des bilans, principalement quand il est question d'un contexte socio-économique donné.

Plus que partout ailleurs, l'héritage géomorphologique du quaternaire au Maghreb (Raynal, 1957 : 885-894) est très sensible, par la nature lithologique de ces terrains et formations superficielles, aux processus d'érosion quand surviennent les précipitations torrentielles dans un contexte de déforestation et de surcultures généralisées. D'autant plus que les pressions engendrées par la double explosion, démographique et urbaine, se soldent inévitablement par la remise en cause générale des équilibres précaires.

En conséquence, la sécheresse récurrente ne fait que reculer les échéances... tôt ou tard, les conséquences signalées à l'issue de l'épisode du 24 au 27 janvier 1992 se conjuguent à celles qui seront liées au mouvement et à l'instabilité des terrains après l'apparition d'intempéries persistantes. Les glissements de terrain ne seront pas exclus en revêtant diverses formes et indépendamment de la submersion de quartiers entiers et récents jusqu'aux villes anciennes en particulier.

Quant à la sismicité, elle ne saurait être sous-estimée puisqu'elle affecte bien les zones surpeuplées du pourtour méditerranéen, particulièrement au Maghreb. Du reste, l'évolution récente de l'Algérie le montre bien. Si, dans le temps, les conséquences étaient réduites, voire peu sensibles, il n'en va pas de même présentement. L'aggravation croissante des bilans de victimes dans le cas d'El Asnam (ex-Orléansville) est en rapport avec l'explosion urbaine de la ville. Les séismes de 1954 et 1980 sont intervenus dans deux phases d'urbanisation diamétralement différentes, la première quand la ville coloniale était encore circonscrite au périmètre de création au milieu du XIX^{ème} siècle avec quelques constructions isolées, alors que la seconde concerne une agglomération beaucoup plus étendue avec des effectifs qui ont plus que doublé en passant de 24000 à 62000 habitants. De plus, les derniers mouvements telluriques (1989) ont trait à des zones récemment urbanisées comme autour de Tipaza et de Mascara en 1994.

En conséquence, de tels faits sont riches d'enseignement et constituent autant de leçons pour tous les responsables chargés de la localisation des projets. Plus que jamais, c'est une cartographie des risques qui doit servir de base à des investigations plus poussées et elle est indispensable à toute prise de décision. Une telle cartographie si incontournable doit précéder aussi et surtout les études préliminaires ainsi que celles de faisabilité

PLUVIOMÉTRIE DES 24-28 JANVIER 1992 (SOURCES : A. LEMOU (1992, 5 : 16))

Stations	Altitude (mètres)	Hauteur (mm)
Baraki	20	135
Birmandreis	40	148
Boufarik	60	168
Bou Medfaa	95	159
Bouharoun	270	254
Mouzaïa-Mines	520	379
Miliana	715	336

VALEURS DES CRUES (HM³) (SOURCES : A. LEMOU (1992, 5 : 16))

Stations	Lieux	Valeurs
Oued Chétif	Arib	61,3
O Mazafran	Zeralda	71,0
O El Harrach	Baraki	105,5
O Isser	Lakhdaria	58,2

Les objectifs de localisation optimale

Administration publique, promoteur et tout décideur, ne peuvent se prononcer clairement sur la justesse de la localisation optimale qu'en fonction d'objectifs préalablement bien définis afin de pouvoir appréhender les incidences sur le plan spatial à moyen et long termes. Pour cela, intéressons-nous plus spécialement aux trois objectifs suivants : la lutte contre les phénomènes de désertification, le désenclavement et la lutte contre les nuisances multiformes.

La lutte contre les phénomènes de désertification

Les phénomènes de désertification, humaine et naturelle, affectent de plus en plus des régions bien déterminées, particulièrement au cours de ces dernières décennies, en se soldant, suivant les cas, soit par les dépeuplements et en définitive l'exode rural, soit encore par la conjonction d'autres phénomènes, d'ordre aussi bien naturel qu'humain.

Dans l'ensemble, les premiers cités concernent de vieux pays, les pays développés. Là le reflux de certaines tranches d'âges (population active) concourt au vieillissement et par voie de conséquence, au repli et à la marginalisation de zones entières et se solde par le transfert des services de proche en proche.

En revanche, dans les régions appartenant aux pays en développement, le reflux peut être déclenché et accéléré non seulement par des phénomènes propres à tout exode rural, mais aussi à la dégradation du milieu, notamment du capital biologique avec perte des sols et déforestation irréversible avec toute une réaction en chaîne très dangereuse (Dubost M, 1992 : 17-28). Au stade de bad-lands la restauration devient pratiquement impossible.

En conséquence, le décideur doit avoir une vision très large et la plus profonde possible pour pouvoir cerner les problèmes et les enjeux, à court et long termes.

Le désenclavement

Tel est bien l'objectif que doit rechercher tout décideur en s'efforçant pleinement de réanimer des contrées condamnées au repli et l'exclusion par l'histoire, la géographie et, de plus en plus, l'économie de marché, alors qu'elles ne sont pas toujours dépourvues de ressources exploitables et rentables si des conditions minimales sont réunies...

Aussi l'objectif à rechercher se confond-il avec la prospection des zones à désenclaver. Plus que jamais, la cartographie s'avère très utile en fixant et focalisant l'attention à la fois sur les faiblesses et les potentialités que représentent les localisations à promouvoir. Cependant l'approche doit être bien motivée en n'oubliant nullement les véritables enjeux du présent et de l'avenir à court et long termes. La recherche d'équilibre régional doit être à l'origine de telles motivations. En définitive, une véritable politique d'aménagement intégré doit être soutenue constamment. Ces conditions sont difficiles à réunir dans un tel contexte géographique, celui des pays du Sud, mais sont indispensables pour la réanimation de vastes contrées.

La lutte contre les nuisances multiformes

La localisation optimale au sein d'un environnement immédiat proche ou même éloigné présentant des nuisances, fusent-elles peu visibles, et a fortiori masquées, mérite une attention particulière. Ces nuisances multiformes ne sont pas toujours celles auxquelles on pense sous l'effet de la médiatisation et qu'on a déjà évoquées. Elles ne peuvent se limiter seulement à l'industrie chimique et à toute sous-traitance du nucléaire. Du reste, la recrudescence des affections respiratoires le prouve aisément, particulièrement à la recherche de sites de substitution dans le cadre à la fois du redéploiement et de facilités ne prenant pas en compte les intérêts des populations des zones d'implantation ... A cet égard, songeons aux zones dépotoirs de matières radioactives révélées « prudemment » de temps à autre par certains médias.

Les problèmes spécifiques des pays du Sud résultent en particulier des conditions d'une urbanisation explosive. Là aussi, l'actualité est très riche d'enseignement si les médias arrivent à poursuivre objectivement leur mission d'information et de sensibilisation, résultat qui ne se manifeste guère qu'à la suite de la résurgence de maladies infectieuses, souvent consécutives à l'infiltration d'eaux usées dans le réseau d'eau potable. Dans tous les cas, la réapparition de plus en plus fréquente de maladies naguère éradiquées souligne la gravité et la complexité des problèmes d'aménagement et de réhabilitation des zones d'habitat spontané. Toute localisation optimale se heurte donc à des obstacles souvent insurmontables alors que la solution doit être urgente et impérative.

Par ailleurs, certaines infrastructures doivent partout focaliser l'attention. Il en est ainsi des aéroports proches des zones d'habitat alors que l'espace-tampon ne cesse de se réduire, mettant en danger la vie des populations, d'autant plus que les normes de sécurité ne sont pas toujours respectées. La cartographie à jour devient ainsi indispensable.

L'approche générale

Les problèmes à résoudre sont donc nombreux et fort complexes. L'arbitrage se complique souvent à la suite de la divergence des intérêts en jeu. Toutefois l'objectivité ne peut être assurée sans une approche d'ensemble solidement étayée. La synthèse doit être exprimée par une cartographie appropriée et enrichie par toute une illustration graphique et des projections pour faciliter la lecture et par voie de conséquence les prises de décision.

L'importance d'une cartographie synthétisée

En synthétisant les principaux résultats obtenus à travers les voies et moyens explorés, l'outil cartographique devient ainsi incontournable.

C'est ainsi qu'en ce qui concerne les objectifs, on parvient à identifier deux types de zones, les unes répulsives et les autres attractives. Les premières expriment le reflux des habitants en fonction des causes indiquées. En conséquence, elles constituent un appel des implantations qu'il convient de susciter et de soutenir par des projets adéquats en mesure d'inverser les mouvements migratoires. C'est aussi le même objectif recherché à travers les initiatives de désenclavement. Dans les deux cas, les localisations optimales seront celles qui parviendront progressivement à résoudre le maximum de problèmes.

En revanche, les zones attractives ne peuvent se développer aux dépens des zones dépressionnaires. Elles ne peuvent donc bénéficier de nouvelles implantations économiques, ni de certains services publics pour ne point attirer davantage la population en provenance des régions limitrophes.

Par ailleurs, on doit croiser cette cartographie avec celle déjà évoquée en premier lieu. En dressant une carte représentant les risques suivant une échelle graduée, l'on parvient à éliminer les zones dangereuses, en sélectionnant les meilleurs sites d'implantation. Les éliminations recouvrent inévitablement les zones inondables, celles à moyenne et forte sismicités, et les terrains instables et en proie à de violentes formes d'érosion.

L'approche doit être complétée aussi par celle propre aux projections de l'urbanisation. Comme déjà souligné ci-dessus, les phénomènes d'urbanisation dans le cas de population à fort accroissement demeurent difficilement maîtrisés. En effet, il convient d'économiser certaines infrastructures ne suivant pas le rythme d'accroissement naturel de la population en question.

Enfin cette approche doit se référer aussi et surtout à la méthode suivante qui vient d'être codifiée et qui a été publiée récemment. Elle correspond parfaitement à nos préoccupations majeures de terrain, de terrain durement éprouvés par les phénomènes présentement analysés (Sari,

1977) et de plus en plus par les effets d'une sécheresse récurrente (Sari, 1994). C'est celle de l'inondabilité.

L'efficience du modèle « inondabilité »

C'est plus qu'une lutte contre les inondations et les fortes crues. Il s'agit avant tout d'un ensemble d'aménagements prenant en compte tout le bassin versant avec des transferts judicieux des volumes d'eau dans l'espace et surtout dans le temps en fonction à la fois des besoins et des sécheresses persistantes, voire récurrentes comme dans le cas propre au Maghreb et bien d'autres zones très exposées. Tel est bien le résultat du programme de recherche du CEMAGREF (Sécheresse 1993, no 3, Vol. 4, p. 171-176).

Reprenant les principaux objectifs dans ce cadre trop étroit :

« Le modèle « inondabilité » présenté ici a pour vocation d'aider la société à reprendre en main un aménagement et une gestion intégrées et équitables de ces eaux continentales. Il a pour objectif principal le ralentissement dynamique de la ressource en eau et pour contrainte structurante la satisfaction de la demande de protection raisonnée contre les inondations. Par raisonné, on entend un choix qui tient compte des effets aval (directs) des protections locales. Le produit « inondabilité » est une carte de synthèse lisible... »

Incontestablement c'est un outil précieux. D'autant plus précieux que les dernières inondations frappant certains pays d'Europe occidentale ont gravement perturbé la vie socio-économique des habitants alors qu'inversement les pays du Maghreb sont de plus en plus affectés durement par les effets de la sécheresse récurrente, ne parvenant pas à préserver les réserves des retenues des grands ouvrages, ni à emmagasiner une partie des fortes précipitations survenant en quelques jours. D'importantes hauteurs sont ainsi charriées dans la mer, au détriment de l'homme et de la nature.

En conséquence, le modèle « inondabilité » donne lieu à l'établissement de cartes avec indication de trois zones : une zone jaune non inondée, une zone rouge dite inondée et une zone verte inondée surprotégée. Pareille cartographie jointe à celle des risques, de population et d'infrastructure constitue une synthèse incontournable pour la prise de toute décision.

Conclusion

Ainsi toute localisation optimale doit-elle répondre à des objectifs bien définis devant tenir compte du maximum de données et de paramètres propres non seulement au milieu socio-économique et culturel mais aussi au milieu naturel, géomorphologique, sismique et météorologique. Toute synthèse, aussi satisfaisante que possible, doit souligner et exprimer lisiblement l'interaction de ces phénomènes et relations intimes avec les retombées et incidences multiples sur le plan spatial proprement dit. Pareille démarche ne saurait être négligée, même quand il est question d'implantation de projets à caractère local et à plus forte raison de service public, souvent condition indispensable à la survie et la revivification d'isolat, de «vieux» pays en voie de dépeuplement... Plus que jamais, le rapprochement administrés-administration s'impose. Principalement dans les milieux ne parvenant pas à participer activement à la chose publique !

BIBLIOGRAPHIE

- CLAUDE J., (1993) Crues et sécheresse, *Sécheresse*, Paris no 3, vol. 4 : 141-142
- DUBOST M., (1992), Les pratiques agricoles et la transformations des terres dans les montagnes et les régions en pente de la Méditerranée, in *Montagnes et forêts - méditerranéennes*, Le Bourget-du-Lac, MAB, UNESCO, CIHERM, SCOPE, pp. 17-18
- LEMOU A (1992), L'épisode pluvieux du 24 au 27 janvier 1992 sur le Centre du pays, *Eaux et Sols d'Algérie*, Alger, Agence Nationale des Ressources Hydrauliques, no 5, pp. 15-19
- OBERLIN G., et al., (1993), Une méthode globale pour la gestion des zones inondables, le programme «Inondabilité» du CEMAGREF., *Sécheresse*, no 3, vol. 4, pp. 171-176.
- QUEZEL P., et al., (1993), Pratiques agricoles et couvert en région méditerranéenne, *Montagnes et forêts méditerranéennes*, Le Bourget-du-Lac, MAB, UNESCO, CIHERM, SCOPE, pp. 71-90.
- RAYNAL R., (1957) Bodenerosion in Marokko, *Wissenschaftliche Zeitschrift*, Martin Luther University, Hall, Math Natur, Klasse, pp. 885-894.
- SARI Dj., (1977), L'homme et l'érosion dans l'Ouarsenis (Algérie), SNED, 625p.
- SARI Dj., (1992), Deux décennies d'urbanisation en Algérie, *Croissance démographique urbanisation*, Paris, AIDELF, pp. 371-379.
- SARI Dj., (1994), La désertification et l'homme en Algérie, *Séminaire international : Population and Environnement in Arid Regions*, Amman, 24-28/94, mulg.

Analyse spatiale et modèles de localisation optimale.

Outils opérationnels d'aide à la décision.

Isabelle THOMAS

F.N.R.S., Département de Géographie de l'U.C.L.

1. Introduction

Les *services* prennent de plus en plus d'importance dans nos économies tant au point de vue du personnel employé, du coût de fonctionnement, de la qualité des services à rendre aux utilisateurs que des équipements ou de la consommation d'espace et des externalités qui en découlent. Aujourd'hui, plus que jamais, leur gestion doit être efficace et optimale. Mettre un service à la disposition du public c'est, notamment, veiller à le rendre géographiquement aussi *accessible* que possible. Par conséquent, la *localisation* d'un service ne peut être choisie approximativement : elle doit répondre à des critères précis et tenir compte de contraintes diverses clairement exprimées. Quelle que soit l'échelle à laquelle un problème d'organisation est étudié, les enjeux spatiaux sont plus que jamais une réalité dont il faut tenir compte. Le rôle du géographe est alors de produire les outils et connaissances utiles pour faciliter la décision de choix spatial.

La *géographie* a pour but (I) de comprendre pourquoi les activités humaines s'installent en certains lieux, (II) d'expliquer les structures spatiales qui en résultent (notamment la manière dont les individus se comportent dans un cadre géographique) et donc aussi (III) de fournir des outils permettant d'agir sur l'espace et sur sa gestion. Cette problématique concerne les géographes mais aussi les économistes et les chercheurs opérationnels. Des éléments de réponse ont été élaborés au sein de ce qu'il est convenu d'appeler la *théorie de la localisation*. Nous n'aborderons pas ici l'ensemble des méthodes mais nous nous limiterons à un seul aspect, à un seul type de démarche et à un exemple. Parmi les activités structurantes de l'espace, nous nous limiterons à la localisation de *services*. Les modèles de localisation se préoccupent de la localisation optimale d'activités mais aussi de l'affectation des clients aux installations: l'objectif des *modèles de localisation-affectation* est donc de déterminer à la fois les localisations optimales de services et l'affectation de la clientèle aux sites choisis. Dans la plupart des modèles proposés jusqu'à présent, ces localisations optimales résultent d'un compromis entre les coûts de fonctionnement du service et les bénéfices que retirent les usagers de la proximité de ce service sous forme de coût de transport, d'accessibilité ou de temps d'intervention. Les exemples sont nombreux : quel bureau de poste supprimer afin de minimiser les inconvénients de la fermeture et la perte de clientèle ? Où localiser un parc à conteneur afin de rencontrer les besoins de la population et de minimiser les effets néfastes ? Comment organiser les zones d'intervention des services d'incendie ? Comment organiser les zones téléphoniques ?

L'*objectif* de cette contribution est d'expliquer le plus simplement possible comment fonctionnent les modèles de localisation optimale, d'illustrer les modèles par un cas concret, d'insister sur quelques problèmes d'applicabilité auxquels la recherche essaie de répondre, et, enfin, de réfléchir sur l'interaction potentielle entre la démographie et la géographie pour faire de ce type d'approche un outil pluridisciplinaire pleinement opérationnel pour les services publics.

2. Les modèles de localisation-affectation

2.1 Définition.

Les problèmes liés à l'organisation spatiale optimale des services publics ont donné naissance à une littérature scientifique qu'il n'est pas dans notre intention d'inventorier ici (Beaumont, 1981; Domschke et Drexl, 1985; Hansen e.a. 1987; Current et Shilling, 1990). L'optimalité spatiale d'un service public comporte diverses facettes : la localisation des installations fixes, l'organisation d'éventuelles unités mobiles, le découpage en zones de compétence, la concurrence entre sites d'offre, l'affectation de la clientèle, le comportement spatial des usagers. Rares sont les modèles tentant d'intégrer plusieurs facettes.

Face à cette multitude de problèmes, les *modèles de localisation-affectation* sont nombreux. Le choix d'un modèle dépend de la formulation précise du problème considéré. Une formulation fréquemment utilisée est celle de la *p-médiane*. Ce modèle présente l'avantage de trouver des solutions exactes et d'avoir une formulation souple se rapprochant ainsi de bien des situations concrètes d'aménagement. Notons aussi, qu'une application récente des modèles de localisation aux services d'incendie a permis de démontrer l'utilité de la *p-médiane* par rapport à d'autres modèles (Richard e.a., 1990). Nous retiendrons la *p-médiane* pour illustrer nos propos.

La *p-médiane* consiste à choisir la configuration géographique des unités d'offre de manière à minimiser la somme des distances parcourues sous une série de contraintes énoncées par l'utilisateur. Le modèle assure la couverture efficace (voire équitable) du milieu. Dans bien des cas de services, l'hypothèse du plus court chemin est tout à fait soutenable. La *p-médiane* présente l'avantage d'être facilement adaptée aux spécificités du problème posé par l'utilisateur et d'être résolue par des méthodes efficaces, rapides et souples (Hanjoul et Peeters, 1985; Hansen, e.a., 1987). Ce modèle permet non seulement de s'attacher à la forme des aires et aux localisations optimales, mais également de suggérer le nombre idéal de services et de simuler des solutions sous diverses contraintes et hypothèses.

Le problème consiste donc à choisir la configuration géographique des unités d'offre j de manière à minimiser la somme pondérée des distances parcourues par les utilisateurs. F^* est notée comme étant la solution qui optimise F :

$$\text{Min} \sum_{i=1}^m \sum_{j=1}^n a_i x_{ij} d_{ij}$$

où :

i : indice des points de demande, j : indice des sites potentiels d'offre, a_i : poids affecté au point i , d_{ij} : coût de transport entre i et j , n : nombre de localisations potentielles d'offre, m : nombre de points de demande, x_{ij} : portion de la demande émanant de i et allouée en j ,

Ce problème est contraint :

A. La totalité de la demande doit être satisfaite par les p sites localisés .

$$\sum_j x_{ij} = 1$$

B. La demande de i allouée en j ne peut être satisfaite que si un site est implanté en j .

$$0 \leq x_{ij} \leq y_j$$

où y_j vaut 1 lorsqu'une facilité est localisée en j , et 0 si non.

C. p est le nombre d'équipements à localiser, ou, le budget alloué pour la construction des unités.

$$\sum_j y_j = p \text{ et } y_j = 0 \text{ ou } 1$$

D. (facultatif) Une contrainte de distance maximale à parcourir par les utilisateurs entre leur localisation i et l'unité d'offre j peut être imposée. Le modèle propose alors une solution telle qu'aucun point de demande ne soit séparé de plus de cette quantité d_{max} du service le plus proche.

A l'optimum, chaque point de demande i est satisfait par l'unité j la plus proche, et la localisation est telle que le coût total d'accès de l'ensemble des points i inclus dans l'aire de service de chaque unité j est minimum.

Traduisons ce problème en termes concrets de planification publique.

2.2 Un exemple

La liste de situations concrètes d'application est longue et impossible à dresser. Les modèles de localisation-affectation s'appliquent tant à des exemples micro-géographiques (la localisation des bulles à verres au sein d'une commune, suppression d'une école primaire, la localisation d'un parc à conteneurs), que des exemples méso-géographiques (organisation des districts de gendarmerie, organisation des zones téléphoniques à l'échelle nationale) ou macro-géographiques (implantation de sièges de multinationales en Europe). Les exemples peuvent tant concerner des services attirant le client (crèches, écoles, piscines), que se déplaçant vers le client (pompiers, ambulances) ou ayant les deux fonctions à la fois (la poste, la police), ou encore des services dont la fréquentation est obligée (écoles primaires) ou non (bibliothèques).

Les problèmes d'aménagements peuvent être très variés : (I) le service peut ne pas encore exister et il convient de le mettre en place. Dans ce cas, on peut imaginer un ensemble de sites potentiels et proposer/évaluer les solutions; (II) l'offre existe déjà et elle est localisée en n points. Dans ce cas, on peut imaginer évaluer la situation existante, devoir supprimer k sites ou au contraire en proposer des additionnels. Il importe dans chaque cas de connaître les localisations et la taille des services ainsi planifiés.

Prenons un exemple afin d'illustrer la démarche modélisatrice : la localisation optimale des bureaux de Poste à Namur (Thomas, 1983, 1984).

2.2.a Le problème

Namur est une commune de 100.000 habitants dont seul le tissu urbain sensu stricto a été délimité à l'aide d'une approche géographique multivariée. Le milieu est représenté par 420 points, chacun correspondant à une portion de secteur statistique (figure 1). La demande est mesurée par le nombre de visites mensuelles à la Poste émanant de chaque point i de demande. Cette quantité a été estimée par enquêtes et les mécanismes étudiés avec précision. Un modèle économétrique a été construit. Dans cet exemple, la demande ne dépend pas de la distance séparant la demande de l'offre : on ne fréquente pas plus le service quand on en est proche mais la règle de fréquentation du bureau de poste le plus proche a été vérifiée. Chaque étude particulière doit être précédée d'une analyse fine du comportement de l'utilisateur : comportement en matière de demande, comportement de choix spatial.

La distance entre chaque point du réseau est mesurée sur carte et exprimée en kilomètres. Il n'est pas tenu compte de la distance-temps. Plus de 100 points de demande sont retenus comme localisation potentielle de l'offre (figure 2).

La p -médiane a été appliquée à cet ensemble de données (1) pour suggérer au décideur quel est le nombre idéal de bureaux pour desservir la zone urbaine namuroise, (2) pour voir quels seraient les bureaux à supprimer tout en maintenant un niveau d'efficacité et d'équité raisonnables, et (3) pour évaluer la situation existante (18 bureaux) face à une situation optimale.

FIGURE 1 :
DISTRIBUTION DES 420 POINTS DE DEMANDE

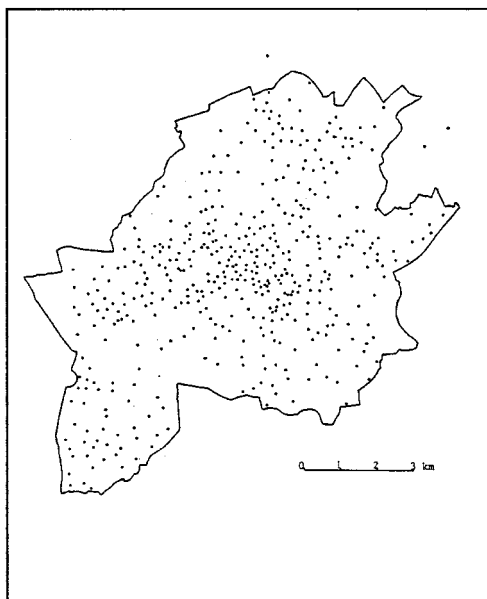
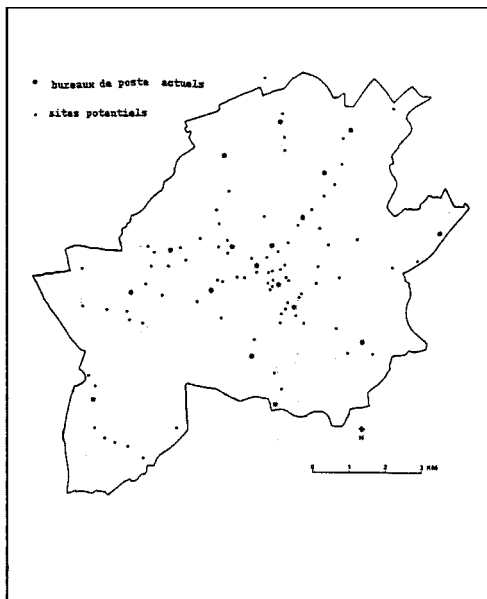


FIGURE 2 :
DISTRIBUTION DES SITES POTENTIELS D'OFFRE

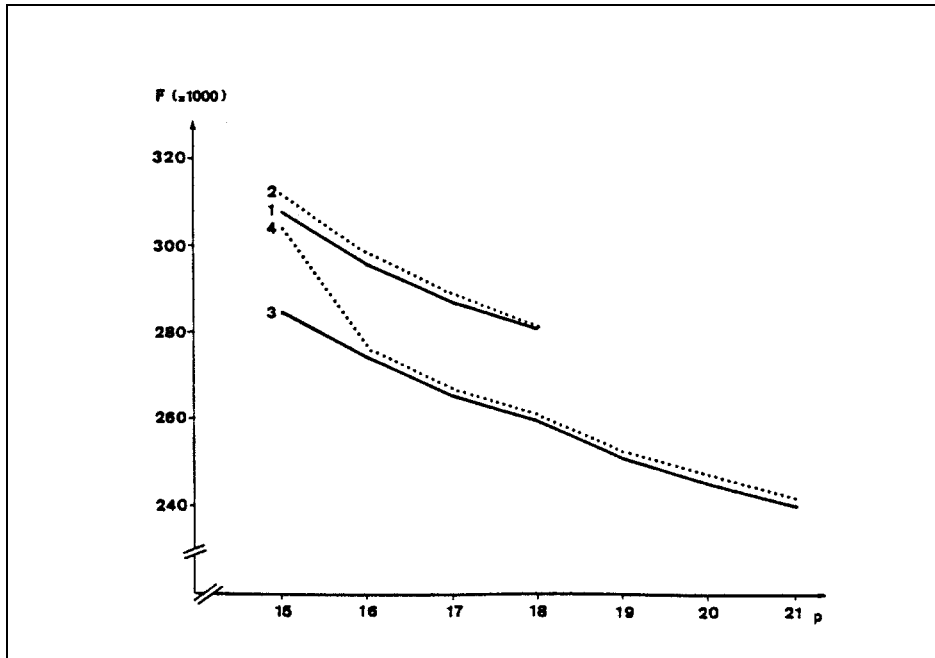


Les outputs de la p -médiane sont : l'accessibilité totale du système, la distance moyenne parcourue par un usager et la distance maximale séparant l'utilisateur le plus mal localisé d'un site d'offre, la localisation optimale des bureaux, l'affectation de la clientèle aux services et la forme des zones desservies.

2.2.b Évaluation en terme de distances

La figure 3 rend compte de la valeur de la fonction objectif F^* , c'est-à-dire le coût total d'accès aux services. Au plus on localise de bureaux au plus la clientèle est proche d'un service et au moins elle parcourt de kilomètres : quand p augmente, F^* diminue. La courbe 1 rend compte de la situation actuelle (18 bureaux) dont on supprime successivement un, deux ou trois bureaux de poste. Supprimer un ou deux bureau(x) ne coûte pas beaucoup en terme d'accroissement du déplacement. En supprimer trois coûte beaucoup plus cher car ces suppressions affectent alors des lieux centraux (tableau 1). La courbe 3 rend compte du coût de déplacement vers les bureaux de poste localisés idéalement : le modèle a choisi p bureaux parmi l'ensemble des sites potentiels; la courbe indique donc que 16 bureaux localisés idéalement à Namur fonctionnent aussi bien que les 18 bureaux actuels ! L'allure de ces deux courbes montre clairement la croissance du coût de déplacement associé à la réduction du nombre d'équipements implantés. En plus, F^* présente une cassure légèrement plus marquée lorsque p vaut 16 : nous pouvons donc considérer cette valeur comme candidat à l'optimum susceptible de retenir l'attention du décideur.

FIGURE 3 : ÉVOLUTION DU COÛT TOTAL DE TRANSPORT F^* AVEC LE NOMBRE DE BUREAUX DE POSTE (p)
 Courbes 1 et 3 pas de contrainte de distance maximale, courbes 2 et 4 contrainte de distance maximale.



Les courbes 2 et 4 rendent compte des résultats du modèle pour lequel une contrainte de distance maximale a été envisagée : la distance maximale que peut parcourir l'utilisateur le plus défavorisé est limitée à 6.1 km. Le choix de cette quantité est basé sur la configuration actuelle : elle correspond à la plus petite valeur nécessaire pour assurer le recouvrement total du milieu par le service. En effet, si la contrainte est trop stricte (valeur trop petite de la distance maximale), le modèle de la p -médiane ne trouvera plus de solution : il existe des points de demande séparés de plus de d_{max} du bureau de poste le plus proche. Or, le modèle postule que tout point de demande doit être affecté à une unité d'offre.

Le tableau 1 résume les principales caractéristiques des solutions proposées : la valeur de la fonction objectif F^* , la distance maximale parcourue par l'utilisateur le plus mal desservi et la distance moyenne parcourue par l'utilisateur. Quelle que soit la restriction effectuée sur le nombre actuel de sites, la distance maximale parcourue par le client le plus mal localisé varie très peu. Cette forte valeur concerne quelques usagers localisés dans les communes périphériques de la ville. Ces communes ont été incluses dans le milieu d'étude pour des raisons administratives de dépendance postale uniquement : elles n'appartiennent pas au milieu homogène urbain namurois défini par ailleurs. La distance moyenne est peu élevée (587 mètres) : les bureaux de poste actuels sont très proches de la majeure partie de la clientèle. La distance actuelle parcourue par l'utilisateur moyen correspond donc au seuil de distance maximal qu'un piéton accepte généralement de parcourir (500 mètres). Cette bonne adéquation au public s'explique partiellement par la manière dont le budget est utilisé : dans le cas des services récréatifs, par exemple, la totalité du budget est affecté aux coûts de fonctionnement. Par contre, dans le cas des services tels que la Poste, une partie de cette somme doit être attribuée au coût de transport à charge de l'État c'est-à-dire les coûts engendrés par la distribution et la collecte du courrier. Cette fonction du service postal est ignorée dans notre modèle, car elle représente un problème

totalemment différent. Nous devons cepandant être conscients de son existence pour l'explication de l'adéquation.

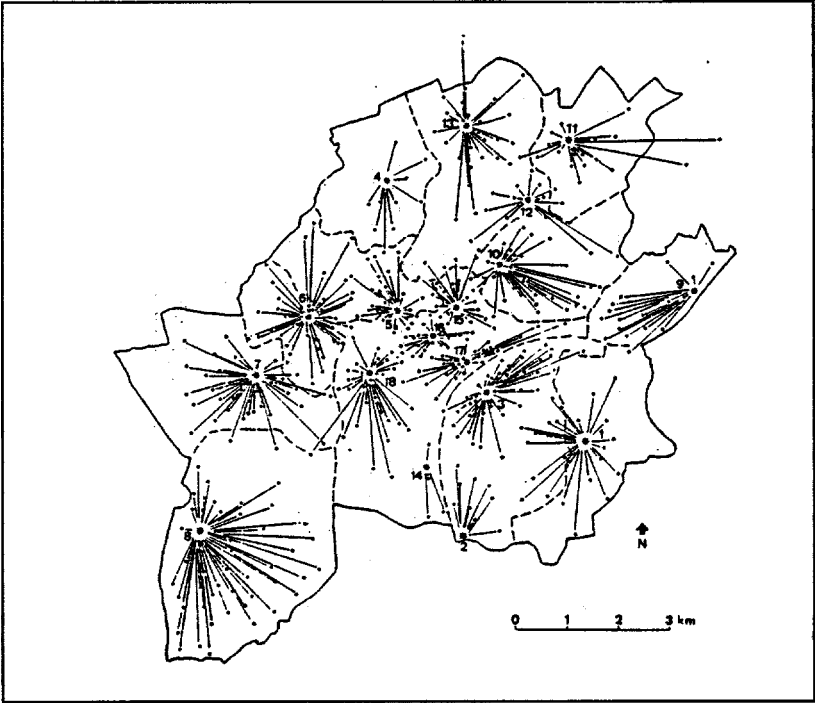
TABLEAU 1 : ÉVOLUTION DE L'EFFICACITÉ (D_{MOY}) ET DE L'ÉQUITÉ (D_{MAX}) EN FONCTION DU NOMBRE DE SITES CHOISIS (P).

p	F^*	(1)	d_{moy}	d_{max}
18	280827	-	587	6100
17	286977	2.2	600	6100
16	295318	5.2	621	6100
15	307801	9.6	643	6100
(1) représente la réduction de la valeur de F^* par rapport à la situation actuelle de 18 unités d'offre.				

2.2.c Évaluation en terme de localisations

La figure 4 rend compte des décisions spatiales suggérées par le modèle. La première suppression affecte la halte postale de Saint Marc, soit l'installation représentant le plus petit volume de travail (0.5 % du volume d'heures de travail prestées, 0.2 % du volume des opérations de guichet enregistrées). Cette fermeture affecte très peu le coût de déplacement à charge de la collectivité (+2.2 %) et réduit faiblement les frais de fonctionnement à charge du décideur. Notons que les deux autres haltes sont maintenues car elles représentent un volume de travail plus important et leurs localisations sont mieux justifiées par rapport aux autres bureaux actuels, à la demande et aux autres services namurois.

FIGURE 4 : SOLUTION PROPOSÉE PAR LE MODÈLE (THOMAS 1984)



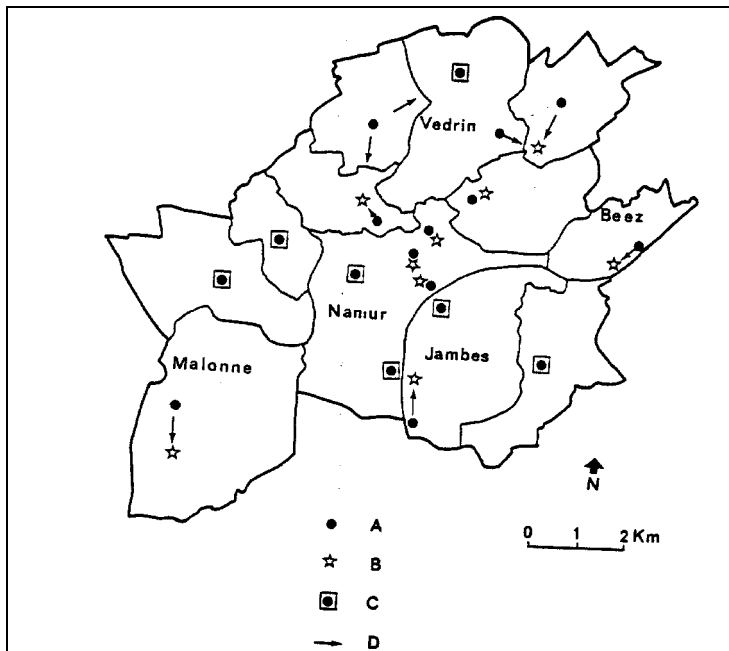
La deuxième suppression touche un autre bureau périphérique : Champion, dont la clientèle est rattachée à un petit bureau actuel (Vedrin 2). L'imposition d'une contrainte de distance maximale à parcourir par l'utilisateur le plus défavorisé modifie cette situation en faveur du bureau supprimé; cette suppression est aussi l'exemple d'un effet de frontière certain. La troisième restriction retenue concerne le bureau de Namur 3 situé en milieu urbain, à proximité d'autres bureaux. Au-delà de trois suppressions, la perte d'efficacité dépasse 10 % et entraînerait certainement une transformation des habitudes de consommation de la part des usagers qui se dirigeraient peut-être vers des services concurrents mieux répartis dans l'espace.

La planification des restrictions en termes d'équipements postaux a été simulée sous deux hypothèses réalistes (I) restrictions simultanées de plusieurs bureaux et (II) restrictions successives. La première suppose qu'en aucun cas, un bureau n'ait déjà été préalablement éliminé. Les 18 bureaux actuels sont affectés par la fermeture simultanée de k sites. La seconde hypothèse envisage des restrictions séquentielles opérées successivement dans le temps : étant donné une première fermeture, quel est, dans l'ensemble des bureaux restant, celui qui correspond au coût supplémentaire minimal en terme de distance à parcourir par l'utilisateur ? Cette deuxième hypothèse est plus conforme à la réalité : la répartition actuelle est le résultat d'une séquence d'actions et non d'une planification systématique. Dans cet exemple simple, les deux types d'approches convergent.

2.2.d Évaluation en terme d'affectation

Puisque le modèle alloue chaque point de demande i à un site d'offre j , la somme des poids a_i affectés en j donne l'*affectation*, la taille du service. Cet output du modèle permet de voir si l'espace planifié est isotrope, c'est-à-dire que tous les services sont de même taille, ou, au contraire, de fortes différences apparaissent, une hiérarchie du service peut être envisagée. Le tableau 2 reprend les écarts entre l'affectation planifiée et l'affectation observée pour les 18 bureaux de Poste existants à Namur.

FIGURE 5 : LOCALISATION ACTUELLE DES BUREAUX ET AFFECTATION OPTIMALE (THOMAS 1986)



Une analyse de la *forme des aires* de service des bureaux de Poste peut également être envisagée. Bien des services accordent de l'importance à cette notion. La figure 5 illustre la situation postale optimale à Namur. L'analyse détaillée des aires de service attendues (figure 4) permet de confirmer l'utilité de la p -médiane en planification, et dans ce cas-ci de l'utilité de la p -médiane pour la planification des bureaux de Poste. La formulation la plus simple retenue dans cet exemple ne rend cependant pas compte de l'attractivité différentielle des sites d'offre de tailles extrêmes (très petits et très grands bureaux). La différence observée entre les aires théoriques et attendues des bureaux moyens varie en fonction des circonstances locales.

Bien d'autres exemples opérationnels ont déjà été développés (piscines, crèches, pompiers, ambulances, distributeurs de billets de banque, écoles, bibliothèques, etc.). Tous confirment l'opérationnalité du modèle utilisé dans sa forme la plus simple ou dans une forme plus complexe incorporant des procédures économétriques d'estimation de la demande ou du comportement spatial du consommateur.

3. Sensibilité des modèles à la formulation des inputs

L'application des modèles de localisation-affectation à des cas concrets soulève malheureusement encore de nombreux problèmes (Leonardi, 1981; Rushton, 1989) : représentation et inclusion du comportement du consommateur et/ou du décideur, estimation de la demande, représentation de la distance, comparaison efficace des solutions, effet de l'agrégation et de l'échelle spatiale sur les résultats, inclusion et mesures de la réalité administrative et décisionnelle, forme du réseau, effet de frontière, sont autant de problèmes non encore résolus. Les résultats des modèles de localisation sont hypothéqués par ces choix. La plupart de ces problèmes trouvent une solution partielle dans l'analyse des données spatiales (Haining, 1990); leur discussion n'est pas neuve pour le géographe mais leur extension aux modèles de localisation-affectation est récente.

Appliquer un modèle de localisation-affectation signifie choisir des mesures qui représentent le fonctionnement du service dans le milieu étudié. Envisageons de façon générale quelques-uns de ces choix dont l'importance n'est pas négligeable dans les résultats opérationnels de modèles.

3.1 La demande

La demande pour un service se localise en un point i et se représente par une quantité a localisée en i . Il importe donc de mesurer avec exactitude la demande exprimée en unités adéquates, de la localiser en des lieux préalablement définis et de connaître les mécanismes qui la génèrent. La demande est hétérogène en quantité, nature et conséquences : crimes, accidents, vols, fausses alarmes sont autant d'exemples faisant la diversité du travail du gendarme. Toute analyse de localisation optimale doit être précédée d'une analyse approfondie de la demande et de ses composantes spatiales. Mieux comprendre ces mécanismes signifie mieux les estimer et donc construire un modèle de localisation plus réaliste et conduire à de meilleures simulations. Des développements théoriques et empiriques sont encore nécessaires dans ce domaine; ils relèvent de l'analyse des données.

Chaque service mène à une *estimation* différente des mécanismes générant la demande. La demande peut éventuellement dépendre de la distance qui la sépare du service (élasticité). Le rôle de l'aménageur est de trouver la bonne estimation de cette quantité a_i qui s'avère souvent être une fraction de la population d'un secteur, quartier ou commune. Il importe de bien cerner cette fraction et éventuellement son comportement dans le temps et l'espace. La sensibilité des modèles à la manière de mesurer la demande a été vérifiée (par ex.: Beguin e.a., 1992, Thomas, 1986 et 1993-b, Richard, 1985). Ces auteurs y démontrent que dans le cas des localisations optimales des bibliothèques, des bureaux de Poste ou des casernes de pompiers, il

n'est pas nécessaire de recourir à l'enquête : la simple estimation de la demande par la population conduit à des résultats similaires. Ceci permet de conclure à une bonne estimation du volume de la demande par une mesure de *population* mais n'exclut aucunement des mécanismes spécifiques liés à certaines études de cas.

Outre la quantité a , il importe également de bien définir les lieux i : le chercheur a rarement le choix du niveau d'*agrégation spatiale* avec lequel il travaille; son choix est contraint par la disponibilité des données. Chaque lieu i correspond à une observation localisée dans l'espace; cette observation est rarement une donnée individuelle et fréquemment un agrégat d'individus (exemple : secteurs statistiques, communes, arrondissements). L'effet de l'agrégation des données sur les résultats d'analyses statistiques spatiales est connu depuis longtemps (voir : Arbia, 1989; Haining, 1990). Le problème des unités surfaciques modifiables (M.A.U.P.) se décompose en problème de taille (ou d'échelle) et de forme (niveau et mode d'agrégation). Ce problème ne connaît pas encore de solution dans le domaine de la localisation optimale (Casillas, 1987; Current et Schilling, 1987; Daskin e.a., 1989; Debeer et Thomas, 1994; Hodgson et Neuman, 1993) : les résultats sont contradictoires et méritent une attention particulière. Le rôle de l'agrégation de la demande dans les modèles de localisation inquiète les chercheurs sans qu'aucune solution efficace et standard ne soit proposée afin d'en éviter les biais. L'aménageur doit être conscient du biais introduit par son choix. Il n'a souvent pas le contrôle de ce biais : les données lui sont données par l'autorité publique.

Le chercheur dispose parfois de données individuelles: lorsque les données sont totalement *désagrégées* dans l'espace (voire même dans le temps), le phénomène est représenté par un nuage de points dans lequel la structure spatiale est souvent difficilement décelable (exemple : la distribution des accidents de la route peut être considérée comme aléatoire). L'existence d'un point de demande serait donc d'autant plus aléatoire que la durée de la période d'observation est courte et/ou l'agrégation spatiale est faible. Récemment plusieurs auteurs ont développé des solutions pour tenir compte de la composante aléatoire de la demande, du temps de déplacement ou de la disponibilité d'un véhicule dans les modèles de localisation (Daskin, 1987; Goldberg et Paz, 1991). Ils ont introduit un modèle de type « file d'attente » comme sous-routine à la p -médiane (Berman, Larson et Parkan, 1987) : cette démarche comble certaines lacunes, mais la solution devient rapidement énorme en termes informatiques.

Si les modèles de type « file d'attente » sont excellents pour réfléchir sur un problème d'organisation spatiale microscopique, ils deviennent vite inapplicables avec des tailles importantes (échelle régionale, par exemple). Doit-on modéliser d'abord une situation générale et ensuite tenir compte de situations particulières et donc rares ? A l'échelle microspatiale, la décision humaine n'est-elle pas aussi opérationnelle que le meilleur modèle stochastique ? En d'autres termes, n'est-il pas opportun de réfléchir sur l'organisation d'un service à une échelle macrospatiale et ensuite résoudre empiriquement certains problèmes locaux telle l'occurrence aléatoire d'un appel supplémentaire, ou l'adresse précise du service à localiser ?

Le choix du modèle serait-il lié au niveau d'agrégation spatiale ? Agrégation spatiale et déterminisme sont étroitement liés. A un niveau microspatial, l'occurrence d'un événement peut être le résultat d'un processus aléatoire, mais un service est rarement créé pour quelques unités de demande. L'analyse de la distribution spatiale de nombreux événements dans un environnement conduit à reconnaître des groupes de faits : certains crimes sont concentrés dans des quartiers urbains défavorisés, certains accidents dangereux se concentrent les nuits de week-end, etc. Ceci signifie que même si UN incident peut être considéré comme aléatoire, l'occurrence de nombreux événements conduit à considérer une explication partiellement déterministe et à faire appel à l'autocorrélation (Odland, 1988; Griffith, 1992). *Agrégation et autocorrélation* spatiale sont intimement associées à toute application d'un modèle spatial et a fortiori aux problèmes de localisation optimale. Leur impact exact et les solutions pour éviter leur biais mériteraient plus d'attention.

Le niveau d'agrégation de la demande et l'estimation de la quantité de demande sont deux problèmes auxquels l'aménageur doit bien réfléchir : comme dans toute démarche modélisatrice, ses résultats peuvent en dépendre.

3.2 Forme du réseau, mesure de la distance et externalités

L'espace étudié est représenté par un réseau composé de noeuds et d'arcs. Les noeuds du réseau correspondent aux sites d'offre et de demande. A chaque arc est associée une mesure de coût de déplacement. La mesure et l'estimation de la *distance* séparant deux lieux i et j est un problème ancien en géographie. Il trouve un nouveau domaine d'application dans les modèles de localisation. En effet, plusieurs auteurs se préoccupent de connaître la sensibilité des résultats aux mesures de la distance (voir par exemple Hodgson et Storrier, 1994); d'autres s'intéressent à l'influence de la forme du réseau sur les résultats (Peeters et Thomas, 1995) soulevant ainsi de nombreux problèmes complexes tels que l'influence de la position optimale d'un périphérique urbain sur les localisations optimales.

Dans la plupart des modèles proposés jusqu'à présent, les localisations optimales résultent d'un compromis entre les coûts de fonctionnement du service et les bénéfices que retirent les usagers de la proximité de ce service sous forme de coût de transport, d'accessibilité ou de temps d'intervention. Or, toute décision d'implantation affecte aussi l'environnement des particuliers, même de ceux qui ne sont pas directement concernés par le service en question. Ceci est connu en économie sous le nom d'*externalités*; celles-ci peuvent être positives lorsque l'environnement est modifié d'une façon favorable (par exemple par une augmentation des valeurs immobilières) ou négatives lorsque le service est source de nuisances (les écoles engendrent du bruit et des problèmes de trafic à certaines heures, les casernes de pompiers ont leurs sirènes et leurs parkings pour engins spéciaux, etc.). L'inclusion des externalités spatiales dans les modèles de localisation soulève une difficulté méthodologique supplémentaire : alors que les déplacements de et vers les usagers s'effectuent généralement au moyen d'un réseau routier, c'est par d'autres canaux (air, cours d'eau, sol, etc.) que se diffusent les externalités, dont les nuisances et pollutions. Si un réseau routier peut être bien défini, le vent - par exemple - est une variable qui se traite de manière beaucoup plus complexe étant donné ses multiples directions potentielles. L'espace géographique est ainsi doté simultanément de deux types de distances : une *distance-transport* (entre l'utilisateur et le site d'offre visité) et une *distance-externalité spatiale* (entre l'utilisateur et tous les points d'offre, telle la distance-pollution). De plus, la demande pour le service en question émanant d'un point du territoire n'est pas nécessairement liée aux externalités dont bénéficie ou que subit la population située en cet endroit. Ce double aspect des implantations de services n'a reçu jusqu'à présent aucun traitement satisfaisant : soit on ignore les externalités comme dans les modèles traditionnels, soit on ne prend en compte qu'elles seules dans une famille de modèles spécifiquement liés aux services polluants.

Appliquer un modèle simple de localisation ne pose aucun problème méthodologique. Incorporer des formes différentes de coûts de transport (congestion, externalités, etc.) rend le modèle plus réaliste mais complique à la fois la récolte des données et son inclusion dans le modèle opérationnel.

3.3 Comment définir l'optimalité ?

Définir « la » fonction objectif n'est pas toujours chose simple : la structure administrative et la disparité des tâches accomplies par un service rend difficile l'optimisation d'un seul critère et oriente la recherche vers des méthodes multicritères. Considérons un exemple de service : la gendarmerie. Habituellement les modèles de localisation optimale minimisent les temps de déplacement; cela suppose que les équipes rentrent immédiatement après l'intervention. Comment tenir compte des « tournées » ? Comment tenir compte des

temps variables passés sur place ? Comment tenir compte du travail de prévention accompli éventuellement sur le chemin de retour ? Faut-il tenir compte des temps de réaction ? des délais d'intervention ? Le temps de déplacement n'est qu'une fraction de la réalité d'intervention.

Sur le terrain, les équipes d'intervention fixes et mobiles coexistent. Comment concilier ces deux types de service ? Toutes peuvent répondre à un appel. Chaque type d'équipe doit-elle correspondre à l'application d'un modèle différent ? Étant donné leurs différences, peuvent-elles être considérées comme un tout ? Comment synthétiser l'optimalité et les disparités spatiales ? L'optimalité spatiale est-elle stable dans le temps ? La plupart des modèles s'occupent de représenter une situation à un moment déterminé. Or, les heures d'ouverture du service, les fluctuations saisonnières ou journalières de trafic, le caractère saisonnier de certains délits, etc. sont autant de fluctuations perturbant l'organisation spatiale d'un service et posant la question de la signification d'une situation moyenne. L'organisation spatiale optimale basée sur une situation moyenne est-elle compatible avec ces sources de variation ? Doit-on construire différents modèles ? Une réflexion plus profonde doit permettre de construire une méthodologie générale pour ce type de problème.

Dans l'étude d'un cas concret, l'organisation réelle du service devient tellement complexe que le choix d'UNE fonction-objectif est difficile, si non impossible. Le coût d'un modèle multiobjectif se justifie-t-il ? Il serait intéressant de tester les différences entre un modèle mono-objectif et un modèle complexe afin de juger du surcoût engendré par une meilleure représentation de la réalité. Chaque facette de la réalité spatiale doit-elle être prise en compte par un modèle particulier ou une situation moyenne est-elle suffisante ? N'est-il pas temps de comparer les résultats de plusieurs techniques (mono-, multi-critères) ?

Théorie et analyse des données sont intimement liées : les données ne sont pas sans conséquences sur les formulations d'hypothèses et sur les choix de modélisation. La diversité des services et des environnements spatiaux engendrent tellement de disparités qu'il sera difficile de définir une seule méthodologie.

3.4 Qu'est-ce qu'une limite de zone d'affectation ?

Soit l'exemple d'un service d'urgence minimisant le temps moyen de déplacement. Au moment où l'appel arrive, le service doit veiller à ce qu'une équipe soit disponible pour répondre à l'appel dans le périmètre de l'appel. Les conséquences d'un découpage spatial en zones ont rarement retenu l'attention des chercheurs. Or, ce découpage influence l'efficacité du service : les auteurs comparent les localisations centrales mais négligent la forme des aires. Leurs limites définissent des aires d'intervention exclusives du service. La signification de ces limites revêt toute son importance lorsqu'on sait que chaque service concerné par un accident (médical, pompier, police, gendarmerie, arrondissements judiciaires, zones téléphoniques, etc.) définit son découpage spatial propre. La coordination des services et la construction d'un modèle spatial unique permettrait de lever bien plus de dysfonctionnements spatiaux que l'optimisation d'un seul service.

Les modèles tel que la p-médiane permettent de définir des aires de service. Ces limites ne constituent pas une règle à ne pas transgresser mais bien un point de référence, comme l'est toute abstraction ou tout modèle. Ces limites seront d'autant plus floues que le comportement du consommateur sera impliqué. En matière d'urgence, la gendarmerie actuelle transgresse ses limites de districts dans moins de 5% des interventions. Faut-il en tenir compte dans une modélisation ? La résolution humaine au cas par cas peut être tout aussi efficace pour ces quelques interventions.

Enfin, n'oublions pas les effets de frontière, d'autant plus importants que l'application concerne un petit espace : le fait d'ignorer le « reste du monde » perturbe les résultats en marge du milieu d'études.

3.5 Urbain ou rural : deux réalités à considérer séparément ?

Les modèles de localisation optimale des services d'urgence ont particulièrement été appliqués en *milieu urbain* : les différences d'efficacité y sont mesurées en secondes plutôt qu'en minutes vu la densité du réseau routier, de l'offre et de la demande. La congestion du réseau routier ou du service pèsent considérablement dans le fonctionnement du service : l'incertitude associée à chaque tronçon est forte en milieu urbain (Rushton, 1989).

En *milieu rural*, la distance kilométrique revêt tout son sens; elle est déterminante dans l'organisation optimale d'un service : les distances sont mesurées en minutes plutôt qu'en secondes (Halseth et Rosenberg, 1991). Le nombre de points d'offre et de demande y est peu élevé, la densité d'usagers potentiels est faible alors que les distances à parcourir sont beaucoup plus longues.

Étant donné ces différences, est-il possible de trouver une formulation commune pour les deux types d'environnements ? Les échelles spatiales sont différentes. Les milieux mixtes (rural, urbain) ont rarement attiré l'attention des chercheurs car il est difficile d'y trouver un bon compromis entre équité et efficacité. Traiter le rural et l'urbain séparément signifie précéder toute étude d'organisation spatiale optimale d'une approche socio-économique afin de définir avec précision les limites de ces deux milieux. Cette procédure est bien connue du géographe; elle est remise à l'ordre du jour avec les modèles de localisation. Le problème régional revêt ici toute son importance : par définition une région comprend villes et campagnes. Une solution consisterait à procéder en *deux étapes* : d'abord définir les mésozones à l'échelle macrospatiale à l'aide de programmation linéaire telle la *p*-médiane; ensuite, à l'intérieur de ces mésozones, concevoir l'organisation spatiale précise avec éventuellement d'autres techniques (hypercube, etc.). Il serait alors important de mesurer dans quelle proportion cette solution en deux temps s'écarte de la solution tenant compte d'un espace unique.

3.6 Les administrations disposent-elles de données fiables ?

Appliquer un modèle de localisation optimale signifie disposer de nombreuses données sur l'environnement géographique, sur le service, sur le comportement des usagers et du décideur. La plupart des données ont une composante spatiale et/ou temporelle, ajoutant une dimension au problème. Les services publics ne disposent généralement pas d'une telle base de données répondant aux exigences du modèle de localisation. Certaines données sont récoltées, d'autres sont informatisées, d'autres encore sont récoltées par différents organismes, posant ainsi le problème de la *compatibilité* et de la *comparabilité* des bases de données. De nombreuses données sont inaccessibles au planificateur. Il n'existe souvent aucune mesure de *fiabilité* des données récoltées, information primordiale en analyse de données.

Citons l'exemple des bureaux de poste : la Poste ne connaît pas la clientèle qui se rend dans ses bureaux. Seule une enquête coûteuse permet de combler cette lacune, particulièrement lorsque l'on veut localiser la clientèle. Autre exemple : les interventions d'urgence à la gendarmerie. Le nombre et la nature des appels d'urgence ne sont connus que par l'intermédiaire d'une statistique manuelle, dans laquelle aucune mention n'est faite - par exemple - du nombre d'hommes/heure nécessaire pour chacune d'entre elles, (c'est-à-dire des charges de travail).

Une des raisons du développement de modèles simples ou du petit nombre d'applications réalistes trouvées dans la littérature peut donc être le manque de données concernant les inputs

du modèle. Le fonctionnement du service est mal compris et certainement mal mesuré. La collecte et la gestion des données devraient être facilitées par l'avènement des SIG (Systèmes d'Information Géographique) et leur inclusion dans les dispatchings. Le succès de ces nouvelles techniques informatiques dépend toutefois de la récolte et de la fiabilité des données. Chercheurs et décideurs gagneraient à collaborer plus étroitement : une meilleure connaissance de l'utilité des statistiques motiverait les administrations publiques à améliorer la qualité des données et à mieux cibler leur récolte.

4. Conclusion

Qu'ils soient publics ou privés, firmes et services sont de plus en plus conscients de l'importance du choix d'une implantation adéquate. Les modèles de localisation-affectation constituent un outil très attrayant pour orienter le choix de telles décisions spatiales. Ils permettent d'aider à la prise de décision par un outil scientifiquement justifié. A l'aménageur - géographe, démographe ou économiste - d'analyser consciencieusement tous les problèmes d'estimation qui sont associés à la localisation optimale : connaissance et compréhension du service à localiser, mesure et localisation des quantités demandées au moment de l'analyse et ultérieurement, connaissances de mécanismes de comportement des usagers et du décideur, estimation du coût de déplacement, etc.

Les potentialités d'applications sont multiples; communes et administrations publiques se posent de nombreux problèmes de restriction (quelle école supprimer ?) ou de création de services (où localiser le tri sélectif des immondices ?). Population et pouvoirs publics usent de leur influence. Le modèle propose un outil simple qui permet d'orienter le débat sans prendre en compte les luttes d'influence. Utilisé simplement, il oriente le débat. Utilisé avec plus de minutie quant aux inputs, il peut proposer des solutions très précises. Enfin, pour le scientifique, ce type de modèle offre encore de nombreuses perspectives de recherche. En effet, comme dans bien d'autres travaux de modélisation, les problèmes soulevés dans cet article relèvent de la recherche scientifique. Soulever ces questions ne signifie pas diminuer l'utilité des modèles mais en poser les limites actuelles : la plupart relèvent de la modélisation conceptuelle mais davantage encore de l'*analyse des données*. Comment associer les problèmes d'analyse spatiale des données au problème étudié ? Le modèle doit-il tenir compte de toute la réalité environnementale, comportementale et administrative ? Un modèle n'est-il pas avant tout une abstraction de la réalité ? Pour arriver à un bon choix, chercheurs et décideurs doivent collaborer. Plusieurs disciplines scientifiques peuvent apporter leur savoir-faire.

Les modèles de localisation-affectation semblent techniquement mûrs, mais de nombreux problèmes généraux se posent quant à leur applicabilité - comme dans bien d'autres types de modèles économétriques ou démographiques. Gageons que le développement des systèmes experts et des systèmes d'information géographiques facilitera l'application et l'applicabilité des modèles, aidera les chercheurs à réfléchir sur les nombreux biais introduits et les aménageurs à faciliter leur mise en application.

BIBLIOGRAPHIE

- ARBIA G. (1989), *Spatial Data Configuration in Statistical Analysis of Regional Economics and Related Problems*. Doordrecht, Kluwer, 256p
- BEAUMONT J. (1981), Location-Allocation Problems in a Plane : A Review of Some Models. *Socio-Economic Planning Science* 15 : 217-229.
- BEGUIN H., THOMAS I., VANDENBUSSCHE D. (1992), Weight Variations within a Set of Demand Points, and Location-allocation Issues: a Case Study of Public Libraries. *Environment and Planning-A*, 24, 1769-1779.
- BERMAN O., LARSON R. et PARKAN C. (1987), The Stochastic Queue p -median Problem. *Transportation Science* 21 : 207-216.
- CASILLAS P. (1987), Data Aggregation and the p -median Problem. In *Spatial Analysis and Location-Allocation Models*. Edited by Ghosh, A. and Rushton, G. 327-344. New-York: Van Nostrand.
- CURRENT A., SCHILLING D. (1987), Location Modeling : Perspective and Overview. *Geographical Analysis* 22 : 1-3.
- DASKIN M. (1987), Location, Dispatching, and Routing Models for Emergency Services with Stochastic Travel Times. In *Spatial Analysis and Location-Allocation Models*. Edited by Ghosh and Rushton. 224-265. New York: Van Nostrand.
- DASKIN M., HAGHANI A., KHANAL M., MALANDRADRAKI C. (1989), Aggregation Effects in Maximum Covering Models. *Annals of Operations Research*, 18, pp. 115-140.
- DEBEER I., THOMAS I. (1994), Sensibilité de la p -médiane à l'agrégation spatiale de la demande. Une illustration. *Acta Geographica Lovaniensia*, 34, 715-727.
- DOMSCHKE W., DREXL A. (1985), Location and Layout Planning. *Lecture Notes in Economics and Mathematical Systems* (Springler Verlag, Berlin and New York) 238.
- GOLDBERG J., PAZ L. (1991), Locating Emergency Vehicle Bases when Service Time Depends on Call Location. *Transportation Science* 25 : 264-279.
- GRIFFITH D. (1992), What is Spatial Autocorrelation ? Reflections on the Past 25 Years of Spatial Statistics, *L'Espace Géographique*, 3, 1-16
- HAINING R. (1990),. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press.
- HALSETH G., ROSENBERG M. (1991), Locating Emergency Medical Services in Small Town and Rural Setting. *Socio-Economic Planning Sciences* 25 : 295-304.
- HANJOUL P., PEETERS D. (1985), Comparison of Two Dual-Based Procedures for Solving the p -Median Problem. *European Journal of Operational Research*. 20:387-396
- HANSEN P., LABBÉ M., PEETERS D., THISSE D. (1987), Facility Location Analysis. *Fundamental of Pure and Applied Economics*, 22, pp.1-70.
- HODGSON M., NEUMAN S. (1993), A GIS Approach to Eliminating Source Aggregation Error in p -Median Models, *Location Science*, 1:2, pp. 155-170.
- HODGSON M., STORRIER A. (1994), Functional Error in Geographical Analysis: the Case of Spatial Separation in Location-allocation. *Geographical Systems* (in press).

- LEONARDI G. (1981), A Unifying Framework for Public Facility Location Problems-Part 1 : A Critical Overview and some Unsolved Problems. *Environment and Planning A* 13 : 1001-1008.
- ODLAND J. (1988), *Spatial Autocorrelation*. London : Sage. 87p.
- PEETERS D., THOMAS I. (1995), The Effect of the Spatial Structure on the p -Median Results. *Transportation Science* (in press).
- RICHARD D. (1985), *La localisation des services d'incendie dans le province de Luxembourg*. Thèse de doctorat, U.C.L., Département de Géographie
- RICHARD D., PEETERS D., BEGUIN H. (1990), The Location of Fire Stations in a Rural Environment : a Case Study. *Environment and Planning A* 22 : 39-52.
- RUSHTON G. (1989), Applications of Location Models. *Annals of Operations Research* 18 : 25-42.
- THOMAS I. (1983), Amélioration du découpage géographique des agglomérations urbaines. Proposition d'une méthode simple d'agrégation. *L'Espace Géographique*. 3, 207-214
- THOMAS I. (1984), Towards the Simplification of Location Models for Public Facilities: the case of the Postal Service. *Papers of the Regional Science Association*, 55, pp. 47-58.
- THOMAS I. (1986), *La localisation optimale des services publics. Une méthode opérationnelle et son application au service postal*. Louvain-la-Neuve, Cabay, Revue des Thèses, n°11, 269 p.
- THOMAS I. (1993-a), Appliquer un modèle de localisation-affectation à une situation concrète : une gageure ? *Revue d'Economie Régionale et Urbaine* 2, pp. 201-214.
- THOMAS I. (1993-b), Sensibilité du découpage spatial optimal des services d'urgence à la définition de la demande. Quelques réflexions sur un exemple mésogéographique. *L'Espace Géographique*. 4, 318-332

Comment appréhender le problème statistique des petits nombres en démographie ?

Pierre ARS¹

Institut de Statistique - Université catholique de Louvain, Belgique

Luc DAL² et **Michel POULAIN**³

GéDAP - Université catholique de Louvain, Belgique

Dans cet article, nous discutons des problèmes statistiques liés à l'analyse démographique de petites populations. Dans ce but, nous construisons des intervalles de confiance pour des taux, quotients ou proportions. Dans le cas de (petites) populations, il est possible de construire des intervalles exacts, tandis que dans les autres cas, on détermine des intervalles de confiance plus précis que ceux utilisés d'ordinaire. En outre, on propose un intervalle de confiance approché et un test exact pour la différence de deux proportions. Les différentes méthodes sont appliquées à deux situations réelles.

1. Les objectifs poursuivis

En démographie, on considère principalement deux types de variables à partir desquelles on peut étudier la population relative à un territoire donné :

1. l'effectif de cette population et de ses différentes sous-populations (selon l'âge, le sexe, l'état matrimonial, la nationalité, ...). Ces nombres sont des entiers non négatifs correspondant à des observations à un instant donné « t ». En les rapportant les uns aux autres, on calculera généralement des proportions caractérisant différentes sous-populations.
2. des nombres d'événements (naissances, décès, mariages ou migrations) rendant compte du mouvement de la population pendant une période donnée $[t, t+1]$. Sur base de ce second type de variables, on définit le plus souvent des taux en rapportant le nombre d'événements à la population moyenne soumise au risque.

Qu'en est-il de l'importance du caractère aléatoire de ces mesures ou dénombrements?

Tout d'abord, on ne peut exclure que, dans certains cas, la mesure démographique en elle-même (le dénombrement des structures ou la comptabilité des événements) soit également entachée d'une marge d'erreur. Cette marge d'erreur pourra comprendre une erreur de type systématique de sous-dénombrement, par exemple, mais également une erreur de type aléatoire qui trouve son origine au cours des différentes phases de la méthode de la collecte. Ce phénomène se constate très bien dans le dénombrement des migrations internes à un pays donné qui font l'objet d'une double mesure de façon quasi indépendante. En fait, tout courant de migration interne est comptabilisé avec un même outil statistique au titre d'émigrations par le lieu de départ et au titre d'immigrations par le lieu de destination. Pratiquement, on observe que les chiffres relatifs à un même courant de migration diffèrent et que la différence entre les deux mesures est proportionnelle à la racine carrée du volume du courant de migration que l'on

¹ ars@stat.ucl.ac.be

² dal@spri.ucl.ac.be

³ poulain@spri.ucl.ac.be

cherche à mesurer. Il diminue par conséquent en termes relatifs avec la racine carrée de l'importance de ces courants de migration⁴.

Par ailleurs, nous nous placerons résolument dans le cadre de l'hypothèse avancée par Keyfitz en 1966 : rappelons que celle-ci postule que la population étudiée n'est autre qu'un échantillon extrait d'une population hypothétique de taille infinie. Prenons un exemple : si la population infinie comprend un certain pourcentage (inconnu) de personnes âgées de 60 ans et plus, la population finie de taille n étudiée affichera, quant à elle, une proportion qui ne sera qu'une estimation de ce pourcentage. Plus la taille n de la population sera importante et plus l'estimation sera meilleure et assortie d'un intervalle de confiance plus réduit.

Enfin, dans le cas spécifique du dénombrement des événements au cours d'une période donnée en matière de fécondité, mortalité ou mobilité spatiale, on est évidemment dans l'impossibilité de maîtriser toutes les variables (comportementales, physiologiques, socio-économiques et autres) pouvant rendre compte de l'occurrence de ces événements. On préférera par conséquent admettre qu'il s'agit d'une occurrence à caractère aléatoire et que le nombre d'événements effectivement observés entre deux instants t_0 et t_1 sera le résultat d'un processus qui comprend une composante aléatoire.

Finalement, peu importe si le caractère aléatoire des mesures démographiques peut se justifier par l'une ou l'autre, voire plusieurs, de ces hypothèses. Ce qui prime, c'est d'estimer l'ampleur de ce caractère aléatoire au niveau de l'analyse et de l'interprétation des indices démographiques que nous calculons couramment. C'est principalement lorsque les populations sont de taille réduite, et/ou que les événements considérés sont rares que ce caractère aléatoire se traduit par un risque d'interprétation erronée des indices calculés. Force est donc de constater que toutes les mesures démographiques sont affectées d'une composante aléatoire qui nécessite le recours à un intervalle de confiance, celui-ci étant d'autant plus restreint que la taille de la population concernée est grande. Dès lors, en matière de petites populations et de petits nombres, les deux questions principales qui se posent au démographe sont les suivantes :

1. indépendamment de toute erreur de mesure, une proportion ou un taux observés sur une population comprend toujours une marge d'erreur (puisque cette proportion est une estimation de la vraie proportion qui elle est inconnue). On constate empiriquement que cette marge d'erreur est, entre autres, une fonction décroissante de la taille de cette population. De quelle manière varie cette incertitude, en fonction de la taille de la population, et quelle confiance peut-on accorder à une telle proportion ou taux ? Autrement dit, quelle précision ou quel crédit peut-on donner à une telle mesure ? La réponse à cette question est évidemment essentielle dans le cadre de l'analyse que l'on fera à partir de cette observation.
2. corrélativement, les différences entre les valeurs des indices démographiques calculés pour deux populations de tailles différentes reflètent-elles des situations ou des comportements résolument distincts, ou sont-elles plutôt le fait d'un processus stochastique lié à la petite taille des populations soumises au risque et/ou aux faibles taux d'occurrence des événements démographiques, le tout se traduisant par un nombre limité d'individus ou d'événements observés ? En d'autres termes, la valeur de cet indice pour la première population est-elle significativement supérieure ou inférieure à celle calculée pour la seconde population ?

On se référera éventuellement à l'annexe 1 pour un rappel succinct des principales distributions statistiques utilisées en démographie.

⁴ Des considérations théoriques permettent d'expliquer cette constatation.

2. Applicabilité des distributions statistiques discrètes usuelles aux phénomènes démographiques

Au sein d'une population de taille n , deux types de variables sont donc prises en considération :

- des événements qui peuvent se réaliser avec une probabilité θ et qui modifient les structures ;
- des structures qui partitionnent la population (selon l'âge, le sexe, l'état civil, ...).

Lorsqu'on étudie cette population, on peut dès lors se poser deux types de questions :

- le premier porte sur les événements : on peut, par exemple, se demander quelle est la probabilité qu'un nombre X d'individus de la population vivent un type d'événement.
- le second type de questions concerne les structures : on peut se demander quelle est la probabilité que Y individus (choisis au hasard au sein de la population supposée parfaitement connue) appartiennent à une structure donnée S dans la population.

Le premier type de questions est un problème de nature statistique étant donné qu'on ignore les paramètres qui définissent la loi de probabilité d'occurrence des événements étudiés et qu'il faut tout d'abord estimer ces paramètres (ou les tester). Le second type de questions est un problème de probabilité, puisqu'il n'y a aucun problème d'estimation. Cet article ne considère que le premier type de problème.

2.1. La distribution hypergéométrique

Celle-ci modélise le problème de l'extraction (non exhaustif) de n individus d'une population de taille N dont R individus ont une caractéristique donnée et $N-R$ ont la caractéristique opposée.

Si on traite un problème de sondage ou d'échantillonnage au sein d'une population finie, il peut être justifié de faire usage de la distribution hypergéométrique. En prenant un modèle d'extraction d'urne, si on connaît R (ou $N-R$), il n'y a aucun problème d'estimation et on est en mesure de calculer immédiatement la probabilité recherchée (en recourant éventuellement à l'approximation binomiale si la taille de la population est grande). Si par contre on ne connaît pas exactement R (ou $N-R$), alors on se trouve face à un problème d'estimation à partir d'un échantillon extrait d'une population de taille finie et ce problème se situe hors de notre contexte, puisque nous avons retenu l'hypothèse de Keyfitz.

Par contre, si on s'intéresse à l'occurrence d'événements, on se tournera vers la loi binomiale dont on essaiera d'estimer le paramètre : même dans le cas de petites populations, la distribution hypergéométrique n'est pas appropriée pour traiter ce type de problème, puisque, avant leur réalisation, on ignore combien d'événements vont se produire : en reprenant le modèle d'urne, on ne connaît exactement pas R , le seul paramètre connu est N , et on se trouve face à un problème d'estimation.

Dès lors, dans notre problématique, l'usage de la distribution hypergéométrique (exacte ou approchée par une distribution binomiale de paramètres connus) constituerait une erreur méthodologique.

2.2. La distribution binomiale

Celle-ci repose sur deux hypothèses (indépendance des événements⁵ et constance de la probabilité) que nous examinons dans le contexte des phénomènes démographiques.

⁵ L'indépendance de deux événements A et B s'exprime par $P[A] = P[B] = P[A \cap B]$ ou encore $P[A|B] = P[A]$ lorsque $P[B] \neq 0$. Cela signifie que la réalisation de B n'a aucun effet sur la réalisation de A .

Première hypothèse : au cours d'une période de temps donnée, on suppose que tous les individus d'une population sont soumis à un risque qui se caractérise par l'occurrence d'un type d'événement. Chaque individu peut vivre cet événement et ceci indépendamment des autres individus. Par exemple, au cours de la période de temps donnée, des individus vont décéder et ces décès auront lieu indépendamment les uns des autres : a priori, il n'y a pas de raison de penser que le décès d'une personne A ait une influence quelconque sur le décès d'une personne B.

On pourrait objecter que dans certaines situations, telles que en présence d'épidémies, de catastrophes, ... il y a un nombre plus important d'événements et donc que le comportement d'un individu A a une influence sur celui d'un individu B et donc que l'on perd l'indépendance. En réalité, il n'en n'est rien, mais dans ces cas, c'est le niveau du risque qui est beaucoup plus élevé. Par exemple, rien ne permet d'affirmer que si un individu A décède cela aura une influence sur la propension d'un autre individu B à mourir. Ce raisonnement peut être généralisé aux autres risques démographiques (par exemple la natalité) et il est donc justifié d'accepter l'hypothèse d'indépendance.

Seconde hypothèse : d'emblée, cette hypothèse pourrait être mise en doute : par exemple dans le cas de la mortalité, on peut objecter que tous les individus ne sont pas affectés d'une probabilité de décès identique, celle-ci dépendant d'un ensemble de facteurs individuels (sexe, âge, état matrimonial, profession, ...) et comportementaux. Cependant, en l'absence d'informations détaillées, on est forcé d'admettre que le niveau du risque étudié est le même pour toute la population soumise à ce risque. Seule une analyse plus détaillée, par exemple de la mortalité selon l'âge, le sexe, ... permettrait de mettre en évidence ces différences, mais elle se situerait dans le cadre de l'analyse explicative et requerrait de ce fait des informations plus détaillées.

Notons cependant que si on stratifie une population en classes homogènes pour le risque considéré (ce qui est théoriquement possible si on travaille, par exemple, pour la mortalité) alors, au sein d'une même strate, la probabilité de subir le risque est constante, par définition. Dans ce cas, un autre problème se pose et qui est lié à un problème d'ajustement. Supposons que l'on ait stratifié la population en k classes caractérisées par les probabilités $(\theta_1, \dots, \theta_k)$. Au sein d'une strate i , le nombre d'événements observés X_i est une variable aléatoire binomiale $Bi(n_i, \theta_i)$ où n_i est la taille de la population de la strate i . Le problème qui se pose alors est de trouver la distribution de la somme des X_i , car la somme de variables binomiales n'est pas, en général, une binomiale. Toutefois, nous avons constaté sur des données réelles que X suit approximativement une loi binomiale $Bi(n, \theta)$ où n est la somme des n_i et θ est la moyenne des θ_i pondérée par les n_i . Ceci s'explique par la règle de Bayes en considérant un mélange de k populations, chacune ayant un poids \dots . Dans ce cas, la probabilité qu'un individu choisi au hasard subisse le risque considéré est alors θ si on admet l'hypothèse d'indépendance ; on retombe alors sur les conditions conduisant à une variable binomiale. Néanmoins, X n'est pas une binomiale de paramètres n, θ et cet argument ne fait que de justifier intuitivement la qualité de l'ajustement.

2.3. La distribution de Poisson

Bien qu'elle soit plus restrictive que la distribution binomiale (puisque 3 hypothèses sont formulées au lieu de 2), la distribution de Poisson peut être utile pour modéliser l'occurrence des événements démographiques, à condition de se placer sur l'axe du temps⁶. Par ailleurs, on rappelle à l'annexe 2 que le loi de Poisson est un cas limite de la loi binomiale (« lorsque n devient grand et θ est petit », ce qui est souvent le cas en démographie).

Première hypothèse : à condition de découper l'intervalle de temps de manière suffisamment fine, il est légitime de penser que deux phénomènes démographiques ou plus ne

⁶ Dans le cadre d'une analyse spatiale, on pourra faire usage de la distribution de Poisson et se placer dans le plan, les événements étant alors des points.

peuvent survenir exactement au même instant (sauf peut-être dans le cas de la nuptialité où nécessairement deux individus se marient au même instant : mais dans ce cas, il n'y a qu'un seul mariage !). Cette hypothèse est donc admissible.

Deuxième hypothèse : celle-ci est assez naturelle : en se restreignant au terme du premier ordre, on peut admettre que le nombre d'événements qui se produisent pendant un intervalle de temps assez court est directement proportionnel à la longueur de cet intervalle de temps.

Troisième hypothèse : l'hypothèse d'indépendance a été discutée et admise au point précédent.

Chacune de ces distributions présente des avantages et des inconvénients : la loi de Poisson est plus facile à manipuler, elle est additive et sa moyenne coïncide avec son espérance, mais elle est plus restrictive. La loi binomiale est plus lourde au niveau des calculs et n'est pas additive, mais en contrepartie, elle est moins restrictive et semble (légèrement) mieux répondre à la réalité.

Aussi, par souci de généralité, le caractère binomial du nombre des événements est celui qui sera retenu dans les lignes qui suivent.

3. Construction des intervalles de confiance et tests

Dans la problématique qui est la nôtre, il s'agit donc :

1. de déterminer des intervalles de confiance pour une proportion θ_0 observée et à partir de là, de calculer des intervalles de confiance sur le nombre d'événements qui définissent θ_0 ;
2. de mettre en évidence, ou de tester, la nullité, ou la non nullité, de la différence entre deux proportions.

Ces deux points font l'objet des paragraphes 3.2 et 3.3.

Avant de les développer, il est essentiel de faire un bref rappel et de repréciser les notions de test et d'intervalles de confiance. La confusion qui règne entre les deux concepts est à l'origine d'erreurs fréquentes. L'élaboration des tests et la recherche d'intervalles de confiance sont deux problèmes relativement proches, mais toutefois différents et ils répondent à des problèmes statistiques distincts.

3.1. Préliminaires

3.1.1. Intervalles de confiance

La construction d'un intervalle de confiance répond à un problème d'estimation : on observe une valeur d'une variable aléatoire dont la distribution dépend d'un paramètre⁷, on se fixe une probabilité élevée (appelée le niveau de confiance qui est fixé souvent à 90 ou 95 %), et à partir de là, on détermine une région qui contient le paramètre observé avec cette probabilité. Si le paramètre est réel, on peut construire des intervalles bilatères⁸ ou unilatères. Les premiers sont de la forme $]T ; T'[,$ tandis que les seconds sont de la forme $[0 ; T[$ ou $]T' ;$

⁷ Celui-ci peut être réel ou vectoriel.

⁸ A ce propos, signalons que l'on trouve dans la littérature deux définitions différentes pour les intervalles de confiance bilatères pour un paramètre θ au niveau $1-\alpha$. On trouve, selon les auteurs, soit $P[T_1 < p < T_2] = 1-\alpha$ (i), soit $P[T_1 \leq p \leq T_2] = 1-\alpha$ (ii). Par exemple, Bickel et al. ainsi que Dagnelie utilisent (ii), tandis que Mood et al. ainsi que Saporta utilisent (i). Sans le mentionner explicitement, Bickel et al. utilisent toutefois la définition (i) lorsqu'ils se placent dans le cas discret. Cette différence (minime) entre les deux définitions n'a pas d'importance dans le cas de distributions continues, mais, par contre, elle en a une dans le cas discret. La distinction est souvent passée sous silence du fait que l'on travaille essentiellement avec des variables continues. Nous retiendrons pour notre part la définition (i) et nous ne considérerons dans l'exposé théorique que les intervalles de type bilatères.

1] où T et T' sont deux statistiques telles que, dans le premier cas, $P[T < \theta < T'] = 1 - \alpha$ ou, dans le second cas, $P[\theta < T] = 1 - \alpha$ et $P[\theta > T'] = 1 - \alpha$. Nous ne parlerons pas ici des intervalles unilatères, les calculs étant analogues à ceux effectués pour les intervalles bilatères.

3.1.2. Tests d'hypothèses

La construction des tests d'hypothèses répond à la recherche d'une règle de décision. Le problème peut se résumer comme suit : on vise à éprouver une hypothèse, H_0 , dont on a de bonnes raisons de penser qu'elle est valide (i.e. elle reste admise jusqu'à preuve du contraire), contre une autre hypothèse, H_1 . Ces deux hypothèses sont des affirmations relatives au(x) paramètre(s) (cas paramétrique) ou à la distribution de ceux-ci (cas non paramétrique) et sont mutuellement exclusives. On mène une expérience et en confrontant les observations à la règle de décision, on peut soit rejeter H_0 au profit de H_1 , ou au contraire ne pas rejeter (ou encore à conserver), H_0 au profit de H_1 . Dans chaque cas, on commet une erreur : si H_0 est vraie et qu'elle est rejetée, on commet alors une certaine erreur α fixée a priori (dite erreur de type I), tandis que si H_1 est vraie et qu'elle est rejetée, on commet alors une certaine erreur β (dite erreur de type II). Les deux hypothèses ne jouent donc pas des rôles symétriques. Les erreurs commises lors de la décision sont reprises ci-dessous.

Décision	H_0 vraie	H_1 vraie
non rejet de H_0	$1 - \alpha$	β
rejet de H_1	α	$1 - \beta$

L'idéal serait évidemment de trouver un test qui minimiserait à la fois α et β , mais malheureusement, un tel test n'existe pas : pour un échantillon de taille n donnée, si on diminue α , alors automatiquement, β augmente.

3.1.3. Construction de tests

Celle-ci peut s'effectuer de diverses manières : on citera la méthode du rapport de vraisemblance généralisée, la méthode du rapport de vraisemblance monotone, la méthode basée sur la statistique de Rao, la méthode basée sur la statistique de Wald.

Une autre méthode se base sur les intervalles de confiance. Bien que souvent assez facile à mettre en œuvre, il faut insister sur le fait qu'elle n'est qu'une méthode parmi d'autres et par conséquent, la recherche d'intervalles de confiance n'est pas équivalente à celle de la construction de tests. Dans les ouvrages élémentaires, on ne mentionne souvent que cette dernière méthode et cette lacune induit souvent la confusion entre les deux concepts.

3.1.4. Tests basés sur les intervalles de confiance

Dans le problème de test sur une proportion, on considère θ un paramètre réel (inconnu) qui caractérise (partiellement ou entièrement) une distribution statistique et θ_0 un réel connu ; tester l'hypothèse $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ au seuil de signification α peut se faire en déterminant un intervalle de confiance, au niveau de confiance $(1 - \alpha)$, noté $]T_1 ; T_2[$ et à vérifier que $\theta_0 \in]T_1 ; T_2[$. Si c'est le cas, on ne rejette pas H_0 , tandis que dans le cas contraire, on rejette H_0 .

De même, dans le problème de test sur une différence de deux proportions, si on considère deux populations indépendantes (de tailles respectives n_1 et n_2) telles que chacun des individus de ces deux populations a des probabilités respectivement θ_1 et θ_2 de subir un risque donné, tester $H_0 : \theta_1 = \theta_2$ contre $H_1 : \theta_1 > \theta_2$, (ou $H_1' : \theta_1 < \theta_2$ ou encore $H_1'' : \theta_1 \neq \theta_2$) peut se faire en trouvant un intervalle de confiance $]T_1' ; T_2'[$ au niveau $(1 - \alpha)$ et à vérifier ensuite que $\theta_1 - \theta_2 \in]T_1' ; T_2'[$. Si c'est le cas, on ne rejettera pas H_0 . On rejettera respectivement H_1

lorsque $\theta_1 - \theta_2 \leq T'_1$, H_1' lorsque $\theta_1 - \theta_2 \geq T'_2$ et H_1'' lorsque l'une des deux inégalités est satisfaite $\theta_1 - \theta_2 \leq T'_1$ ou $\theta_1 - \theta_2 \geq T'_2$.

3.2. Intervalles de confiance et tests sur le paramètre θ d'une distribution binomiale

3.2.1. Intervalles de confiance bilatères et tests exacts (sans biais)

Soit X une variable aléatoire $Bi(n, \theta)$. θ est un réel compris entre 0 et 1 et est inconnu. On souhaite construire un intervalle de confiance bilatère exact (ou sans biais) au niveau de confiance $(1 - \alpha) \%$ pour θ . Il s'agit donc de trouver deux statistiques T_1 et T_2 [c'est-à-dire des fonctions de X et des paramètres connus (ici, il n'y a que n)] telles que : $P[T_1 < \theta < T_2] = 1 - \alpha$. L'intervalle de confiance sera donc $]T_1 ; T_2[$. Par intervalle de confiance sans biais, nous entendons que l'égalité précédente est vérifiée.

Ce problème n'est pas évident lorsque la distribution de la statistique suffisante (c'est-à-dire X ici) est discrète, en particulier la distribution binomiale dans le cas que nous considérons.

Il existe plusieurs méthodes conduisant à leur élaboration. Nous en décrivons deux : la méthode « pivotale » et une méthode « alternative ».

Méthode « pivotale »

Cette méthode est celle qui est la plus souvent utilisée et qui est décrite dans les ouvrages « classiques » de statistique élémentaire, mais elle ne conduit donc pas, dans le cas discret, à des intervalles sans biais. Nous en rappelons son principe : il s'agit de trouver une variable aléatoire Y qui ne dépende que des observations et du paramètre (éventuellement vectoriel) à estimer et dont la loi est connue exactement ou asymptotiquement (notée W) et ne dépend pas du paramètre. Il suffit d'écrire alors $P\left[W_{\frac{\alpha}{2}} < Y < W_{1-\frac{\alpha}{2}}\right]$ et ensuite d'isoler le paramètre d'intérêt. Y porte le nom de « variable de pivot » exact ou approximatif, selon le cas. L'existence des quantiles exacts $W_{\frac{\alpha}{2}}$ et $W_{1-\frac{\alpha}{2}}$ requiert évidemment la continuité de la distribution. Elle est donc inapplicable dans le cas qui nous intéresse ici.

Méthode « alternative »

Nous notons X_{obs} la valeur observée de la variable X qui dépend d'un paramètre réel θ inconnu. Soit alors θ une valeur comprise entre 0 et 1. On pose alors $P_\theta [X \leq x]$ la probabilité qu'une variable binomiale $Bi(n, \theta)$ soit inférieure ou égale à x .

Proposition

Sous les hypothèses et avec les notations précédentes,

a) il existe deux nombres T_1 et T_2 uniques et compris entre 0 et 1 tels que

$$P_{T_1}[X \geq X_{\text{obs}}] = \frac{\alpha}{2} \quad (\text{i})$$

$$P_{T_2}[X \leq X_{\text{obs}}] = \frac{\alpha}{2} \quad (\text{ii})$$

b) $]T_1, T_2[$ est un intervalle de confiance exact au niveau de confiance α .

Démonstration :

- a) On montre ii) , la démonstration de i) se faisant de manière similaire.

Considérons la fonction suivante : $h : \begin{cases} [0,1] \rightarrow [0,1] \\ \theta \rightarrow h(\theta) = P_\theta[X \leq X_{\text{obs}}] \end{cases}$

Cette fonction est continue sur $[0, 1]$, dérivable sur $]0, 1[$ (ceci est trivial) et strictement décroissante (ceci est nettement moins trivial). En outre, $h(0) = 1$ et $h(1) = 0$. Le théorème des valeurs intermédiaires implique alors le résultat. Notons que T_2 ne dépend que de X_{obs} et de n , et est donc bien une statistique.

- b) Il suffit de montrer que $P[\theta \geq T_2] = \frac{\alpha}{2}$. On montre de la même façon que $P[\theta \leq T_1] = \frac{\alpha}{2}$, ce qui permet de conclure. Mais cela résulte de l'identité entre les deux événements suivants grâce à la définition de h et de T_2 : $\{\theta \geq T_2\} \Leftrightarrow \{P_\theta[X \leq X_{\text{obs}}] \leq \frac{\alpha}{2}\}$

Cette proposition permet alors de construire numériquement des intervalles de confiance exacts : T_1 et T_2 s'obtiennent en résolvant numériquement les équations suivantes : $P_{T_2}[X \leq X_{\text{obs}}] = \frac{\alpha}{2}$ et $P_{T_1}[X \geq X_{\text{obs}}] = \frac{\alpha}{2}$, soit dans notre cas précis :

$$\sum_{k=X_{\text{obs}}}^{k=n} C_n^k \theta^k (1-\theta)^{n-k} - \frac{\alpha}{2} = 0 \text{ et } \sum_{k=0}^{k=X_{\text{obs}}} C_n^k \theta^k (1-\theta)^{n-k} - \frac{\alpha}{2} = 0.$$

La méthode de Newton-Raphson converge rapidement vers les résultats. Par ailleurs, on peut faire usage d'une relation exacte basée sur la distribution de Fisher : au seuil α , la borne inférieure de l'intervalle bilatère est donnée par $\frac{X_{\text{obs}}}{X_{\text{obs}} + (n - X_{\text{obs}} + 1)F_{2(n - X_{\text{obs}} + 1), 2X_{\text{obs}}, 1 - \frac{\alpha}{2}}}$, tandis que

la borne supérieure est donnée par $\frac{(X_{\text{obs}} + 1)F_{2(X_{\text{obs}} + 1), 2(n - X_{\text{obs}}), 1 - \frac{\alpha}{2}}}{n - X_{\text{obs}} + (X_{\text{obs}} + 1)F_{2(X_{\text{obs}} + 1), 2(n - X_{\text{obs}}), 1 - \frac{\alpha}{2}}}$.

Le calcul est très facile à réaliser, même avec une machine à calculer scientifique de poche disposant des fonctions statistiques.

Toutefois pour de « grandes » valeurs de n (pratiquement n supérieur à 2000), les méthodes numériques exactes peuvent ne pas fournir de solution satisfaisante. En optimisant les calculs, grâce aux relations de récurrence entre les termes successifs des sommes, on est confronté à des problèmes d'instabilité numérique : ceci se produit par exemple lorsqu'on multiplie des puissances très élevées de quantités proches de 0 par des grands nombres. Il en résulte une instabilité numérique dont les effets sont catastrophiques. Les relations basées sur la distribution de Fisher peuvent également ne pas fournir de résultats valables lorsque le nombre de degrés de liberté est trop grand.

Dans ce cas seulement, il est justifié de se tourner vers des approximations.

3.2.2. Approximations

Les théorèmes et propriétés relatifs aux convergences de variables aléatoires (cf. annexe 2) permettent, sous certaines hypothèses, de faire usage d'approximations. Les propriétés asymptotiques des distributions « lorsque n devient grand » permettent de remplacer, dans certains cas, ces distributions par des distributions normales.

On trouve dans la littérature un ensemble de conditions plus ou moins restrictives permettant ces approximations. Classiquement, on préconise de faire appel à l'approximation de Poisson lorsque, θ étant suffisamment petit, n est grand et que le produit $n\theta$ reste constant. Les différents auteurs ne sont pas unanimes sur les valeurs de n et de θ . On trouve les inégalités suivantes : $n \geq 20$ et $\theta < 0.05$, $n \geq 100$ et $\theta < 0.05$ ou encore $n \geq 100$ et $n\theta < 10$. En ce qui concerne l'approximation normale, on trouve également diverses conditions sur n et θ : on préconise $n\theta$ et $n(1-\theta) \geq 5$, $n\theta(1-\theta) > 9$, $n\theta(1-\theta) \geq 10$, $n > 9 \max \left\{ \frac{\theta}{(1-\theta)}; \frac{(1-\theta)}{\theta} \right\}$ et encore

$$\theta \pm 2 \sqrt{\frac{\theta(1-\theta)}{n}} \text{ dans } [0, 1].$$

En fait, toutes ces conditions sont valables : tout dépend de la précision sur le résultat que l'on souhaite obtenir (cf. Leemis and Trivedi (1996)), et, au risque de se répéter, il est évident que si on peut éviter de recourir à une approximation en faisant usage des lois exactes (ce qui constitue de moins en moins un obstacle), c'est de loin la solution préférable. Les graphiques repris en annexe permettent de visualiser la qualité des approximations.

Approximation poissonnienne : intervalle de confiance sans biais pour le paramètre d'une loi de Poisson

Si X une variable aléatoire de Poisson de paramètre λ , alors, on montre que $P[X \leq k] = P[Y > 2\lambda]$, où Y est une variable aléatoire $\chi^2_{2(k+1)}$. Dès lors, si on dispose d'un échantillon de taille n , et de moyenne \bar{x} , on peut en déduire l'intervalle de confiance bilatère exact pour λ au niveau $(1 - \alpha)$: $\frac{1}{2n} \chi^2_{2n\bar{x}; \frac{\alpha}{2}} < \lambda < \frac{1}{2n} \chi^2_{2n(\bar{x}+1); 1-\frac{\alpha}{2}}$ où $\chi^2_{k;\alpha}$ est le α -quantile d'une χ^2 à k degrés de liberté. Par conséquent, cette relation facile à mettre en œuvre permet de trouver facilement un intervalle de confiance exact pour le paramètre d'intérêt.

Approximations normales : intervalles de confiance pour la moyenne

Approximation usuelle

Puisqu'on sait que X est asymptotiquement $N\left(\theta, \sqrt{\frac{\theta(1-\theta)}{n}}\right)$, on peut déduire facilement un intervalle de confiance (assez grossier) pour θ . L'intervalle dont les bornes sont données par $\hat{\theta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$ (où $\hat{\theta} = \frac{x}{n}$ et $z_{\frac{\alpha}{2}}$ est le quantile au niveau $\alpha/2$ de la distribution normale centrée réduite), est donc un intervalle de confiance pour θ au niveau α .

Approximation « améliorée »

En résolvant l'équation du second degré suivante : $(\hat{\theta} - \theta)^2 = z_{\frac{\alpha}{2}}^2 \frac{\theta(1-\theta)}{n}$ dans laquelle θ est le paramètre à estimer et $\hat{\theta}$ est la proportion observée, on déduit les bornes suivantes

$$\frac{\left(2\hat{\theta} + \frac{z_{\alpha}^2}{n}\right) \pm \sqrt{\frac{z_{\alpha}^4}{n^2} + 4\hat{\theta}\frac{z_{\alpha}^2}{n}(1-\hat{\theta})}}{2\left(1 + \frac{z_{\alpha}^2}{n}\right)}. \text{ Elles sont un peu plus précises que les précédentes puisqu'on}$$

on ne fait pas appel à une estimation de la variance.

Approximation tenant compte de la correction de continuité

L'approximation normale est meilleure lorsque l'on tient compte d'une correction de continuité qui permet de prendre en compte le fait que l'on remplace une distribution discrète par une distribution continue. Au seuil de signification α , les bornes de l'intervalle de confiance sont

données par $\hat{\theta} + \frac{1}{2n} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$. Cet intervalle se déduit de la formule proposée par Yates (cf. infra). Le terme correctif tend évidemment vers 0 lorsque n tend vers l'infini.

Approximation arcsinus

Pour être complet, signalons enfin l'existence d'une approximation basée sur la fonction arcsinus. On peut montrer que, pour de petites valeurs de θ , $\arcsin(\sqrt{X})$ est asymptotiquement

$N\left(\arcsin(\sqrt{\theta}), \frac{1}{4n}\right)$. Partant de ce résultat, il est alors facile de construire un intervalle de confiance pour la moyenne θ : il est donné par les bornes suivantes $\sin^2\left(\arcsin\sqrt{\hat{\theta}} \pm z_{\frac{\alpha}{2}} \frac{1}{2\sqrt{n}}\right)$.

Son principal avantage réside dans le fait que ses bornes seront toujours comprises entre 0 et 1 et donc, cette méthode est appropriée lorsqu'on travaille avec des proportions ou des taux.

3.2.3. Comparaison des différents intervalles de confiance

On pourra se référer à l'article de Leemis et Trivedi (1996) mentionné plus haut.

Pour les différentes méthodes, il est possible de calculer l'erreur commise en utilisant les intervalles de confiance obtenus par les différentes méthodes approchées qui ont été présentées ci-dessus. Si on note par $[l_j; u_j]$ les intervalles trouvés par la méthode « j » où

- $j = P$ dans le cas de l'intervalle de Poisson ;
- $j = N$ dans le cas de l'intervalle de normal simple ;
- $j = C$ dans le cas de l'intervalle de normal tenant compte de la correction de continuité ;
- $j = M$ dans le cas de l'intervalle de normal « amélioré » ;
- $j = A$ dans le cas de l'intervalle trouvé au moyen de l'approximation arcsinus ;
- $j = E$ dans le cas de l'intervalle exact,

alors, deux mesures possibles de l'erreur commise, en remplaçant la vraie distribution par une approximation, sont données par $\max \left\{ \left| \frac{l_E - l_j}{l_E} \right|; \left| \frac{u_E - u_j}{u_E} \right| \right\}$ et $\max \{ |l_E - l_j|; |u_E - u_j| \}$. Il s'agit d'une erreur relative et d'une erreur absolue.

Nous donnons en annexe 2 une représentation graphique de ces deux mesures pour les intervalles de confiance calculés par les différentes méthodes pour différentes valeurs de n et pour différentes valeurs de θ . On trouvera, par ailleurs, différents intervalles de confiance calculés par les différentes méthodes pour $n = 10$, $n = 50$, $n = 100$, pour les valeurs de θ valant 0.1, 0.02 et 0.01.

De manière générale, plus la taille n augmente, plus faibles sont les erreurs et donc meilleures sont les approximations. Dans le cas des approximations normales, ceci est d'autant plus vrai que θ s'approche de la valeur de 0.5. A l'opposé, plus θ s'approche de 0, les approximations de Poisson sont les meilleures.

En ce qui concerne les petites valeurs de n , les choses se présentent différemment : pour les valeurs de θ inférieures à 0.1, l'approximation de Poisson est la meilleure tandis que les approximations basées sur la normale sont dans l'ensemble assez médiocres, sauf peut-être celle qui fait appel à la transformation arc sinus) : elles peuvent conduire à des bornes inférieures qui peuvent être négatives. C'est le cas par exemple, lorsque θ est inférieur à 0.3 et $n = 10$ ou $\theta < 0.07$ et $n = 50$. Mais de toutes manières, dans cette situation, l'utilisation de l'intervalle exact est facile et elle évite de la sorte ces problèmes d'approximation.

Lorsque n est « grand », et que θ n'est pas trop petit, par exemple, $n \geq 500$ et $\theta \geq 0.1$, on pourra recourir aux approximations normales. Les erreurs commises sont faibles et on examinera les courbes présentées ci-dessus pour déterminer les erreurs (relative et absolue) que l'on commet.

3.3. Intervalles de confiance et tests sur la différence des paramètres $\delta = \theta_1 - \theta_2$ de deux distributions binomiales

Supposons que l'on souhaite comparer deux proportions et que l'on veuille donc tester

$$H_0 : \theta_1 = \theta_2 \text{ contre } H_1 : \theta_1 \neq \theta_2 \text{ ou } H_0 : \theta_1 = \theta_2 \text{ contre } H_1 : \theta_1 > \theta_2$$

Si X_1 et X_2 sont deux distributions binomiales indépendantes notées respectivement $Bi(n_1, \theta_1)$ et $Bi(n_2, \theta_2)$, alors le problème peut se ramener à celui de la détermination d'intervalles de confiance pour la différence $\delta = \theta_1 - \theta_2$. Les estimateurs de maximum de vraisemblance de θ_1

et θ_2 sont respectivement $\hat{\theta}_1 = \frac{x_1}{n_1}$ et $\hat{\theta}_2 = \frac{x_2}{n_2}$. Dès lors, par le théorème de Zehna, l'estimateur

de maximum de vraisemblance de δ est donné par $\hat{\delta} = \frac{x_1}{n_1} - \frac{x_2}{n_2}$. On a que $E(\hat{\delta}) = \delta$ et

$$V(\hat{\delta}) = \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}. \text{ Il suffirait donc de construire les intervalles de confiance exact}$$

pour répondre au problème posé, mais ce problème est compliqué. Actuellement, à notre connaissance, il n'existe pas de solution à ce problème.

3.3.1. Intervalles de confiance bilatères basés sur des approximations

La distribution exacte de δ n'est pas connue exactement car c'est une combinaison linéaire de binomiales de paramètres n_1 et n_2 a priori différents. Il n'est donc pas possible de trouver, en général, un intervalle de confiance exact par la méthode pivotale rappelée au

paragraphe 3.2.1., ni même par la méthode qualifiée d'alternative. Actuellement, la construction des intervalles de confiance reste basée sur des approximations. Nous en distinguerons trois qui sont toutes basées sur les propriétés asymptotiques des binomiales et sur les propriétés de la loi normale.

Intervalle de confiance normal simple

Selon Dagnelie (STAT1 p. 278), l'approximation de la loi binomiale par la loi normale permet d'écrire, au niveau de confiance $(1 - \alpha) \%$, l'intervalle de confiance sur δ donné par les

$$\text{bornes } \hat{\delta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}.$$

Intervalle de confiance corrigé (Yates)

Il est possible d'améliorer la précision de l'intervalle précédent en tenant compte d'une correction de continuité qui prend en compte le fait que l'on fait usage d'une distribution discrète à la place d'une distribution continue.

$$\text{Yates propose l'intervalle suivant : } \hat{\delta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{p}_2(\hat{\theta}-\hat{\theta}_2)}{n_2}} + \left\{ \frac{1}{2n_1} + \frac{1}{2n_2} \right\}$$

Intervalle de confiance « amélioré » (Hauck et Anderson)

En utilisant l'estimateur non biaisé de la variance de la différence δ (et ceci se justifie d'autant plus que les populations sont de petite taille), Hauck et Anderson proposent, quant à

$$\text{eux, l'intervalle amélioré suivant : } \hat{\delta} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{p}_2(\hat{\theta}-\hat{\theta}_2)}{n_2}} + \left\{ \frac{1}{2\min(n_1, n_2)} \right\}$$

Les relations proposées précédemment permettent de construire immédiatement un intervalle de confiance approximatif et donc de tester la différence des proportions.

Le problème de la recherche d'un intervalle de confiance exact pour la différence entre deux paramètres de binomiales, problème simple en première apparence, est donc loin d'être évident et actuellement, ce problème n'a pas encore trouvé, à notre connaissance, de réponse. Il y a encore une recherche à poursuivre dans ce domaine. En attendant cette réponse, nous préconisons de recourir à une des deux approximations prenant en compte une correction de continuité (Yates ou Hauck et Anderson).

3.3.2. Test exact

Si on souhaite seulement tester l'hypothèse $H_0 : \theta_1 = \theta_2$ contre $H_1 : \theta_1 \neq \theta_2$, alors, pour de « petites valeurs de n_1 et n_2 , il existe une solution.

Dans un article récent, Berger (1996) propose un test qui répond à la question.

Soit le test $H_0 : \theta_1 = \theta_2$ contre $H_1 : \theta_1 < \theta_2$.

Si X et Y sont deux binomiales $Bi(m, \theta_1)$ et $Bi(n, \theta_2)$,

$$\text{Si } Z(x, y) = \frac{\hat{\theta}_2 - \hat{\theta}_1}{\sqrt{\hat{\theta}(1-\hat{\theta})\left(\frac{1}{m} + \frac{1}{n}\right)}} \text{ où } \hat{\theta} = \frac{x+y}{m+n},$$

Alors la p - valeur du test est égale à $p_z = \sup \sum_{(a,b) \in R_z(x,y)} C_m^a p^a (1-p)^{m-a} C_n^b p^b (1-p)^{n-b}$,

où le sup est pris sur l'ensemble des p compris entre 0 et 1,

et $R_z(x,y) = \{(a,b) \in \{0,1,\dots,m\} \times \{0,1,\dots,n\} \text{ tels que } Z(a,b) > Z(x,y)\}$.

On rejette H_0 ssi p_z est inférieure au seuil α fixé au départ.

Partant de ce résultat, il suffit de tester :

1) $H_0 : \theta_1 = \theta_2$ contre $H_1 : \theta_1 < \theta_2$. On déterminera une p -valeur p_{z1}

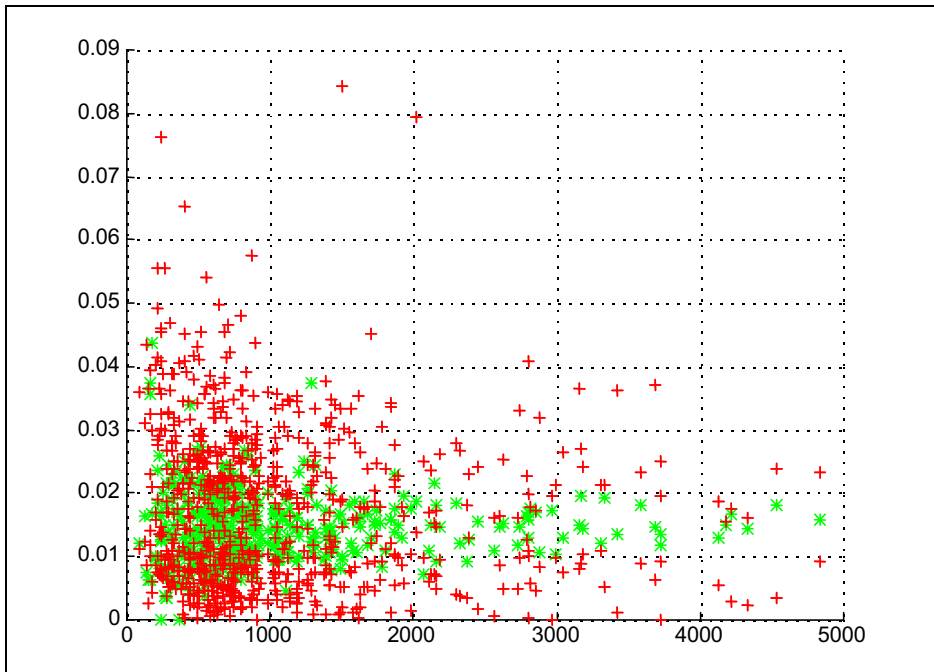
2) $H_0 : \theta_2 = \theta_1$ contre $H_1 : \theta_2 < \theta_1$. On déterminera une p -valeur p_{z2}

On rejettera l'hypothèse ssi $\min(p_{z1}, p_{z2})$ est inférieure au seuil α .

4. Application à des données réelles

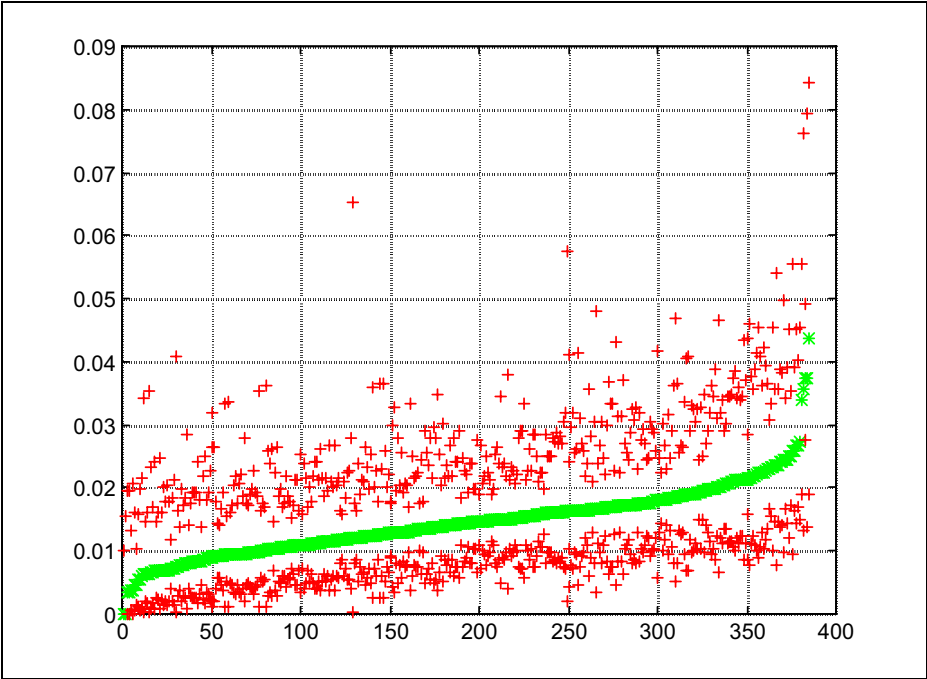
4.1. Intervalles de confiance sur des taux bruts de mortalité

Pour chacun des taux bruts de mortalité observés en 1910 sur 385 communes, la figure suivante donne en fonction de la taille de la population le taux brut de mortalité observé (indiqués par *) et les bornes des intervalles de confiance exactes (indiquées par +) au niveau de confiance de 95 %.



On notera la tendance à la décroissance des intervalles de confiance en fonction de la taille de la population soumise au risque : en moyenne, plus cette taille est petite, plus l'intervalle correspondant à une longueur importante et inversement. Par ailleurs, ces taux bruts de mortalité fluctuent autour d'une valeur proche de 0.015.

En ordonnant les taux bruts de mortalité par valeurs croissantes, les mêmes données se présentent comme suit :



Il est intéressant de constater que pour des taux bruts de mortalité proches, les intervalles de confiance peuvent être très différents. Ceci s’explique par l’effet de la taille de la population.

4.2. Intervalles de confiance sur le nombre d’événements

On se propose dans l’exemple suivant de calculer les taux bruts de mortalité de quelques quartiers de la ville de Namur entre 1991 et 1996 et d’en déduire un intervalle de confiance pour le nombre de décès. Nous retenons les quartiers de Wierde, de Jambes centre et de Loyers, ainsi que l’ensemble de la ville de Namur.

Pour ces quartiers le tableau ci-dessous donne la population moyenne, le nombre de décès qui ont été observés sur la période, le taux brut de mortalité observé, les bornes de l’intervalle de confiance exact pour le taux brut de mortalité et le nombre de décès correspondant à ces bornes.

Quartier	Pop. moyenne	Nbre décès	TBM	Borne inf.	Borne sup.	Décès inf.	Décès sup.
Wierde	443	24	0.05418	0.03502	0.07954	15 (15)	35 (33)
Jambes centre	5811	428	0.07365	0.06707	0.08067	390 (399)	469 (467)
Loyers	1452	38	0.02617	0.01859	0.03575	27 (26)	52 (50)
Namur ⁹ (total)	104502	6430	0.06153	0.06004	0.06305	6274 (6277)	6589 (6582)

Les valeurs numériques obtenues avec les méthodes faisant appel aux approximations conduisent à des résultats différents. On a repris entre parenthèses les nombres de décès

⁹ Pour l’ensemble de la ville de Namur, l’intervalle exact ne peut être calculé, car on se heurte à des problèmes numériques. Vu la faible valeur du taux brut de mortalité, et la taille élevée, on se tournera vers l’approximation de Poisson qui présente l’avantage de permettre un calcul d’intervalle exact.

obtenus en faisant usage de l'approximation normale simple. On peut apprécier le gain de précision qui est obtenu en recourant à l'intervalle exact.

Dans le cas de Wierde, la prise en compte de la correction de continuité mène à un nombre de décès compris entre 15 et 34, tandis que l'approximation basée sur la transformation arcsinus donne 16 et 34 décès et l'approximation de Poisson fournit 15 et 36 décès. Ces différences absolues sont faibles (une unité en plus ou en moins), mais l'erreur relative commise est relativement importante (de l'ordre de 3 à 4 %).

4.3. Intervalles de confiance sur la différence de taux bruts de mortalité

La question qui était posée au départ était la suivante : « est-ce que les niveaux de mortalité sont (significativement) différents entre deux quartiers ? ». Les deux méthodes exposées ci-dessus (test exact et intervalle de confiance) permettent maintenant répondre à cette question.

Pour ce faire, nous déterminerons un intervalle de confiance à 95 % sur la différence entre les deux taux bruts de mortalité observés. La taille des populations étant assez grande, le calcul se fait en utilisant la méthode de Hauck et Anderson. Si l'intervalle trouvé contient la valeur nulle, alors, la conclusion est qu'il n'y a pas de différence significative, donc égalité, entre les deux taux.

En reprenant les mêmes quartiers de la ville de Namur que précédemment, on détermine les intervalles de confiance suivants :

Quartier	Wierde	Jambes centre	Loyers	Namur (total)
Wierde	- -			
Jambes centre	[-0.0405 0.0038]	- -		
Loyers	[0.0065 0.0518]	[0.0372 0.0584]	- -	
Namur (total)	[-0.0274 0.0149]	[0.0053 0.0191]	[-0.0434 -0.0267]	- -

Les cellules en gras indiquent des couples de quartiers qui présentent des taux bruts de mortalité significativement non différents, puisque les intervalles de confiance contiennent la valeur 0, tandis que les autres présentent des différences significatives.

La comparaison des taux bruts de mortalité de Wierde ($n = 443$, $d = 24$, $\theta = 0.05418$) et de Loyers ($n = 1452$, $d = 38$ et $\theta = 0.02617$) à l'aide du test proposé par Berger conduit aux résultats numériques suivants :

$$H_0 : \text{TBM}_{\text{Wierde}} = \text{TBM}_{\text{Loyers}}$$

$$H_1 : \text{TBM}_{\text{Wierde}} < \text{TBM}_{\text{Loyers}}$$

$$p_{z1} = 0.9976$$

$$H_0 : \text{TBM}_{\text{Loyers}} = \text{TBM}_{\text{Wierde}}$$

$$H_1 : \text{TBM}_{\text{Loyers}} < \text{TBM}_{\text{Wierde}}$$

$$p_{z2} = 0.0044$$

Les conclusions du test sont donc : pour le premier test, on rejette H_1 avec quasi certitude (et ceci était prévisible) tandis que pour le second test, on rejette H_0 pour tout seuil inférieur à 0.0044. En d'autres termes, on accepte avec un risque très minime de se tromper l'hypothèse de différence non nulle entre les deux taux. (On retrouve ainsi le résultat obtenu approximativement au moyen du test basé sur l'intervalle de confiance, ce qui est rassurant !).

Il y a donc bien une différence significative entre les deux taux bruts de mortalité. Les deux quartiers ont donc des mortalités significativement différentes, et dans une étape

ultérieure, il pourrait être utile d'en rechercher la cause, ce qui nécessiterait évidemment des informations supplémentaires.

Remarque

Il importe de manipuler les égalités trouvées avec beaucoup de précautions et de prendre attention de tirer des conclusions abusives : ces égalités sont des relations statistiques et donc toujours entachées d'une erreur. Ce ne sont pas des égalités stricto sensu et ne vérifient pas toutes les propriétés usuelles. En particulier, la transitivité n'est pas satisfaite (i.e. si $a = b$ et si $b = c$ alors, on ne peut affirmer que $a = c$). Ainsi, Wierde et Jambes centre ont le même taux brut de mortalité et il en est de même pour Wierde et Namur (total), mais Jambes centre et Namur (total) ont des taux bruts de mortalité significativement différents.

5. Conclusions

A partir des mesures de comptage des occurrences d'un phénomène démographique régi par une loi binomiale de paramètre θ inconnu, nous avons déterminé un intervalle de confiance exact sur le paramètre de cette loi : chaque mesure de comptage peut donc être affectée d'une précision, et ce pour n'importe quelle taille de population.

En outre, nous pouvons comparer le niveau d'un phénomène démographique sur deux espaces différents grâce à des méthodes exactes ou approchées.

Les méthodes de calcul présentées sont facilement et directement applicables à partir de données réelles. Seule la contrainte numérique peut encore éventuellement constituer un obstacle, auquel cas, il est opportun de se tourner vers les approximations : pour des événements rares (θ petit), il est recommandé de travailler avec l'approximation poissonnienne (pour laquelle il est facile de trouver des intervalles exacts), tandis que pour des événements non rares, on optera pour les intervalles déterminés à l'aide de l'approximation normale tenant compte d'une correction de continuité.

Pour la comparaison de deux paramètres de binomiales, les méthodes approchées permettent de procéder aux tests usuels. La méthode proposée par Hauck et Anderson présente l'avantage de donner des résultats acceptables pour des valeurs faibles de θ_1 et θ_2 . Selon la précision requise, on fera appel à l'une ou à l'autre méthode. Un test exact a été proposé par Berger, mais il se révèle assez lourd à mettre en oeuvre pour les situations mettant en jeu des populations de taille « importante ».

Les méthodes développées ci-dessus sont prometteuses et ouvrent une voie de généralisation possible à d'autres paramètres démographiques usuels tels l'espérance de vie, l'indice conjoncturel de fécondité, Ces méthodes doivent encore être développées, mais nous pensons au stade actuel qu'elles constitueront une étape qui donnera à la démographie la possibilité de se libérer de la contrainte des « grands nombres » et d'analyser statistiquement l'étude des populations de tailles restreintes.

BIBLIOGRAPHIE

- BERGER, R.L. (1996), More Powerful Tests From Confidence Interval p Values, *The American Statistician*, 50, 4, 314 - 318.
- BICKEL, P.J. and DOKSUM, K.A. (1977), *Mathematical Statistics*, Holden - Day, Inc, Oakland, California, 492 p.
- DAGNELIE, P. (1992), *Statistique Théorique et Appliquée*, tome 1, Presses Agronomiques de Gembloux, 492 p.
- HAUCK, W. W. and ANDERSON, S. (1986), A Comparison of Large-Sample Confidence Interval Methods For The Difference of Two Binomial Probabilities Distributions, *The American Statistician*, 40, 1, 318 - 322.
- LEEMIS, L. M. and TRIVEDI, K. S. (1996), A Comparison of Approximate Intervals Estimators for the Bernoulli Parameter, *The American Statistician*, 50, 1, 63 - 68.
- MOOD, A.M., GRAYBILL, F.A. and BOES, D.C. (1974), *Introduction to the Theory of Statistics*, Mac Graw - Hill, 564 p.
- SAPORTA, G. (1990), *Probabilités Analyse des données et Statistique*, Technip, PARIS, 493 p.
- TASSI, P. (1989), *Méthodes statistiques*, Economica, PARIS, 474 p.

ANNEXE 1 : Distributions statistiques usuelles

1.1. Distribution binaire (ou indicatrice, ou de Bernoulli)

C'est la loi d'une variable aléatoire X qui peut prendre deux valeurs 1 (succès) ou 0 (échec) avec les probabilités respectives θ et $1 - \theta$ ($0 \leq \theta \leq 1$). On montre immédiatement que l'espérance, $E(X) = \theta$ et la variance, $V(X) = \theta(1 - \theta)$. Cette distribution est appropriée pour décrire et étudier les proportions et les taux. Par exemple, dans une population infinie, dont certains individus possèdent une caractéristique donnée, un individu tiré au hasard possède la caractéristique ($X = 1$) ou il ne la possède pas ($X = 0$).

1.2. Distribution binomiale

On répète dans des conditions identiques (succès ou échec et probabilité de succès constante au fil des épreuves) n expériences indépendantes de Bernoulli. En d'autres termes, on observe n fois l'occurrence (succès) ou la non occurrence (échec) d'un événement pour lequel la probabilité d'obtention d'un succès est constante à chaque expérience et vaut θ . Ce processus est connu sous le nom de « schéma de Bernoulli ».

Si on note par X la variable aléatoire « nombre de succès au cours des n expériences », alors X est la somme de n variables de Bernoulli X_i indépendantes et de même paramètre θ . Par définition, X est une variable binomiale notée $Bi(n, p)$ et on sait alors que la probabilité d'observer k succès au cours des n expériences est donnée par $P[X = k] = C_n^k \theta^k (1 - \theta)^{n-k}$ (avec $k = 0, 1, \dots, n$ et $C_n^k = \frac{n!}{k!(n-k)!}$ où

$k! = k(k-1)\dots 2$ et par convention $0! = 1$). Par ailleurs, on montre que $E(X) = n\theta$ et $V(X) = n\theta(1 - \theta)$. Par exemple, le tirage au hasard, dans une population infinie, de n individus possédant chacun l'un des deux caractères opposés et le prélèvement de n individus dans une population finie, avec remise au fur et à mesure des prélèvements, correspondent à ce schéma.

La distribution binomiale suppose donc deux hypothèses :

- l'indépendance des expériences donnant chacune lieu à deux résultats exclusifs (succès ou échec);
- la constance de la probabilité de « succès » au cours des n expériences.

1.3. Distribution hypergéométrique

Soit une urne contenant N billes dont R sont rouges et B sont bleues, et supposons que l'on procède à un tirage sans remise¹⁰ de n billes parmi ces $R+B$ billes. On considère alors la variable aléatoire X qui prend pour valeurs le nombre de billes rouges extraites parmi les n . Le tirage s'effectuant sans remise, il

n'y a pas indépendance entre les extractions et on a alors que $P[X = k] = \frac{C_R^k C_B^{n-k}}{C_N^n}$. On a alors $E(X) =$

$n \frac{R}{N}$ et $V(X) = n \frac{R}{N} \frac{B}{N} \frac{N-n}{N-1}$. Lorsque N est grand que le rapport $\frac{R}{N}$ tend vers une constante θ , alors on

montre facilement que la loi hypergéométrique tend vers la loi binomiale de paramètres N et θ .

1.4. Distribution de Poisson

Si X désigne le nombre de survenances d'un phénomène aléatoire dans un hypervolume V d'un espace vectoriel tel que :

- la probabilité que plus d'une survenance se produise dans un hypervolume dont la mesure tend vers 0, est négligeable devant la probabilité qu'une seule survenance se produise dans cet hypervolume

¹⁰ Si par contre le tirage se fait avec remise, on se retrouve dans les conditions d'application des hypothèses qui définissent la loi binomiale de paramètres $R+B$, θ où θ est connu et vaut $\frac{R}{N}$.

- la probabilité qu'une seule survenance se produise dans un hypervolume V donné est au premier ordre proportionnel à la mesure de cet hypervolume,
- il y a indépendance entre les survenances,

Alors, $P[X = k] = \frac{e^{-\lambda} \lambda^k}{k!}$ où λ est une constante positive qui dépend de la mesure de V . X suit alors

une loi de Poisson et on note dans ce cas $X \sim P_0(\lambda)$. X modélise l'occurrence des événements « rares ». On a immédiatement $E(X) = \lambda$ et $V(X) = \lambda$.

Un des principaux avantages de cette distribution réside dans le fait que sa moyenne coïncide exactement avec sa variance et que celles-ci sont indépendantes de la taille de la population étudiée. De plus, la somme de deux variables indépendantes de Poisson de paramètres respectifs λ_1 et λ_2 est une variable de Poisson de paramètre $\lambda_1 + \lambda_2$, ce qui n'est pas le cas pour la somme de deux binomiales de paramètres θ_1 et θ_2 différents.

1.5. Distribution normale (ou de Gauss)

Lorsqu'on effectue une mesure d'une quantité qui peut prendre une valeur réelle quelconque, le résultat de cette mesure est toujours entaché d'une erreur. Par exemple, la mesure de la position d'un point, ou de la longueur, ou de la masse d'un corps physique, ... est toujours connue avec une précision finie. Au niveau macroscopique, tous les instruments de mesure ont une limite à la précision qu'ils peuvent fournir, tandis qu'au niveau microscopique, ceci découle immédiatement du principe d'incertitude d'Heisenberg. Si on répète un très grand nombre d'opérations de mesure, on observe que ces mesures se distribuent autour d'une valeur centrale et présentent une dispersion plus ou moins importante autour de cette valeur centrale. Ce constat permet de définir une variable gaussienne. On dira qu'une variable aléatoire X suit une distribution normale de paramètres μ ($\in \mathbb{R}$) et σ ($\in \mathbb{R}^+$) lorsque $\forall a \in$

$\mathbb{R} : P[X \leq a] = \int_{-\infty}^a f(t) dt$ où $f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}$. On note alors $X \sim N(\mu; \sigma^2)$. On a alors

immédiatement $E(X) = \mu$ et $\text{Var}(X) = \sigma^2$. Cette distribution joue un rôle capital en probabilité et en statistique.

On consultera, par exemple Dagnelie, Saporta ainsi que Mood et al. pour une description des propriétés de ces distributions.

ANNEXE 2 : Convergences des variables aléatoires

Les convergences des variables aléatoires présentent un grand intérêt à la fois théorique et pratique, car lorsque la taille de l'échantillon est « grande », elles permettent, sous certaines conditions assez larges, de faire abstraction de la distribution de ces variables aléatoires et de considérer que celles-ci sont régies par une loi normale. Sans entrer dans les détails, nous rappelons ci-dessous les principaux théorèmes, ainsi que les principales propriétés qu'on en déduit.

2.1. Loi faible des grands nombres

Soient X_i ($i = 1, \dots, n$) n variables aléatoires indépendantes de moyennes respectives μ_i , et de variances respectives σ_i^2 , supposées finies.

Si $\frac{1}{n} \sum_{i=1}^n \mu_i \xrightarrow[n \rightarrow \infty]{} \mu$ et $\frac{1}{n^2} \sum_{i=1}^n \sigma_i^2 \xrightarrow[n \rightarrow \infty]{} 0$, alors $\frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \xrightarrow{P} \mu$.

(C'est la convergence dite « faible » ou en probabilité, c'est-à-dire $\forall \varepsilon > 0, \forall \eta > 0 \exists n_0 \in \mathbb{N}$ tel que $\forall n > n_0 : P[|\bar{X}_n - \mu| > \varepsilon] < \eta$)

2.2. Loi forte des grands nombres

Soient X_i ($i = 1, \dots, n$) n variables aléatoires indépendantes de moyennes respectives μ_i , et de variances respectives σ_i^2 .

Si les μ_i et les σ_i^2 sont telles que $\frac{1}{n} \sum_{i=1}^n \mu_i \xrightarrow[n \rightarrow +\infty]{} \mu$ et $\sum_{i=1}^n \frac{\sigma_i^2}{i^2}$ est une série convergente,

alors $\frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \xrightarrow[p.s.]{} \mu$. (C'est la convergence forte ou presque sûre, ce qui signifie que $P\left[\left\{\omega \text{ tels que } \lim_{n \rightarrow \infty} \bar{X}_n(\omega) \neq \mu(\omega)\right\}\right] = 0$: l'ensemble des points pour lesquels l'égalité n'est pas satisfaite est de mesure nulle)

2.3. Théorème central - limite

Soient X_i ($i = 1, \dots, n$) n variables aléatoires indépendantes de même loi (éventuellement inconnue)

de moyenne μ et de variance σ^2 , alors $Y_n = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^n X_i - n\mu}{\sigma} \xrightarrow{L} N(0;1)$ (C'est la convergence en loi : elle exprime ici la convergence, en tout point x réel, des fonctions de répartition F_{Y_n} vers la fonction de répartition d'une variable aléatoire normale réduite).

Propriétés

Il découle immédiatement de ces théorèmes précédents les propriétés suivantes :

- Si X_n est une suite de variables aléatoires binomiales $Bi(n, \theta)$ telles que le produit $n\theta \rightarrow \lambda$ avec $\lambda > 0$ lorsque $n \rightarrow +\infty$, alors $X_n \xrightarrow{L} P_0(\lambda)$.
- Si X_n est une suite de variables aléatoires binomiales $Bi(n, \theta)$ alors $\frac{X_n - n\theta}{\sqrt{n\theta(1-\theta)}} \xrightarrow{L} N(0;1)$.
- Si X_n est une suite de variables aléatoires de Poisson $P0(\lambda)$ telles que $\lambda n \rightarrow \lambda$, alors $\frac{X_n - \lambda}{\sqrt{\lambda}} \xrightarrow{L} N(0;1)$.

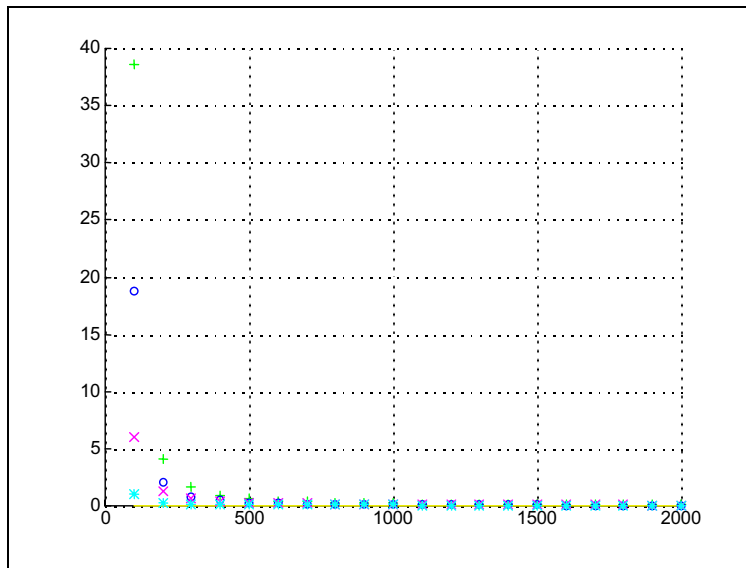
L'intérêt pratique de ces propriétés est évident : dès que la taille d'un échantillon est « suffisamment » grande, et pour peu que les conditions d'application soient satisfaites, on peut approximer la loi binomiale, soit par la loi de Poisson, soit par la loi normale. De plus, si n est assez grand, on peut sous certaines conditions approximer la loi de Poisson par une loi normale.

ANNEXE 3 : Qualité des approximations - intervalles de confiance exact et approches

Variation de $\max \left\{ \left| \frac{I_E - I_j}{I_E} \right|; \left| \frac{u_E - u_j}{u_E} \right| \right\}$ pour $\theta = 0,01$ et n variant de 100 à 2000 par pas de 100.

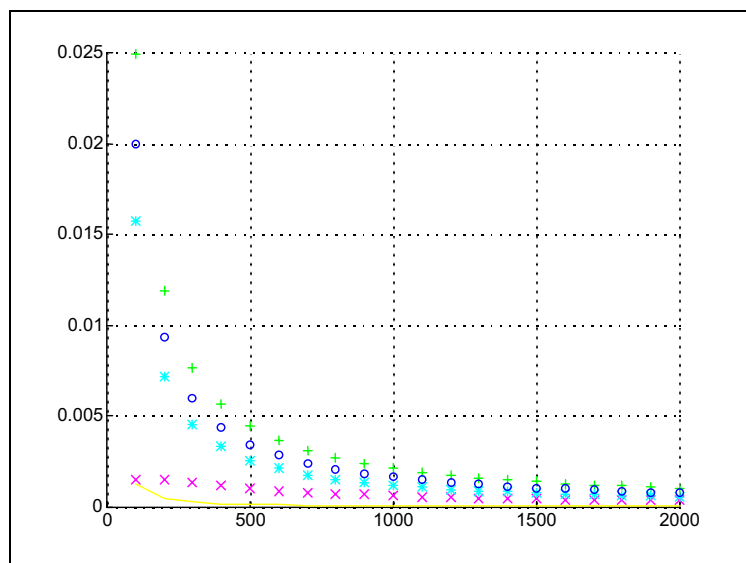
$\alpha = 5\%$

(En trait continu : intervalle de Poisson, + : intervalle normal simple o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus).



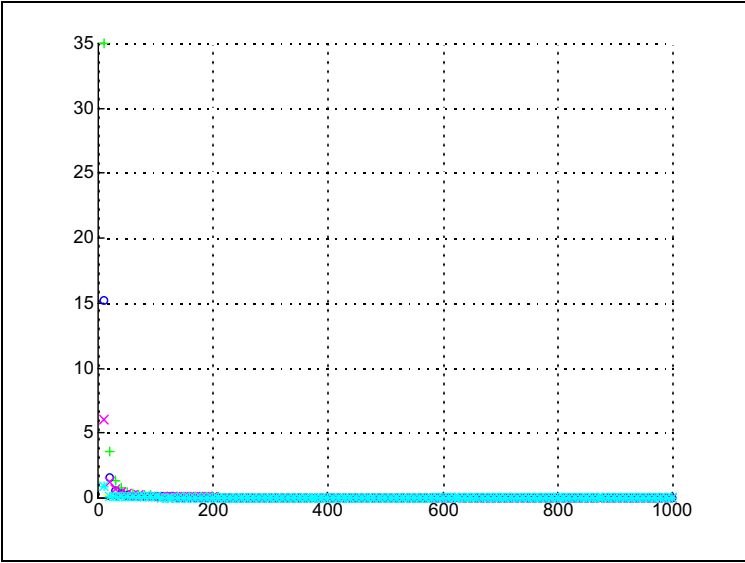
Variation de $\max \left\{ |l_E - l_j|; |u_E - u_j| \right\}$ pour $\theta = 0,01$ et n variant de 100 à 2000 par pas de 100. $\alpha = 5\%$.

(En trait continu : intervalle de Poisson, + : intervalle normal simple o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus).



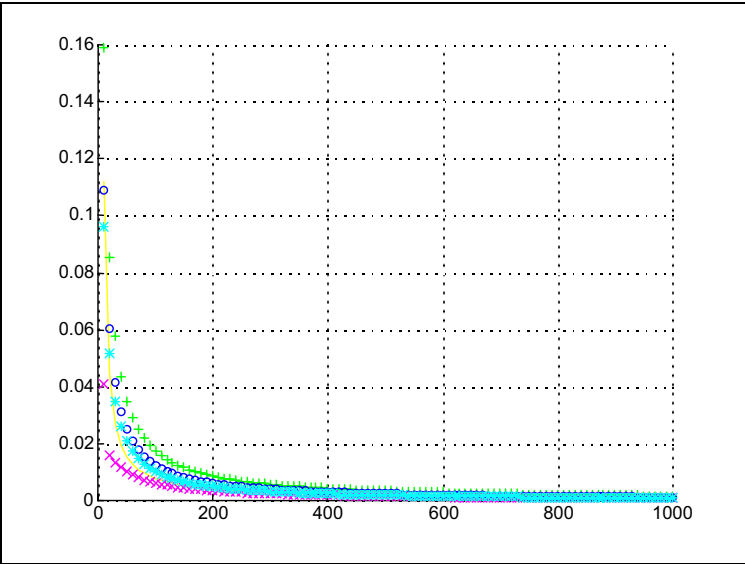
Variation de $\max \left\{ \left| \frac{l_E - l_j}{l_E} \right|; \left| \frac{u_E - u_j}{u_E} \right| \right\}$ pour $\theta = 0,1$ et n variant de 10 à 1000 par pas de 10. $\alpha = 5\%$.

(En trait continu : intervalle de Poisson, + : intervalle normal simple o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus).



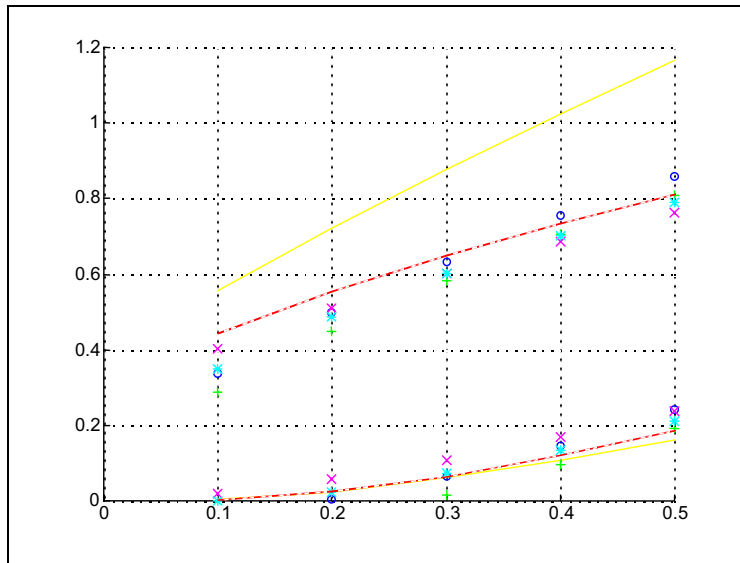
Variation de $\max\{|l_E - l_j|; |u_E - u_j|\}$ pour $\theta = 0,1$ et n variant de 10 à 1000 par pas de 10. $\alpha = 5 \%$.

(En trait continu : intervalle de Poisson, + : intervalle normal simple o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et approximation arcsinus).



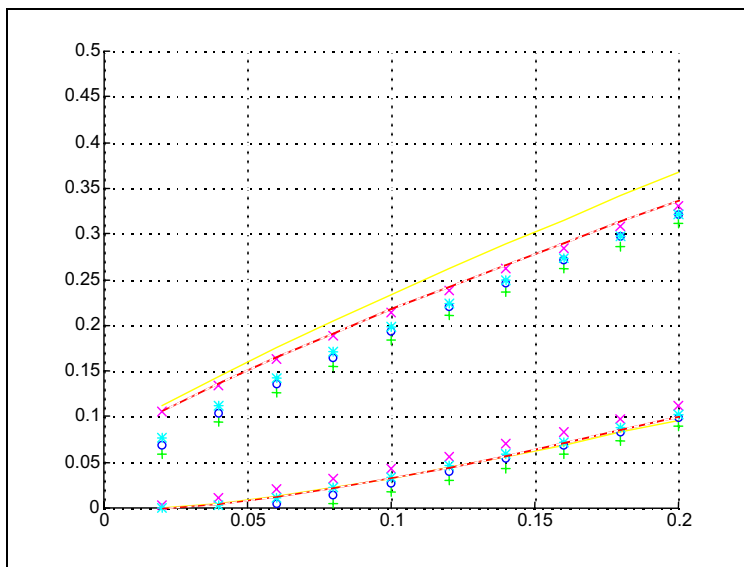
Intervalle de confiance ($\alpha = 5 \%$) pour le paramètre d’une binomiale : méthode exacte et approximations pour $n = 10$ et θ varie de 0.1 à 0.5 par pas de 0.1

En trait pointillés : intervalle exact, en trait continu : intervalle de Poisson, + : intervalle normal simple o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus.



Intervalle de confiance ($\alpha = 5\%$) pour le paramètre d'une binomiale : méthode exacte et approximations pour $n = 50$ et θ varie de 0.02 à 0.2 par pas de 0.02.

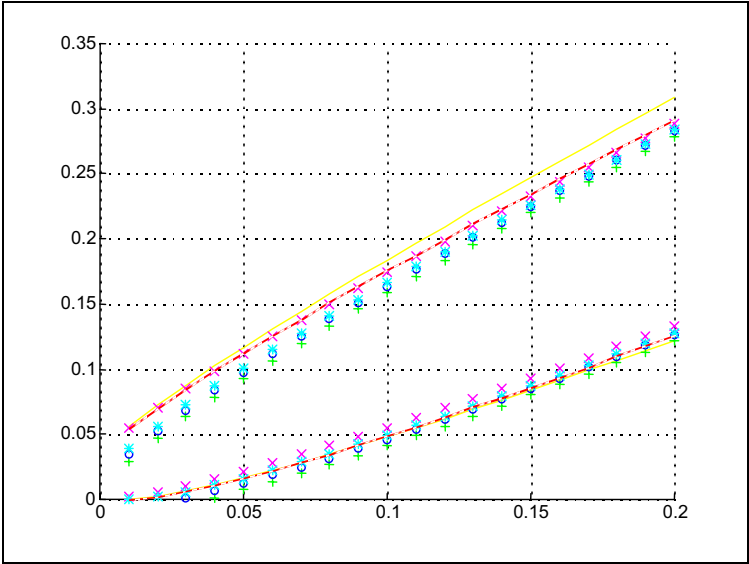
En trait pointillés : intervalle exact, en trait continu : intervalle de Poisson, + : intervalle normal simple
o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus.



Intervalle de confiance ($\alpha = 5\%$) pour le paramètre d'une binomiale : méthode exacte et approximations pour $n = 100$ et θ varie de 0.01 à 0.2 par pas de 0.01

En trait pointillés : intervalle exact, en trait continu : intervalle de Poisson, + : intervalle normal simple

o : intervalle normal avec correction de continuité, x : intervalle normal « amélioré » ; * : intervalle normal et arcsinus.



La démographie post censitaire communale et ses facteurs de qualité

Christophe BERGOUIGNAN

IEDUB – Université de Bordeaux

Les communes qui sont un exemple usuel de petite population, ont une double particularité, non seulement elles ont parfois des effectifs relativement faibles, mais aussi elles occupent des territoires restreints. Cette restriction implique un renouvellement accéléré de telles populations, puisque les échanges entre ensembles ont souvent une fréquence inversement proportionnelle à leurs superficies respectives. De cette double particularité découle parfois des difficultés d'étude statistique des évolutions démographiques dans les petites populations. On signale notamment la difficile observation des événements rares, effet direct de la faiblesse des effectifs, et la forte sensibilité des indices aux transformations de la population, effet combiné de la faiblesse des effectifs et du fort renouvellement.

Le plus souvent en l'absence d'une théorisation précise du lien entre le phénomène étudié et la variabilité statistique de sa mesure, seule la faiblesse des effectifs est retenue comme facteur de variation résiduelle des indices. Néanmoins, il peut exister une forte dépendance entre le phénomène étudié et les facteurs de renouvellement de la population dans laquelle l'étude est menée¹. Il est donc logique d'ajouter à l'effectif des populations, d'autres critères pour expliquer l'ampleur des variations résiduelles des mesures de comportements ou de tendances démographiques. On peut ainsi retenir des indices de répartition de la population selon les facteurs d'hétérogénéité face aux phénomènes mesurés, ou encore des indicateurs de renouvellement de la population.

La population des communes après le recensement : questions et orientations pratiques

Outre la difficulté associée aux petits effectifs et au fort renouvellement, la démographie communale est confrontée, en France et dans de nombreux pays, à la discontinuité de l'information statistique sur les stocks de population. La connaissance directe de l'effectif de la population totale d'une localité n'est en effet possible que lors des recensements généraux de population. Pour pallier ce manque, on peut néanmoins tirer profit des dénombrements fournis par certaines sources administratives continues. Ces derniers ne sont, bien entendu, que des symptômes imparfaits des évolutions démographiques communales, et leur mise en valeur, notamment pour les périodes post censitaires, requiert des traitements adéquats.

Une telle démarche d'actualisation éveille naturellement l'intérêt des décideurs locaux. Au delà du simple aspect informatif, ces derniers sont en effet soucieux de l'adéquation des équipements collectifs aux évolutions démographiques. De même, l'évaluation de l'impact démographique récent des options d'aménagement antérieurement choisies ne les laisse pas indifférents.

Dans la pratique, le démographe est alors mis à contribution pour confronter un chiffre assez précis aux approximations plus ou moins implicites de l'évolution post censitaire de la population produites par les agents municipaux ou certains élus.

¹ Par exemple, la naissance du 3^{ème} enfant intervient très souvent dans un délai très rapide après l'accession à la propriété, c'est à dire après l'installation dans une commune. Ainsi les vagues d'emménagement sont facteurs de fluctuations de grande ampleur des taux bruts de natalité, sans que pour autant la population nouvelle ait une intensité ou un calendrier (en termes d'âge) féconds différents de celle déjà résidente dans la commune ou que la structure d'âge de la population communale étudiée soit beaucoup modifiée.

Lorsqu'il s'agit de communes, souvent périurbaines qui voient se poursuivre la forte extension de leur bâti ayant débuté dans les années 75, 80, les approximations de ces acteurs locaux restituent la tendance générale tout en surestimant la population. En effet, elles procèdent :

- soit d'une multiplication du nombre post censitaire, de résidences principales fiscales ou de nouveaux logements par la taille moyenne des ménages 90, or depuis celle ci a baissé dans 90% des communes de la Gironde²,
- soit (dans le cas le plus fréquent) en prolongeant des tendances passées qui reflètent souvent l'explosion démographique du passé récent.
- Plus problématique est le cas des communes dont la tendance durant les dernières périodes intercensitaires, était à la baisse ou à la stagnation de population ou dont l'environnement économique, politique, social et culturel s'est modifié. On signale notamment la fermeture ou l'installation d'une grande entreprise, ou encore le développement d'un pôle universitaire. Pour ces communes souvent urbaines ou rurales, contrairement aux zones de forte construction, l'incertitude des acteurs locaux est grande. Leurs approximations issues de travaux plus ou moins opaques produisent la plupart du temps des résultats largement contradictoires. C'est évidemment dans ce contexte que les évaluations post censitaires du démographe vont apporter a priori le plus au décideur local.

Lorsque l'on est confronté à de telles demandes, il importe de rendre compte du niveau d'incertitude associé aux évaluations post censitaires produites. Si ce niveau est par trop élevé il convient d'en déterminer l'origine. Pour contrôler les risques issus de l'imperfection des sources administratives et de l'échelon communal, on peut alors construire des indicateurs de qualité qui peuvent s'apparenter à des mesures de dispersion résiduelle. En quantifiant l'ampleur de la variabilité restant inexpliquée par les tendances des sources et les hypothèses choisies, ils renseignent sur la fiabilité des ingrédients et des résultats de la démarche d'évaluation post censitaire de populations communales. Une faible fiabilité est ainsi supposée se manifester par une forte dispersion résiduelle.

De façon évidente l'ampleur de cette dispersion résiduelle est influencée par les sources administratives utilisées dans la démarche et par les hypothèses choisies pour exploiter ces sources. Néanmoins cette dispersion peut dépendre aussi de certaines caractéristiques de la commune étudiée, qui interagissent plus ou moins avec les sources et les hypothèses. Ces caractéristiques sont bien entendu multiples mais l'échelon communal laisse supposer que parmi elles les variables produisant les incertitudes associées à l'étude des petites populations jouent un rôle significatif. Comme on l'a dit, les facteurs d'incertitude associés aux petites populations, sont entre autres la taille de l'effectif étudié et son degré de renouvellement. Aussi, va-t-on chercher à mesurer leur rôle respectif dans la formation des dispersions résiduelles issues de l'irrégularité de certaines séries et de l'inadéquation de certaines hypothèses aux évolutions post censitaires.

Avant d'étudier la fiabilité des résultats et ses facteurs, il convient de présenter les sources et leur mode de mise en valeur communale, ainsi que des méthodes reproductibles et systématiques d'exploitation des données et de calcul des indicateurs de qualité.

Ce travail qui avait été initialement entrepris pour les 27 communes de la Communauté Urbaine de Bordeaux (CUB), se trouve ici élargi et systématisé par l'observation des 542 communes de la Gironde. Parmi elles, certaines³ illustreront les aspects concrets de la démarche d'évaluation post censitaire et de ses conclusions.

² Résultat issus des évaluations post censitaires ici présentées.

³ Celles ci font partie d'un sous ensemble de 189 communes qui a fait l'objet d'un travail pour la Caisse d'Allocations Familiales : « Démographie post censitaire des cantons Girondins et de certaines communes de la Gironde », Christophe Bergouignan, 1997.

1 Les évaluations post censitaires communales

La présentation des sources administratives et de leur mise en valeur pour appréhender la démographie post censitaire communale est nécessaire pour comprendre l'analyse des facteurs conditionnant la fiabilité des résultats qui en découle. On abordera ainsi les principes généraux, les sources et leur potentialités d'utilisation post censitaire, les méthodes systématiques d'exploitation et la formation des indicateurs de qualité.

1.1 Le contexte post censitaire communal

En France, la connaissance objective des stocks de population découle des Recensements Généraux de Population effectués par l'INSEE⁴. Durant les périodes intercensitaires, la variation de ces stocks n'est pas systématiquement mesurée sans être pour autant a priori négligeable. Ainsi, pour la commune de Biganos, le taux de croissance annuel moyen de la population totale pour la période intercensitaire 82-90 est de l'ordre de 3,2%, soit sur 8 ans de 28%. Inversement Le Verdon a perdu 17% de sa population en 8 ans soit 2,3% par an. Au niveau national, on pallie ces incertitudes en actualisant le stock de la population issu du dernier recensement, via les flux naturels et des éléments de connaissance des migrations internationales, leur relative faiblesse rendant acceptable les imprécisions sur ces dernières.

Au niveau communal, on ne peut procéder de la sorte. En effet, les migrations intercommunales sont l'élément prédominant des variations de peuplement. A titre d'exemple, 47% des personnes recensées en 90 à Ste Foy la Grande n'y résidaient pas au recensement de 82. La connaissance de telles migrations intercommunales est quasi inexistante en période post censitaire. Face à une telle incertitude, la pratique courante, présumant de l'inertie des phénomènes à l'origine des variations, consiste à effectuer de simples extrapolations des tendances observées antérieurement. Dans les faits, une telle inertie présente un caractère exceptionnel qui invalide la portée des évaluations produites. Aussi, même dans ses aspects les plus élémentaires, comme la connaissance de l'évolution de l'effectif de la population totale, la démographie post censitaire communale est chargée d'incertitude. C'est pourquoi, pour résoudre de telles interrogations, on peut mettre en valeur certaines sources dites administratives et continues pour approximer les variations post-censitaires de population.

1.2 Les sources administratives continues et leur mise en valeur

Certaines administrations sont amenées dans le cadre de leurs activités (facturation, fichiers d'usagers,...) à dénombrer les effectifs communaux d'entités reflétant potentiellement le nombre de logements occupés ou non et certains effectifs de population. La périodicité annuelle de tels dénombrements permet de qualifier ces sources administratives d'information de sources continues. Ces sources sont a priori très nombreuses⁵, mais seules peu d'entre elles fournissent des dénombrements offrant des séries suffisamment longues et couvrant chaque territoire communal girondin. Il s'agit des permis de construire de logements commencés, des résidences principales fiscales et de leurs personnes à charge et des inscrits sur les listes électorales. Ces sources ne sont pas pour autant parfaites et ne permettent pas de dénombrer directement les effectifs communaux de logement et de population. En effet, les permis de construire avec lesquels on actualise le parc de logements comprennent parfois des

⁴ Institut National de la Statistique et des Études Économiques. Le dernier recensement général de population INSEE date de 90.

⁵ Permis de construire, résidences principales fiscales et leurs personnes à charge, électeurs, compteurs domestiques de gaz, d'électricité et d'eau, assurés sociaux, abonnés téléphoniques, élèves des établissements scolaires, allocataires des prestations familiales.

dépendances non chaînées au logement principal⁶. De même les foyers fiscaux représentent les ménages sous la condition que les personnes ayant deux logements n'interviennent pas le principal et le secondaire⁷. Les personnes à charge sont supposées correspondre à la population de moins de 20 ans⁸, et les électeurs à la population majeure de nationalité française. En fait, certains majeurs restent fiscalement à la charge de leurs parents⁹ et certains majeurs nationaux ne s'inscrivent pas sur les listes électorales ou subsistent inscrits dans une commune où ils ne résident plus, comme c'est fréquent dans les communes rurales¹⁰. De plus, les rythmes d'inscription sur les listes électorales sont souvent associés au calendrier politique, ce qui génère parfois de fortes irrégularités dans les séries de données. C'est pourquoi il est préférable d'utiliser les variations des effectifs dénombrés par ces sources, comme symptômes des évolutions post censitaires des grandeurs démographiques, plutôt que de chercher une approximation immédiate des stocks, fut-elle fondée sur des tentatives de reconstitution très élaborées.

On peut ainsi mesurer ces évolutions par le calcul de taux ou de facteurs d'accroissement¹¹, ou en les visualisant graphiquement. Toutefois, pour qu'une telle approche conserve une certaine rigueur il convient, avant tout travail post censitaire, de mettre en regard les tendances intercensitaires des grandeurs démographiques étudiées et des effectifs issus des sources administratives. Certaines communes girondines vont illustrer l'ensemble de la démarche d'utilisation des sources administratives continues pour l'analyse des évolutions démographiques intercensitaires 82-90, puis post censitaires 90-96. On étudiera ainsi les évolutions des effectifs des différentes sources depuis 82 qui constituera la référence (indice 100).

Il existe des communes pour lesquelles les évolutions 82-90 données par les sources correspondent à celles mesurées pour les grandeurs supposées correspondantes par les deux recensements généraux de population.

⁶ Bien qu'il existe une procédure plus ou moins efficace de chaînage.

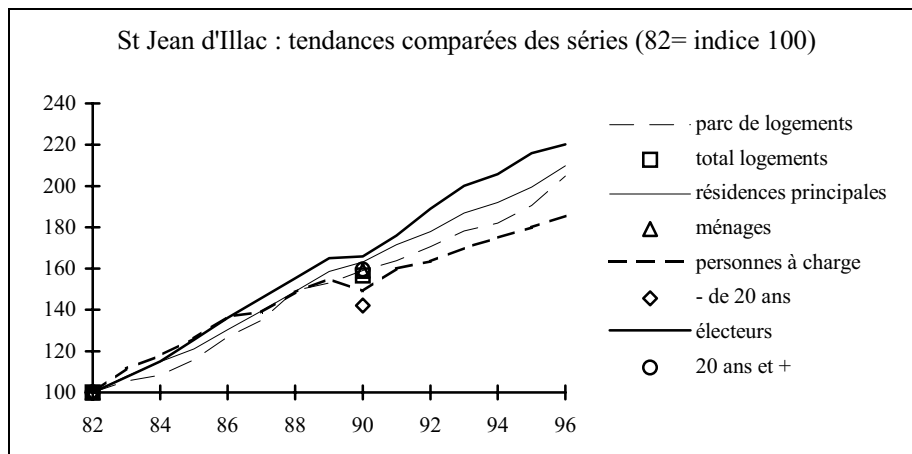
⁷ Il existe en fait certains aspects de la réglementation fiscale qui peuvent inciter à de telles pratiques.

⁸ Associer personnes à charge par foyer fiscal et taille moyenne des ménages peut aussi être envisagé, bien que la complexité d'une telle démarche limite considérablement la vérification de son niveau de fiabilité.

⁹ Ce phénomène est en fait moins massif qu'il n'y paraît en raison du système de déclaration fiscale des pensions alimentaires.

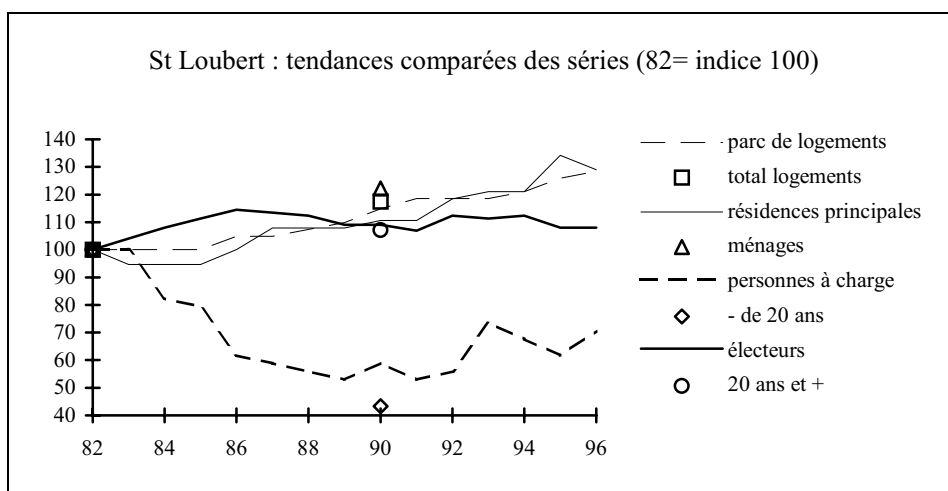
¹⁰ Dans certaines zones rurales on comptait en 90 plus d'inscrits sur les listes électorales que de personnes recensées.

¹¹ Facteur d'accroissement = 1 + taux d'accroissement.



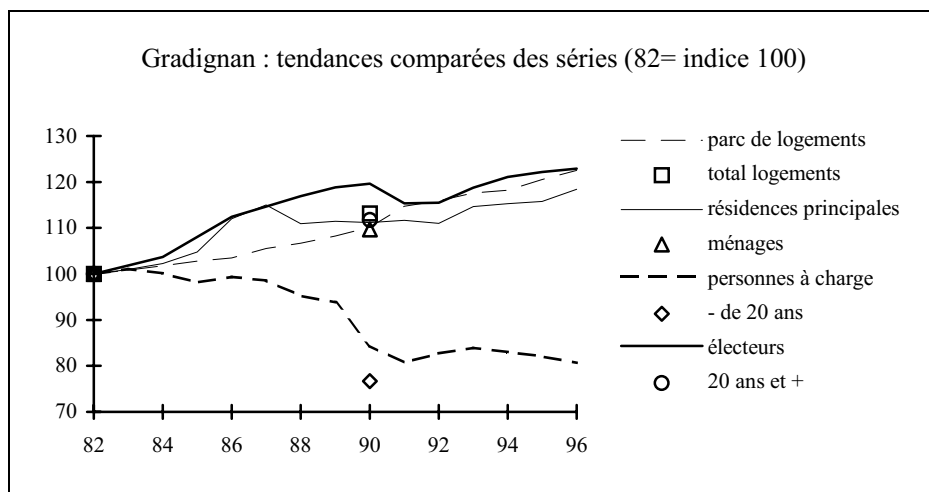
C'est notamment le cas de la commune périurbaine de St Jean d'Illac, qui est passée de 912 habitants en 62 à 3880 en 90, avec une très forte croissance à partir de 68. Ainsi comme les sources le montrent parfaitement, parc de logements, nombre de ménages, populations mineure et majeure y ont crû à un rythme très soutenu entre 82 et 90. Les tendances des effectifs issus des sources depuis 90, font quant eux, apparaître une croissance forte mais dont le rythme fléchit légèrement par rapport à la dernière période intercensitaire. Ce fléchissement paraît affecter principalement le groupe des plus jeunes (personnes à charge). Il s'agit probablement du remplacement incomplet des enfants des ménages installés pendant la très forte croissance 75-82, devenus majeurs et ayant décohabité ou non. L'accroissement de l'écart entre tendance des listes électorales et tendance des personnes à charge pourrait témoigner de la subsistance de certains d'entre eux dans la commune ou de leur inscription malgré qu'ils n'y résident pas.

Pour d'autres communes on constate un écart entre les tendances 82-90 des effectifs issus des sources administratives et les évolutions démographiques données par les recensements. Dans la grande majorité des cas les grandes lignes restent néanmoins tout à fait comparables.



Pour St Loubert, commune rurale caractérisée par une très petite population en baisse depuis 62 (de 129 habitants en 62 à 103 en 90), c'est sans doute la petitesse des effectifs traités (20 personnes à charge en 90), qui cause ce léger écart. Elle peut transformer une différence d'une unité en différence de plusieurs points. Ainsi, bien que la tendance générale soit conservée la croissance du parc de logements et du nombre de ménages et la forte baisse de la population de moins de 20 ans, sont décrites dans une proportion moins exacte que la croissance modérée de la population adulte. Dans cette perspective les évolutions montrées par les effectifs des sources après 90 doivent être interprétées avec précaution. Elles permettent néanmoins de supposer une légère croissance ou une stabilisation de toutes les composantes démographiques étudiées.

Pour Gradignan, commune de banlieue dont la croissance démographique très forte entre 62 et 75 (passage de 6803 à 18691 habitants) a vu son rythme progressivement s'amoindrir depuis, les explications des écarts sont à chercher du côté des spécificités sociologiques communales.



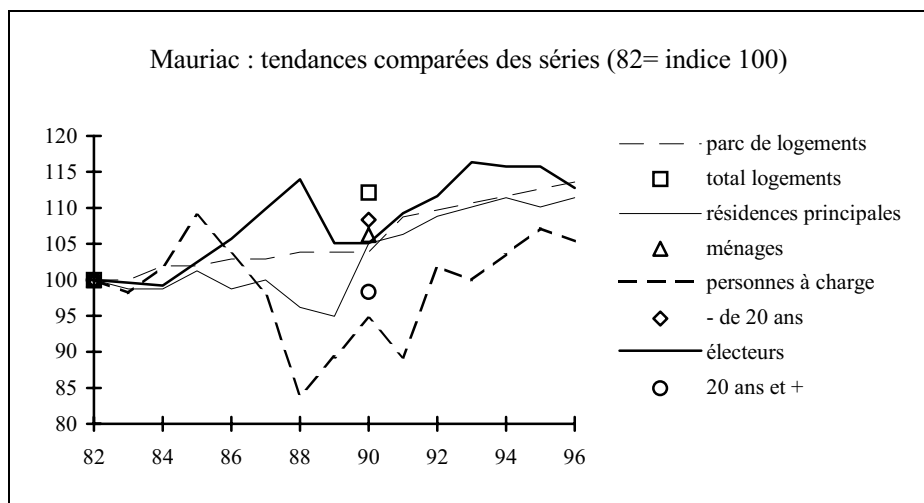
Globalement les sources traitant de personnes (personnes à charge et électeurs) tendent ici à surestimer les hausses et à sous estimer les baisses. La baisse notable du nombre de personnes à charge est en effet inférieure à celle réellement observée pour la population de moins de 20 ans entre 82 et 90. Cette imparfaite illustration de la décohabitation des enfants de familles souvent installées lors de la période de forte croissance 62-75, peut provenir des caractéristiques de ces familles dont le niveau de revenus est en moyenne très supérieur à celui des ménages girondins. Ainsi, il est probable que de nombreux jeunes adultes ayant quitté le domicile familial et la commune restent fiscalement à charge de leurs parents¹². L'excès de croissance du nombre d'électeurs, peut se comprendre par l'inscription des jeunes adultes décohabitants sur les listes électorales de la commune parentale mais aussi par le calendrier électoral (forte mobilisation pour les présidentielles de 88). Si les évolutions démographiques 82-90 de Gradignan sont imparfaitement décrites par les séries issues des sources leur signification globale reste claire, à savoir une faible croissance de la population totale et une diminution du nombre de jeunes. Compte tenu de ces remarques le tassement des tendances

¹² Le système de la pension alimentaire connaissant des plafonds, il est désavantageux pour les familles les plus riches. La rareté globale de ces familles limite ce biais à quelques communes de forte concentration de familles aisées. Gradignan en fait partie puisqu'elle comptait en 90 20% de personnes dont la personne de référence du ménage était cadre ou d'une profession intellectuelle supérieure, contre 11% pour l'ensemble du département.

observé après 90 dans les séries des sources administratives laisse supposer une légère croissance ou une stabilisation de la population avec une moindre diminution de la part des moins de 20 ans.

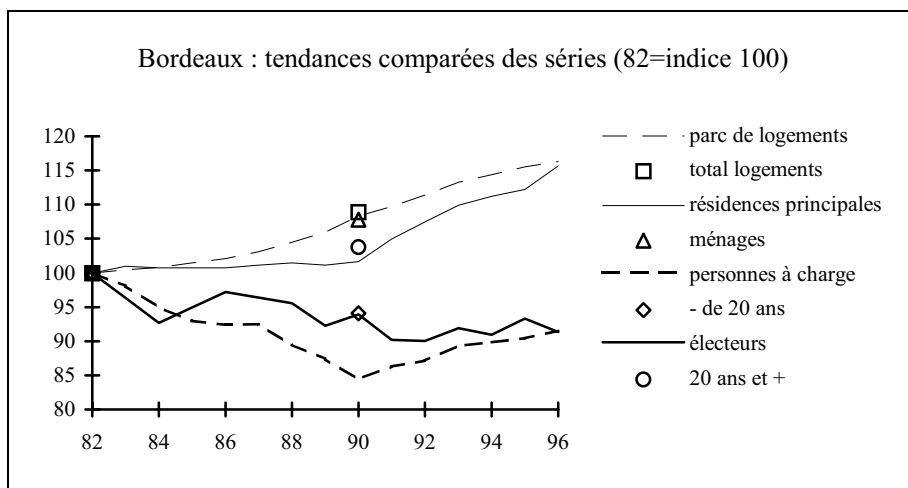
Il existe enfin des communes pour lesquelles on observe un divorce patent entre tendances des effectifs issus des sources administratives et évolutions démographiques mesurées par les derniers recensements.

C'est le cas de Mauriac, petite commune rurale dont la population après avoir fortement diminué entre 62 et 75 (passage de 282 à 219 habitants) croît très modérément après (230 habitants en 90). Ici les séries présentent un aspect très irrégulier à l'exception du parc de logements actualisé par les permis de construire. Une telle instabilité peut provenir de la faiblesse des effectifs associée à un renouvellement important de la population reflet de la rurbanisation de certaines zones de la Gironde (41,3% des personnes recensées en 90 à Mauriac n'y résidaient pas en 82, contre 23,3% seulement à St Loubert). Cette instabilité rend les tendances moins lisibles. Aussi, on constate une baisse du nombre de personnes à charge alors que les recensements indiquent que la population jeune a augmenté entre 82 et 90. De même, on observe une hausse du nombre d'électeurs alors que la population adulte recensée a diminué de 82 à 90. Après 90, les tendances des sources doivent donc être interprétées avec circonspection malgré une moindre irrégularité globale des séries et une certaine cohérence générale laissant supposer une hausse très légère de chaque grandeur étudiée.



C'est le cas aussi de Bordeaux, la capitale régionale, qui après avoir connu une forte baisse de sa population entre 62 et 82 (passage de 278403 à 208159 habitants), connaît une légère reprise depuis (210467 habitants en 90). Ici deux éléments sont de nature à expliquer la différence entre les évolutions démographiques 82-90 mesurées par les recensements et les tendances fournies par les sources administratives pour cette même période. Le premier est le risque d'omissions d'habitants des grandes villes par recensement général de 82 réputé de qualité inférieure. On constate en effet, des tendances toujours inférieures pour les sources notamment fiscales par rapport aux évolutions données par les recensements. Les résidences principales fiscales stagnent presque alors que pour les recensements le nombre de ménages augmente significativement. La baisse du nombre de moins de 20 ans reflet du remplacement des familles avec enfants par d'autres populations est beaucoup moins importante (-6% pour les recensements 82 et 90) que celle du nombre de personnes

fiscalement à charge (-15%). L'écart entre la baisse du nombre d'électeurs et la hausse du nombre d'adultes recensés peut aussi faire intervenir la deuxième explication qui est de nature sociologique. En effet, l'accroissement intercensitaire du nombre d'adultes provient d'une forte poussée de la population étudiante sur la ville (environ +40% en 8 ans). Or les étudiants sont un groupe comptant une proportion significative de non inscrits sur les listes électorales (comme les autres jeunes adultes) mais plus encore une partie importante d'entre eux reste inscrit sur les listes de la commune des parents (surtout si celle-ci est rurale). Les tendances des sources après 90 doivent ainsi être considérées avec prudence. Il est clair que la population adulte et fait nouveau celle des plus jeunes se sont accrues (principalement de 90 à 93) mais le niveau de cette légère croissance reste délicat à évaluer plus précisément. Si le développement de la sous inscription se poursuit, l'accroissement a toutes les chances de se situer entre 0,5 et 1% par an, sinon il est mineur (entre 0 et 0,5% par an). Quoiqu'il en soit des déterminants sociologiques le renouveau de croissance paraît éphémère vu la moindre croissance de toutes les tendances des sources à partir de 94.



1.3 De l'analyse locale des tendances à des méthodes systématiques d'évaluation

Si dans la plupart des communes les tendances des sources administratives 82-90 indiquent des évolutions conformes à celles mesurées par les deux derniers recensements, les coïncidences parfaites restent exceptionnelles alors que des divergences profondes se présentent parfois. Ces dernières résultent souvent de déterminants sociologiques locaux, des faibles effectifs ou du fort renouvellement des populations. Beaucoup plus rarement des incertitudes quant à la qualité des opérations de collecte du recensement peuvent intervenir. Ces facteurs peuvent se cumuler et on peut souvent observer l'association de l'irrégularité des séries et de la divergence. L'influence des calendriers électoraux sur le nombre d'électeurs n'est à ce titre pas la seule manifestation de ce risque particulièrement présent dans les données fiscales des petites communes. Compte tenu de ces réserves chaque analyse communale des données 82-96 des sources et des recensements 82 et 90 permet comme on l'a vu de formuler des conclusions quant aux évolutions démographiques post censitaires selon les contextes communaux. Ces analyses dont on perçoit à travers les difficultés évoquées le caractère symptomatique local, paraissent peu automatisables dans l'absolu. Pourtant elles comprennent de nombreux points invariables d'une commune à l'autre. Il est donc possible, non pas de restituer toutes les spécificités communales, mais de mettre en valeur par des procédures systématiques les séries issues des sources administratives et d'approximer le niveau

d'incertitude entourant les évaluations post censitaires ainsi produites. On peut ainsi concevoir des techniques d'actualisation de la population des communes à une date post censitaire récente afin de produire simultanément des résultats assortis d'indicateurs de fiabilité pour un grand nombre de communes. Il existe dans cette optique plusieurs orientations possibles. Elles ont toutes en commun le lissage préalable des données brutes issues des sources administratives continues de façon à réduire les irrégularités des séries utilisées.

Sur le plan de l'ordonnancement de l'utilisation des données issues des sources administratives deux grands principes s'opposent.

Le premier est le plus spontané, il correspond aux exemples précédemment présentés et fait primer la spécialisation des séries d'effectifs issus des sources pour l'évaluation de chaque grandeur démographique étudiée. Autrement dit, on cherche à mesurer les évolutions post censitaires du parc de logements par les permis de construire, celles du nombre de ménages par la tendance des résidences principales fiscales, celles de la population d'adultes par la tendance du nombre d'électeurs et celles de la population jeune par la tendance du nombre de personnes à charge¹³. Dans cette démarche les seuls traitements envisagés se bornent à tenter de corriger l'écart éventuellement constaté en période intercensitaire entre les évolutions de chaque grandeur étudiée et la tendance de la source supposée correspondante. Ainsi, à Bordeaux on essaiera de corriger le défaut local d'inscription électorale des étudiants de façon à exploiter la série post censitaire du nombre d'électeurs pour appréhender la population adulte. Une telle approche a pour avantages de ne mettre en valeur que des informations dont la définition recoupe a priori celle des grandeurs étudiées. Elle fournit de plus des résultats plus diversifiés que la seule population totale. Elle a pour inconvénient de ne pas permettre de confronter entre eux les résultats issus du traitement des différentes sources puisque chaque résultat concerne une grandeur différente.

A l'inverse le second principe fait primer la confrontation finale des résultats issus du traitement des différentes sources. Ainsi, chaque effectif issu d'une source administrative continue sert de base à la production des évaluations post censitaires de quelques grandeurs démographiques¹⁴. On peut de la sorte obtenir quatre série d'évaluations post censitaires de la population totale des communes. La justification de cette approche tient en deux points. D'une part comme on l'a vu dans les exemples l'adéquation entre tendances des sources et évolutions intercensitaires des grandeurs supposées comparables est parfois très imparfaite, ce qui relativise les vertus de la spécialisation des sources. D'autre part, la confrontation est la seule possibilité de juger de la fiabilité relative des résultats après le travail d'évaluation post censitaire. En effet, en plus des corrections entre tendances comparables déjà évoquées, ce travail sollicite un certain nombre d'hypothèses concernant l'évolution post censitaire des paramètres liant les séries d'effectifs administratifs utilisées et les grandeurs démographiques étudiées. On extrapole, ainsi les variables liant les grandeurs comparables aux effectifs de sources avec la population totale selon un jeu d'hypothèses qui se combine donc avec les corrections des biais liés aux sources. Parmi ces variables on a principalement le taux de vacance des logements, la taille moyenne des ménages, la part des adultes et la part des jeunes. On formule alors trois familles d'hypothèses.

¹³ On peut aussi envisager une estimation de l'évolution de la taille moyenne des ménages par la tendance du nombre moyen de personnes à charge par foyer fiscal. Utilisée indirectement par une reconstitution des stocks (Modèle de Reconstitution Analytique, Christophe Bergouignan, Michel Libert, 1994) et directement par une analyse de tendance (Méthode de Reconstitution et de Régression des Facteurs d'accroissement, Christophe Bergouignan, 1997). Cette démarche a priori intéressante repose cependant sur des relations complexes difficiles à valider empiriquement.

¹⁴ Méthode de Régression des Facteurs d'accroissement, Méthode de Correction Spécifique des Facteurs d'accroissement, Méthode de Correction Spécifique des Stocks Christophe Bergouignan, 1996.

La première suppose la stabilité des paramètres après le dernier recensement. Cela signifie entre autres que taux de vacance des logements, taille moyenne des ménages, part des adultes et part des jeunes sont estimés constants et égaux à leur valeur à la date du dernier recensement. Les tendances de chaque source sont ainsi directement transformées en rythme d'évolution de la population totale. Cela correspond à la méthode MC2S'.

La seconde suppose une variation post censitaire des paramètres poursuivant exponentiellement les variations de la dernière période intercensitaire. Cela revient à considérer constants les rapports entre facteurs d'accroissement¹⁵ des effectifs administratifs et de la population totale. Cela correspond à la méthode MCSF. Une telle extrapolation exponentielle de l'agrégation des biais tendanciels des sources et des variables liant les grandeurs comparables aux effectifs de sources avec la population totale, a pour avantage la prise en compte du passé. Elle a néanmoins pour inconvénient le risque de répercussion post censitaire d'écarts de tendances conjoncturels spécifiques à la dernière période intercensitaire.

La troisième répond à ce souci de réduction de la répercussion de distorsions conjoncturelles transitoires dans les paramètres de traitement des séries issues des sources administratives continues. Elle procède par estimation de paramètres correction pour un ensemble comprenant plusieurs communes. Son principe est de réaliser, pour un nombre suffisant de communes, une régression linéaire entre les facteurs d'accroissement de la dernière période intercensitaire des effectifs issus des sources et ceux de la population totale. Les coefficients obtenus sont alors utilisés pour traiter les facteurs d'accroissement post censitaires des effectifs administratifs de façon à estimer la tendance post censitaire de variation de la population totale. Cela correspond à la méthode MRF. Outre la réduction de l'impact des distorsions communales conjoncturelles dans l'estimation des paramètres de traitement des séries administratives, cette démarche a l'avantage de mesurer le degré d'association entre les évolutions communales passées des effectifs administratifs et de la population totale. Il est possible dans cet esprit de constituer des sous ensembles communaux caractérisés par des formes particulières de relation entre tendances intercensitaires passées. Son principal inconvénient est le revers de ses qualités, c'est à dire la non prise en compte des situations durablement atypiques.

1.4 Les dispersions résiduelles comme critères de qualité des évaluations produites

D'un point de vue pratique il existe trois critères d'examen de la fiabilité des résultats d'une démarche d'évaluation post censitaire de grandeurs démographiques communales à partir des séries d'effectifs issus de sources administratives continues.

Le premier repose sur le fait que pour utiliser convenablement les séries issues des sources, celles ci doivent être porteuses d'une tendance qui émerge nettement des fluctuations fréquemment observées. En effet, dans le cas contraire les évolutions post censitaires déduites sont peu significatives à moyen terme et reflètent davantage l'instabilité temporelle des défauts de collecte associés aux sources que des variations durables. Dans cette optique le degré de dispersion des résidus après extraction de la tendance des taux d'accroissement annuels constitue un bon indicateur a priori du niveau d'irrégularité des séries administratives. On parle ainsi de variabilité accidentelle autour de la tendance¹⁶ (VAAT).

¹⁵ Soit $FAM X(t_0 ; t_n)$ facteur d'accroissement annuel moyen de la grandeur X sur la période $t_0 ; t_n$. On a $FAM X(t_0 ; t_n) = \sqrt[n]{\frac{X(t_n)}{X(t_0)}}$. On préfère les facteurs d'accroissement aux taux d'accroissement en raison de leur signe

toujours positif rendant les paramètres plus robustes.

¹⁶ La tendance des taux d'accroissement annuels est estimée par moyenne mobile, c'est un instrument fruste, mais il a le mérite d'être également adapté à chaque profil communal puisqu'il s'affranchit des intuitions subjectives de spécification des fonctions d'ajustement. Il est par conséquent une nouvelle garantie de comparabilité entre communes.

Le second qui n'est pas une mesure de dispersion résiduelle à proprement parler est assez évident. Il est nommé écart entre tendances (EET). C'est la valeur absolue de la différence entre le taux d'accroissement annuel moyen de la dernière période intercensitaire des effectifs issus de chaque source et des grandeurs comparables données par les recensements. Cela revient à vérifier la capacité des sources à correctement mesurer les variations de la dernière période intercensitaire des grandeurs démographiques a priori comparables. C'est à dire contrôler le niveau d'adéquation entre tendances passées des sources et des recensements. On peut par exemple chercher à mesurer le degré de correspondance entre les variations du nombre d'électeurs durant la dernière période intercensitaire et celles de la population adulte donnée par les recensements.

Le troisième, mesure le degré de dispersion relative des résultats selon les séries issues des sources de l'évaluation post censitaire d'une même grandeur. On étudiera ici la population totale puisqu'elle est l'exemple le plus évocateur. Cela suppose bien entendu que plusieurs effectifs administratifs évaluent sa tendance. Ce qui restreint l'analyse aux résultats des méthodes faisant primer la confrontation sur la spécialisation, c'est dire les méthodes MC2S', MCSF et MRF. Comme on l'a dit, ces méthodes incluent des hypothèses d'évolution post censitaire de certaines variables démographiques communales comme le taux de vacance, la taille moyenne des ménages, la part des jeunes ou des adultes. Si toutes ces hypothèses correspondent à la réalité l'évaluation de la population totale obtenue à partir du traitement de chaque série issue d'une source va aboutir à la véritable valeur de la population totale. Autrement dit, tous les résultats vont coïncider en une même valeur. C'est pourquoi, bien que la convergence des évaluations post censitaires produites à partir des séries de chaque source n'implique, théoriquement, pas forcément la validité des hypothèses, elle en constitue un indice relativement probant. Aussi, on mesure l'inadaptation temporelle des hypothèses de chaque méthode par la dispersion relative des évaluations issues du traitement des différentes sources¹⁷. On parle ici de Dispersion Relative en Ordre de Grandeur (DROG)¹⁸. Plus cette dispersion est importante moins les hypothèses posées sont adaptées aux réalités démographiques post censitaires. Cet indicateur, qui porte sur les résultats obtenus, peut sembler le plus pertinent. En fait, sa principale limite tient à sa portée uniquement interne, puisqu'il est construit, non pas à partir d'une confrontation avec une valeur optimale, mais à partir d'une confrontation entre valeurs issues du système étudié.

2 L'analyse empirique des écarts et dispersions résiduelles

L'analyse des valeurs obtenues pour chacun de ces indicateurs dans les 542 communes girondines comprend des explications méthodologiques d'une part (type de source en cause et méthode de traitement choisie) et des explications communales d'autre part. Ces dernières visent à caractériser les communes pour lesquelles les indicateurs décèlent une bonne ou une mauvaise fiabilité relative des éléments du processus d'évaluation post censitaire.

Au delà du constat de la présence d'effets d'interaction très complexes, cette analyse permet de conclure à l'existence d'un important effet source sur la dispersion autour de la tendance des séries de données (VAAT). Les effectifs de personnes à charge des foyers fiscaux

Cet indicateur de dispersion autour de la tendance des taux d'accroissements annuels est calculé pour chaque commune de la Gironde (542) et les quatre sources administratives continues étudiées. Il s'agit du parc total de logements actualisé par les permis de construire collectés par la Direction de l'Équipement (PLDRE), du nombre de résidences principales de la taxe locale d'habitation (RPTLH), du nombre de personnes à charge qui y sont associées (PACTLH) et du nombre d'inscrits sur les listes électorales (ILE).

¹⁷ Le coefficient de variation en est un bon indicateur.

¹⁸ Les DROG ici considérés correspondent à chaque méthode ici prise en compte (MRF, MCSF, MC2S'), chaque commune de la Gironde (542) pour 96 dernière date post censitaire pour laquelle des évaluations ont été produites.

montrent une tendance bien moins régulière que les inscrits sur les listes électorales. Cela peut sembler surprenant, vu la sensibilité de ces dernières à la conjoncture politique et aux vérifications périodiques des commissions. L'irrégularité des séries fiscales associe sans doute, problèmes de gestion ou de saisie informatique des données¹⁹ et variabilité des mesures. Les évolutions du parc total de logements montrent la tendance la plus régulière, sans doute en raison de programmes pluriannuels de construction.

Les écarts entre tendances intercensitaires passées (EET) les plus importants concernent aussi les effectifs de personnes à charge des résidences principales fiscales. Ceux ci présentent en moyenne entre 82 et 90 des tendances supérieures à celles du nombre de jeunes. Cela résulte sans doute du maintien prolongé de ces derniers dans le ménage parental, causé par l'allongement de la durée des études et les difficultés d'insertion économique. Les électeurs présentent une meilleure adéquation de leurs tendances intercensitaires avec celles du nombre d'adultes. Les écarts de tendances existant vont dans le sens de tendances 82-90 inférieures pour les électeurs. Cela peut s'expliquer par une probable baisse de la proportion d'adultes inscrits au fil du temps. Tendances du parc de logement actualisé par les permis de construire et des résidences principales fiscales montrent en moyenne de faibles écarts avec les tendances du parc de logements et du nombre de ménages mesurées par les recensements. Ces écarts lorsqu'ils existent ne présentent pas de sens réellement dominant.

Un important effet méthode, sur la dispersion, autour de la valeur centrale, des évaluations issues du traitement de chaque source (DROG), peut être mis en évidence. Les régressions entre facteurs d'accroissements de la dernière période intercensitaire d'un ensemble de zones (méthode MRF) montrent ainsi en moyenne des dispersions plus faibles c'est-à-dire la meilleure robustesse. En revanche la prise en compte du seul passé intercensitaire local (méthode MCSF) aboutit en moyenne aux dispersions les plus importantes ce qui démontre la fragilité globale d'une telle procédure de correction.

La relative importance de ces effets source et méthode est cependant loin d'écraser l'effet commune sur les indicateurs de fiabilité des ingrédients, tendances et résultats de la démarche d'évaluation post censitaire communale. Les fortes corrélations entre les indicateurs d'une même nature (les 4 VAAT entre eux, les 4 EET entre eux et les 3 DROG entre eux) et entre ceux de nature différente pris deux à deux, indiquent, une forte cohérence entre les différents aspects de la fiabilité. Cette cohérence entre indicateurs signifie probablement la présence de facteurs communaux qui déterminent en partie la valeur des mesures de qualité. En effet de telles corrélations signifient que si un indicateur signale une faible fiabilité dans une commune, les autres indicateurs la signaleront aussi. Les caractéristiques communales supposées conditionner ainsi les indicateurs de fiabilité des ingrédients, tendances et résultats de la démarche, sont naturellement la taille de la population au dernier recensement, mais aussi le degré de renouvellement de la population. Celui ci sera mesuré par la proportion d'individus ayant déclaré résider à l'avant dernier recensement dans une autre commune que celle où ils résidaient au dernier recensement. Cette mesure du degré de renouvellement reste bien entendu imparfaite, puisqu'elle fait l'impasse sur les flux réels de migrations intercensitaires. Elle ne tient en particulier aucun compte de l'hétérogénéité du parc de logements vis à vis du renouvellement. Il existe en effet des portions de parc locatif qui changent d'occupants chaque année, or la mesure proposée ne fait aucune différence entre un logement ayant abrité 8 ménages différents en 8 ans et 2 ménages différents en 8 ans. Reste qu'en l'état actuel des informations statistiques disponibles, cette proportion appelée par abus de langage proportion d'immigrants est la seule à couvrir l'ensemble des ménages et personnes.

¹⁹ Des valeurs nulles aberrantes apparaissent en 94 pour des communes de moins de 100 habitants. Voir aussi à ce sujet l'influence de la connexion IRTD au niveau des départements français (Jean Claude Labat, 1996).

2.1 L'influence de l'effectif des population communales

Dans la logique de la théorie de l'échantillonnage, on peut supposer que le premier élément du contexte communal à influencer la dispersion d'une mesure démographique est la taille de la population. Cette hypothèse est fréquente, bien qu'une population communale ne soit pas à proprement parler un échantillon tiré aléatoirement dans un ensemble précisément défini. Parfois, on étudie l'influence de la taille des populations sur la dispersion d'une mesure en comparant la valeur que prend un indicateur de dispersion dans de grands ensembles géographiques avec ses valeurs dans les sous-ensembles qui le composent. Ainsi, les indicateurs de dispersion des résultats (DROG 96) des trois méthodes, pour l'ensemble de la Communauté Urbaine de Bordeaux (CUB), donnent une dispersion relative bien inférieure à la moyenne de ceux de chaque commune de la CUB, quelle que soit la méthode²⁰. Cependant, pour la commune de Saint-Médard-en-Jalles on fait le constat opposé. La dispersion relative moyenne des évaluations post censitaires de la population des quartiers est plus faible que celle des évaluations portant sur la commune entière²¹.

En fait, ces résultats font davantage la preuve de l'existence d'interactions entre le type de découpage et d'autres facteurs, dont la taille de la population, que celle d'une influence directe de la taille sur la dispersion des mesures. En effet, si la nature du découpage n'est évidemment pas sans relation avec la taille, elle l'est aussi avec le niveau de renouvellement et la variabilité des structures de sous populations. Aussi, par souci de comparabilité, et de façon à démêler les influences respectives des divers éléments constituant le contexte communal, il est préférable de ne mettre en regard que des indicateurs correspondant à un même niveau de découpage géographique. Autrement dit, il convient de ne prendre que des communes pour individus statistiques dans l'analyse du rôle de la taille et du niveau de renouvellement sur les indicateurs de fiabilité étudiés. En effet, dans ce travail les communes constituent l'ensemble le plus important d'unités géographiques (542 éléments), et celui qui offre en son sein la plus grande diversité de valeurs pour les indicateurs de qualité.

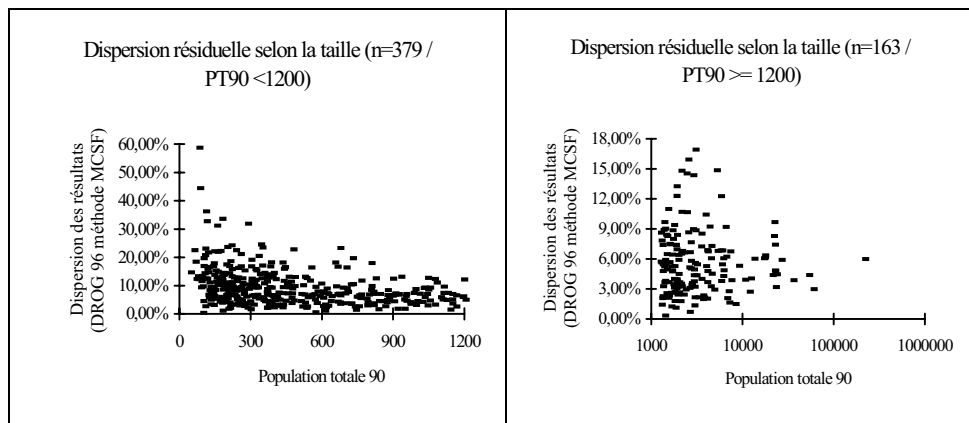
Les associations bivariées²² observées pour l'ensemble des 542 communes de la Gironde entre la taille de l'effectif communal et chaque indicateur de qualité sont peu élevées. Le constat d'un faible effet taille vaut tant pour la signification de la tendance que pour les écarts entre tendances passées ou la dispersion des résultats. Malgré tout, la taille est presque toujours corrélée négativement aux différents indicateurs de qualité. Cela correspond au schéma de « bon sens », d'écarts ou de dispersions résiduelles moyennes d'ampleur inversement proportionnelle à la taille de la population étudiée. La faiblesse des associations laisse néanmoins supposer l'existence d'autres relations explicatives, puisque, pour les communes les plus peuplées, il peut exister une association positive entre taille et niveau de renouvellement. Pour éviter l'influence de ces phénomènes, on distingue parmi les 542 communes de la Gironde, les 379 communes de moins de 1200 habitants²³ et les 163 restantes.

²⁰ 0,88% contre 2,41%, 1,86% contre 3,35% et 2,41% contre 3,87%.

²¹ 2,18% contre 2,93%.

²² On trouvera en annexe la matrice des corrélations entre les indicateurs correspondant à chaque dimension de la qualité et la taille de la population des communes, mesurée par l'effectif au dernier recensement général 90 INSEE. Les évaluations post censitaires synthétiques de population communale produites selon les différentes méthodes ne pouvant être utilisées dans ce but puisqu'elles portent une incertitude qui est l'objet même de cette étude.

²³ On trouvera en annexe la matrice des corrélations entre les indicateurs de dispersion et la taille des communes pour ce groupe de 379 communes et pour les 163 restantes.



Bien qu'une part des écarts ou dispersions résiduelles reste à expliquer, les associations bivariées, avec la taille sont très nettement plus intenses²⁴ pour le groupe des 379 communes les moins peuplées. Leur sens corrobore nettement les présupposés, malgré l'existence de très faibles écarts ou dispersions résiduelles dans certaines communes ayant une très petite population. Aussi, peut-on parler de diminution de l'ampleur des écarts ou dispersions résiduelles potentielles avec la taille. A contrario, les faibles corrélations observées pour les 163 communes les plus peuplées laissent supposer qu'au-delà d'un certain nombre d'habitants (entre 800 et 1400), le rôle de la taille est supplanté par d'autres facteurs. Il peut s'agir d'éléments méthodologiques déjà identifiés (nature des sources utilisées et méthode de traitement des séries), mais surtout de facteurs communaux alternatifs. Ce résultat confirme les résultats initiaux concernant les 27 communes relativement peuplées de la CUB, pour lesquelles la valeur des indicateurs de qualité n'était que faiblement expliquée par la taille, surtout après la prise en compte de variables complémentaires.

2.2 L'influence du degré de renouvellement de la population des communes

Comme on l'a vu, l'influence de l'effectif de la population communale sur les indicateurs de qualité pour les 163 communes de 1200 habitants ou plus est légère. Quant aux 379 petites communes une partie des phénomènes d'écarts de tendance ou de dispersions résiduelles reste parfois à expliquer malgré un important rôle de la taille. Aussi, il convient de tenter d'enrichir avec d'autres facteurs communaux les éléments d'explication de la précision des mesures. D'après les hypothèses déjà émises et les associations bivariées observées pour les 379 petites communes et les 163 plus peuplées on ne retient qu'une variable communale alternative révélatrice d'un important renouvellement de la population. Comme on l'a dit, il s'agit de la proportion de personnes recensées en 90 qui résidaient dans une autre commune en 82²⁵.

On procède ainsi à une analyse de l'impact du degré de renouvellement, sur les indicateurs de qualité des éléments du processus d'évaluation post censitaire à partir de l'observation des

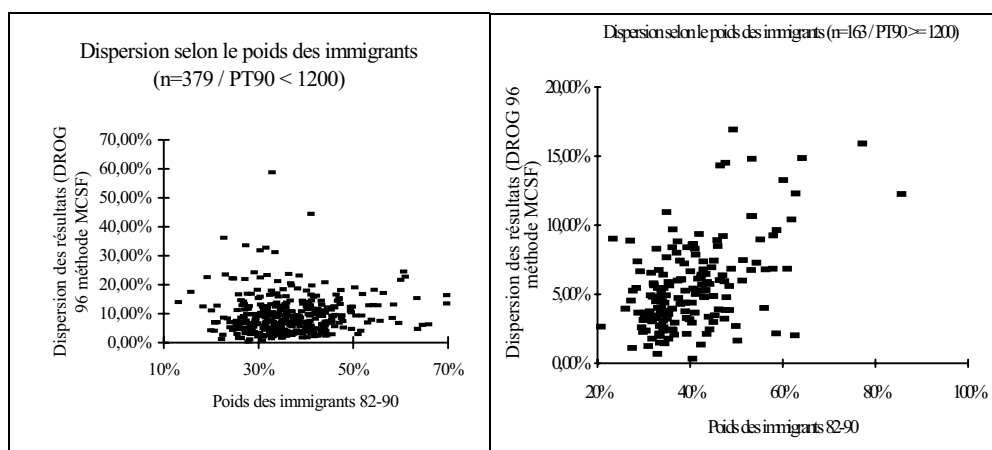
²⁴ L'exclusion ou l'inclusion des quelques valeurs nulles aberrantes des séries de données des communes de moins 100 habitants, ne change rien à l'intensité des associations. Le fait que les quelques valeurs nulles aberrantes issues des problèmes de gestion ou de saisie des fichiers ne touche que des petits effectifs communaux pourrait être un autre versant de l'effet taille.

²⁵ Informations issues du recensement général de population INSEE de 90.

coefficients de corrélation²⁶. En effet, la faible colinéarité entre taille et degré de renouvellement²⁷, ne nécessite pas l'analyse des coefficients de corrélation partiels.

Pour le groupe des 163 communes les plus peuplées, on constate que le degré de renouvellement vérifie des corrélations souvent plus intenses que la taille et dans un sens presque toujours conforme aux hypothèses. Plus de renouvellement se manifeste en moyenne par plus d'écarts ou de dispersions résiduelles. Cette constatation s'explique sans doute par les problèmes de gestion de données rencontrés par les organismes chargés de la collecte, lorsque le renouvellement des administrés est important. Dans un tel cas, les risques de doubles comptes ou d'omissions des administrations responsables des fichiers sont en effet largement accrus.

Pour le groupe des 379 communes les moins peuplées, l'introduction d'une variable de renouvellement n'est pas en mesure de concurrencer le rôle de la taille sur les écarts et les dispersions résiduelles. En effet, contrairement au groupe des 163 communes les plus peuplées, ces fortes corrélations négatives entre taille et écarts ou dispersions résiduelles sont dans la majorité des cas d'une intensité très supérieure à celles faisant intervenir le degré de renouvellement. Ces dernières sans être toujours négligeables, ne permettent guère de conclure à des effets très nets.



L'effet commune, sur l'ampleur des dispersions correspondant aux dimensions de la qualité des éléments du processus d'évaluation post censitaire de la population des communes, n'est donc pas aussi manifeste que l'on pouvait le croire à l'issue des travaux exploratoires. Il reste cependant significatif et son sens est presque toujours conforme aux hypothèses : une taille importante et un faible degré de renouvellement de la population réduisent en moyenne les écarts ou les dispersions résiduelles.

Dans les communes les moins peuplées, l'effet commune est plus fort. Il est presque uniquement le fait de la taille de la population communale, alors que le degré de renouvellement de la population est quasiment sans effet. Ainsi, plus les effectifs de population sont faibles plus on rencontre de communes pour lesquelles les dispersions sont importantes.

Dans les communes plus peuplées, le rôle propre de la taille est marginal, sinon négligeable, au regard du degré de renouvellement de la population. L'approximation du poids des immigrants intercensitaires dans la population communale exprime en partie le degré de

²⁶ On trouvera en annexe la matrice des corrélations entre les indicateurs de qualité et le degré de renouvellement pour le groupe de 379 communes et pour les 163 restantes.

²⁷ $R=0,113$ pour les 379 communes de moins de 1200 habitants, $R=0,025$ pour les 163 communes de 1200 habitants ou plus

disjonction entre les informations statistiques utilisées et les réalités de l'évolution démographique. Autrement dit, plus il y aurait eu en période intercensitaire²⁸ d'arrivées de population, plus il serait difficile de mesurer l'impact et l'éventuelle persistance de telles mutations en période post censitaire.

Ces graphiques montrent aussi que certaines communes très peu peuplées présentent de très faibles dispersions. Ainsi, en toute rigueur si un petit effectif ou un fort degré de renouvellement de la population d'une commune n'impliquent pas nécessairement des mesures très dispersées, on considérera plutôt qu'ils induisent des écarts ou dispersions résiduelles potentiellement plus importants. Le tableau d'exemples communaux suivant montre en effet la diversité des sensibilités à l'effet taille ou degré de renouvellement des écarts ou dispersions résiduelles.

La méthode extrapolant les relations communales passées entre tendances, pour passer des séries de chaque source aux évolutions de la population totale (MCSF), est de beaucoup la plus sensible. C'est pourquoi la dispersion de ses résultats illustre d'autant plus nettement l'influence des facteurs pris en compte. On constate en effet pleinement la réalité des impacts de la taille et du degré de renouvellement, mais aussi leur caractère non systématique. Ainsi Cleyrac qui présente une dispersion de résultats très importante illustre parfaitement les risques associés à un très faible effectif de population. Cela n'empêche aucunement St Genès de Lombaud d'associer petite taille et très faible dispersion des résultats. Dans les communes plus peuplées le rôle du degré de renouvellement est assez net. En effet, Floirac commune à fort renouvellement montre compte tenu de sa taille une forte dispersion des résultats, alors que Léognan caractérisé par un faible renouvellement montre une dispersion très légère. Reste que Gujan Mestras qui compte presque 50% d'immigrants intercensitaires, montre une assez faible dispersion. Enfin Labarde illustre bien l'effet cumulatif d'un effectif de population modeste et d'un degré de renouvellement exceptionnel, la dispersion des résultats y est aussi forte que dans les plus petites populations chargées des plus grosses incertitudes.

Communes	Population totale 90	Part 90 des immigrants	Dispersion des résultats pour la méthode d'extrapolation des tendances passées locales (DROG _{MCSF} 96)
Cleyrac	143	46,44%	15,39%
Floirac	16857	46,10%	6,36%
Gujan Mestras	11430	49,02%	2,72%
Labarde	695	69,10%	16,39%
Léognan	7992	32,57%	1,48%
St Genès de Lombaud	205	29,29%	2,09%

Ces résultats laissent légitimement imaginer l'existence d'une théorie des risques d'écarts et de dispersions résiduelles des tendances démographiques et des estimations des grandeurs en population ouverte. Elle verrait s'appliquer des principes probabilistes, sur des paramètres fonctions de l'interaction entre degré de renouvellement de la population, sources mises à contribution et méthodes de traitement. L'impact de tels principes probabilistes sur les résultats serait alors maximal dans les plus petites populations où il écraserait celui des autres critères. Il ne disparaîtrait, derrière d'autres variables que dans des populations un peu plus conséquentes notamment lors de travaux fondés sur la correction de données imparfaites. Parmi ces critères communaux influant sur les dispersions résiduelles des ensembles les plus peuplés, le degré de

²⁸ Il était délicat d'utiliser les évaluations synthétiques post censitaires de population comme composante des variables explicatives vu qu'elles sont la synthèse des composantes de variables à expliquer.

renouvellement occuperait une position privilégiée sans pour autant avoir l'importance que l'on était en droit d'attendre.

Cette hypothèse de cohérence théorique est confortée par les associations observées entre les indicateurs de qualité de la démarche d'évaluation post censitaire et les dispersions résiduelles des séries temporelles des taux bruts de mortalité et de natalité autour de leur tendance. Ces séries temporelles qui sont très heurtées à l'échelon communal participent en effet à cette démarche d'appréhension post censitaire des évolutions démographiques communales. Elles peuvent constituer dans le contexte post censitaire un outil approximatif d'analyse de l'évolution de la part des personnes les plus âgées, ainsi que de celle des adultes d'âge fécond.

Dans l'attente d'une formalisation théorique aboutie, le travail en petite ou moyenne population reste tout à fait envisageable. Les situations d'importante incertitude comme les évaluations post censitaires de population communale, les estimations de tendance de flux relatifs ou les projections, apportent alors des difficultés supplémentaires, qui ne sont pas pour autant insurmontables. En effet, si les écarts et dispersions résiduelles sont en moyenne plus forts pour les petits effectifs communaux, cette situation n'est pas systématique. Ainsi pour certaines petites populations communales le risque est aussi faible que pour les communes les plus peuplées du département. C'est pourquoi plutôt que de fixer une stricte limite d'effectif, mieux vaut concevoir des indicateurs de fiabilité adaptés à chaque contexte d'étude, de façon à renseigner sur le niveau d'incertitude entourant les résultats, et les ingrédients. Ces indicateurs doivent permettre de répondre non seulement à la question « de combien risque-t-on de se tromper ? », mais aussi à la question « pourquoi risque-t-on de se tromper de tant ? ».

Conclusions pratiques et illustrations

Les constats précédents montrent l'influence potentielle de la taille et du degré de renouvellement des populations communales sur la fiabilité des mesures démographiques post censitaires. L'existence de communes échappant à la majoration des incertitudes associée à un faible effectif de population et/ou à un fort renouvellement montre toutefois le caractère non systématique de tels facteurs de qualité. Aussi, face à la demande d'actualisation des tendances démographiques émanant de décideurs locaux, aucune considération *a priori* sur la taille de la population ne saurait motiver un refus de chercher et d'informer. La diversité des situations et des besoins implique en fait une réflexion cas par cas qui n'est pas pour autant un affranchissement des conditions de rigueur. Au contraire, c'est l'exigence, quelle que soit la taille ou le niveau de renouvellement de la population d'étude, de mesurer par des indicateurs adaptés le degré de signification et de cohérence des ingrédients et résultats. Il n'est cependant guère concevable d'associer à chaque communication de résultats une kyrielle d'indicateurs renseignant sur le degré de respect de chaque contrainte des options de traitement. A l'inverse, on ne peut décemment livrer sans commentaire annexe le résultat considéré comme le plus probable. En fait, plutôt que d'engendrer l'incompréhension ou l'erreur, il convient de construire des bornes minimales et maximales de fluctuation de la valeur évaluée. Ce calcul se fonde sur les indicateurs de dispersion des résultats (DROG). Il tient aussi compte du risque lié aux erreurs existant dans les données de recensement. Même très faibles des erreurs de sens contraire à deux recensements successifs biaisent la tendance intercensitaire qui sert de référence dans les méthodes d'évaluation. On présente ici, les valeurs maximales et minimales de la population totale en 96 pour certaines communes.

Confirmant l'influence moyenne de la taille sur les indicateurs de qualité, les intervalles de fluctuation proportionnellement les plus larges concernent principalement des petites populations communales. C'est le cas notamment de St Antoine, mais aussi dans une moindre mesure de Croignon et de Landerrouat. Réciproquement les plus petits intervalles relatifs touchent majoritairement des communes dont la population est assez importante comme

Arcachon, St Emilion mais aussi Talence. Reste qu'une commune relativement peuplée comme Lacanau génère un intervalle d'ampleur supérieure au quart de son effectif estimé (valeur centrale de l'intervalle). A l'inverse des communes aux effectifs modestes comme Camps sur l'Isle ou La Sauve génèrent des intervalles d'ampleur relativement faible.

Commune	Taux 82-90 de croissance annuel moyen	Population totale 90	Borne inférieure 96 d'évaluation	Borne supérieure 96 d'évaluation	Ampleur relative de l'intervalle
Arcachon	-1,47%	11794	11126	11807	5,77%
Camps sur l'Isle	0,00%	377	383	401	4,59%
Croignon	2,39%	349	335	373	10,89%
Lacanau	2,60%	2411	2344	2954	25,30%
Landerrouat	0,67%	193	175	193	10,06%
St Antoine	-1,83%	296	260	305	15,25%
St Emilion	-1,32%	2799	2578	2695	4,18%
St Julien Beychevelle	-1,73%	874	759	865	13,05%
La Sauve	1,14%	1100	1145	1220	6,34%
Talence	-0,03%	34523	34450	36781	6,54%

Au delà de la question de l'ampleur relative des intervalles, c'est à dire du degré de précision des évaluations post censitaires communales, se pose le problème de l'apport des intervalles proposés en termes de spécificité des contextes communaux. Il n'est pas indifférent d'étudier une commune dont la forte croissance démographique antérieure et les récentes constructions de lotissements rendent l'augmentation de la population évidente pour tous les décideurs locaux ou de chercher à distinguer les évolutions beaucoup plus incertaines d'une commune réhabilitant son parc de logements. En effet, dans ce dernier cas les changements de nature des ménages occupant les immeubles avant et après travaux, rend les transformations quantitatives de la population insaisissables pour les acteurs de terrain. De même les transformations du tissu rural sous l'impact conjugué de l'implantation de rurbains et du vieillissement et de la mortalité des anciens habitants, sont assez difficiles à appréhender sans une démarche scientifique. Aussi les exigences de précision varient elles d'un contexte communal à l'autre. On peut donc regrouper les résultats en fonction de leur apport informationnel selon le contexte communal.

Les résultats communaux les plus significatifs en termes d'intervalles de fluctuation des évaluations post censitaires concernent Camps sur l'Isle et St Emilion. Dans ces deux cas, le niveau de précision des réponses correspond tout à fait à ce que peut attendre un décideur local. Pour Camps sur l'Isle, commune rurale ayant connu une forte croissance dans les années 60, qui s'est fortement ralentie depuis pour stagner entre 82 et 90, on est assuré d'une légère reprise de croissance de 90 à 96 (entre 0,25 et 1% annuels). St Emilion commune viticole ayant connu un fort dépeuplement depuis 62, voit très probablement celui ci se poursuivre de 90 à 96 mais à un moindre rythme (entre -0,6% et -1,3% annuels) que durant la dernière période intercensitaire.

A un niveau de signification encore conséquent mais déjà moindre on trouve La Sauve, Landerrouat et St Julien Beychevelle. Ici les résultats proposés ont une valeur plus informationnelle qu'opérationnelle et peuvent servir de base au redressement de certaines démarches prospectives. Pour Landerrouat, commune rurale connaissant de faibles variations de population de 62 à 90, il est certain que la population n'a pas augmenté entre 90 et 96, on ignore cependant si il s'agit d'une stagnation ou d'une baisse significative (-1,6% annuels). La Sauve,

dont la rurbanisation s'est engagée relativement tôt pour donner lieu à une croissance régulière et assez forte depuis 75, connaît certainement une poursuite de celle ci à un rythme voisin (entre 0,7 et 1,7% annuels) de celui connu de 82 à 90. St Julien Beychevelle commune viticole qui perd des habitants depuis 68, connaît une continuation de son dépeuplement entre 90 et 96, mais avec un rythme qu'il est difficile de cerner avec certitude (entre -0,2 et -2,3% annuels).

Arcachon et Talence ont des intervalles dont l'apport en matière d'évolutions démographiques post censitaires est relativement modeste. Il peut toutefois servir de repère général aux acteurs de terrain, notamment en termes d'évaluation de projections antérieures. Talence banlieue universitaire qui après une croissance soutenue voit sa population stagner depuis 75, est caractérisée par la seule assurance de non diminution de sa population entre 90 et 96. Cependant on ne peut dire avec certitude si il s'agit d'une stagnation ou d'une croissance modérée (+1,1% par an). Arcachon, ville touristique et centre d'un pôle d'urbanisation, perd des habitants depuis 68, et n'en a probablement pas gagné de 90 à 96, mais on ne peut préciser si il s'agit d'une stabilisation de la population ou d'une diminution ralentie (-0,97% annuels).

Pour Croignon et St Antoine le niveau d'incertitude est tel qu'il devient quasiment impossible de proposer des conclusions susceptibles d'intéresser les décideurs locaux. Tout au plus peut on dire pour Croignon, commune ruraine en fort peuplement depuis 75, que la croissance soutenue 82-90 est assez loin d'être atteinte de 90 à 96. Il est cependant impossible de dire si on a affaire à une poursuite de la hausse à moindre rythme (1,1% annuels), à une stabilisation ou à une amorce de baisse (-0,7% par an). Quant à St Antoine, seul le ralentissement entre 90 et 96 d'une éventuelle poursuite du dépeuplement 82-90 constitue une assurance, ce qui n'empêche pas de pouvoir conjecturer une stabilisation ou une hausse modérée (0,5% par an).

Reste Lacanau commune touristique littorale dont l'intervalle est si étendu qu'il comprend la reproduction de 90 à 96 de la croissance soutenue 82-90 ainsi que l'amorce d'une diminution de la population. Une telle incertitude rend les résultats tout à fait inutilisables au plan pratique.

Ainsi, il est clair que les évaluations post censitaires de population fondées sur les tendances des effectifs issus des sources administratives, ne constituent pas une panacée en termes de démographie appliquée à l'aide à la décision communale. On ne peut par exemple pas les comparer loin s'en faut aux registres de population, qui constituent l'instrument privilégié des dénombrements continus. Pour autant, rejeter cette option d'information ou la limiter aux très grands effectifs de population serait un gaspillage évident vu le faible coût de revient de tels travaux. D'une part, pour quelques communes les résultats obtenus ont toutes les caractéristiques de l'excellence. D'autre part, pour la plupart des autres communes les résultats apportent des indications qui sans être d'une grande précision ont le mérite d'être objectives contrairement aux extrapolations plus ou moins explicites auxquelles se livre de toutes façons tout décideur non informé.

ANNEXE TECHNIQUE

VAAT_s : Variabilité Accidentelle Autour de la Tendance des effectifs de la source s pour chaque commune on a :

$$VAAT_s = \sqrt{\text{Var}(e_{s;t})} \quad \text{avec} \quad \begin{array}{ll} AA_{s;t} & = \text{taux de croissance de l'année } t \text{ des} \\ & \text{effectifs issus de la source continue } s \\ s & = \text{PLDRE; RPTLH; PACTLH, ILE} \\ t & = 82 \text{ à } 95 \end{array}$$

$$e_{s;t} = AA_{s;t} - ma_{s;t}$$

$$ma_{s;t} = \frac{1}{3} \sum_{k=-u \text{ à } u} AA_{s;t+k}$$

EET_s : Ecart Entre Tendances de la source s et de la grandeur comparable issu des derniers RGP pour chaque commune on a :

$$EET_s = |AAM_{s;82-90} - AAM_{G \text{ INSEE};82-90}| \quad \text{avec} \quad \begin{array}{ll} AAM_{s;82-90} & = \text{Taux d'accroissement} \\ & \text{annuel moyen 82-90 de} \\ & \text{l'effectif de la source } s \\ AAM_{G \text{ INSEE};82-90} & = \text{Taux d'accroissement} \\ & \text{annuel moyen 82-90 de} \\ & \text{la grandeur comparable} \\ & \text{issu des RGP 82 et 90} \end{array}$$

DROG_{p;t} : Dispersion Relative en Ordre de Grandeur autour de l'évaluation centrale de la méthode p l'année post censitaire t pour chaque commune on a :

$$DROG_{p;96} = \frac{\sqrt{\text{Var}(E_{p;s;PT;96})}}{E_{p;PT;96}}$$

$E_{p;s;PT;96}$ ²⁹ = Évaluation autonome de la population totale au 01/01/96 à partir de la source administrative *continue* s, par la méthode p

$E_{p;PT;96}$ = Évaluation centrale de la population totale au 01/01/96 par combinaison linéaire des évaluations autonomes $E_{p;s;PT;96}$

²⁹ Les techniques de calcul des $E_{p;s;PT;96}$ selon les méthodes MC2S (conforme à la MC2S' à une extrapolation près), MCSF et MRF sont explicitées dans la communication du X^e colloque national de Démographie de Bordeaux : Démographie post censitaire communale, Christophe Bergouignan, 1996.

TABLEAUX STATISTIQUES (542 COMMUNES DE LA GIRONDE)Description univariée des indicateurs de qualité et des variables communales³⁰ (n=542)

Variables	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'	DRO G MCSF	DRO G MRF	EET PLDRE	EET RPTLH
Minimum	0,001	0,005	0,008	0,004	0,004	0,003	0,005	0,000	0,000
Maximum	0,053	0,149	0,242	0,217	0,358	0,588	0,161	0,046	0,043
Moyenne	0,010	0,027	0,052	0,020	0,063	0,082	0,040	0,005	0,005
Écart type	0,006	0,017	0,033	0,017	0,039	0,060	0,021	0,005	0,005

Variables	EET PACTLH	EET PACTLH	EET ILE	PT90	PIMM
Minimum	0,000	0,000	0,000	40	0,125
Maximum	0,134	0,134	0,066	210467	0,848
Moyenne	0,016	0,016	0,008	2240	0,372
Écart type	0,016	0,016	0,007	10173	0,094

Matrice de corrélation des indicateurs de qualité entre eux (n=542)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
VAAT RPTLH	0,189				
VAAT PACTLH	0,099	0,648			
VAAT ILE	0,059	0,258	0,230		
DROG MC2S	0,094	0,393	0,406	0,198	
DROG MCSF	0,165	0,403	0,436	0,240	0,652
DROG MRF	0,214	0,439	0,446	0,276	0,731
EET PLDRE	0,629	0,165	0,148	0,020	0,079
EET RPTLH	0,188	0,427	0,336	0,104	0,193
EET PACTLH	0,084	0,165	0,328	0,122	0,169
EET ILE	0,050	0,187	0,162	0,175	0,065

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH
VAAT RPTLH					
VAAT PACTLH					
VAAT ILE					
DROG MC2S					
DROG MCSF					
DROG MRF	0,625				
EET PLDRE	0,093	0,165			
EET RPTLH	0,254	0,268	0,176		
EET PACTLH	0,287	0,173	0,000	0,238	

³⁰ PT : population totale; PIMM : poids des immigrants. Recensement général de population INSEE 90.

EET ILE	0,133	0,168	-0,025	0,130	0,133
---------	-------	-------	--------	-------	-------

Corrélation des indicateurs de qualité avec la taille des populations communales (n=542)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
PT 90	-0,063	-0,170	-0,176	-0,043	-0,088

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH	EET ILE
PT 90	-0,089	-0,092	-0,084	-0,062	-0,077	0,022

Corrélation des indicateurs de qualité avec la taille des populations communales (n=379 / PT 90 < 1200)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
PT 90	-0,152	-0,559	-0,557	-0,182	-0,322

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH	EET ILE
PT 90	-0,334	-0,396	-0,100	-0,256	-0,177	-0,091

Corrélation des indicateurs de qualité avec la taille des populations communales (n=163 / PT 90 ≥ 1200)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
PT 90	-0,085	-0,189	-0,206	0,057	-0,058

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH	EET ILE
PT 90	-0,016	-0,053	-0,077	0,012	-0,084	0,102

Corrélations des indicateurs de qualité avec le poids des immigrants (n=379 / PT 90 < 1200)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
PIMM	0,158	0,031	0,045	0,048	-0,042

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH	EET ILE
PIMM	0,047	0,03	-0,108	0,005	0,046	0,238

Corrélations des indicateurs de qualité avec le poids des immigrants (n=163 / PT 90 ≥ 1200)

	VAAT PLDRE	VAAT RPTLH	VAAT PACTLH	VAAT ILE	DROG MC2S'
PIMM	0,326	0,213	-0,134	0,290	0,082

	DROG MCSF	DROG MRF	EET PLDRE	EET RPTLH	EET PACTLH	EET ILE
PIMM	0,509	0,445	0,017	0,207	-0,023	0,176

Diversité et dynamiques dans les micro-populations

Jean Louis RALLU

INED

L'étude des petits nombres ne concerne pas seulement les petites populations mais peut bien sur concerner des sous populations peu nombreuses de grandes populations. Par ailleurs, travailler sur de grandes populations n'empêche pas d'être confronté à de petits nombres, tout démographe travaillant sur des enquêtes finit toujours par aboutir à des classes trop petites posant des problèmes pour l'analyse. Quelle méthodologie adopter devant les petits nombres? Les modèles statistiques de plus en plus sophistiqués permettent de considérer des variables toujours plus nombreuses et de dépasser ainsi les problèmes des petits nombres. Cependant il arrive toujours un moment où les modèles ne répondent plus à la diversité du réel et donnent des résultats « chaotiques ». Quelle stratégie d'analyse adopter alors? Nous allons nous placer dans l'optique de l'utilisation de statistiques simples et présenter notre démarche pour l'étude d'une sous-population issue de petites populations, les migrants de retour dans trois pays insulaires du Pacifique sud : Kiribati (72 335 habitants en 1990), Tuvalu (9043 habitants en 1991) et les Îles Cook (18 552 habitants en 1991). Nous considérerons deux approches, par les recensements et par une enquête et nous nous interrogerons particulièrement sur le choix du type d'enquête : enquête statistique représentative ou enquête « qualitative ».

Définition et évolution de la population étudiée.

Les migrants de retour ont été définis dans le cadre d'une approche à partir des recensements par la réponse à la question sur la résidence antérieure. Il s'agit des personnes nées dans le pays, recensées comme résidant dans le pays et ayant déclaré résider à l'étranger en réponse à la question sur la résidence antérieure. Cette approche a permis de dénombrer 1528 migrants de retour à Kiribati, 373 à Tuvalu et 858 aux Îles Cook. Nous sommes bien dans le domaine des petits nombres. Ce fut d'ailleurs une des surprises du chercheur d'être confronté à de si petits nombres dans des pays réputés avoir une migration importante. Certes la migration est importante, mais les retours sont peu fréquents, de plus l'approche par les recensements ne retient qu'une partie des migrants de retour, ceux résidant à l'étranger à la date retenue par la question sur la résidence antérieure, habituellement 5 ans avant le recensement. Ces migrants de retour méritaient cependant d'être étudiés, de même que les migrants lorsqu'on les étudie dans les recensements ; ils représentent des retours relativement récents, au cours des 5 dernières années, c'est-à-dire des personnes encore dans la phase de réinsertion.

L'étude démographique d'une population pose d'abord le problème de son renouvellement, c'est à dire comment on entre dans cette population et comment on en sort. Par définition, on est migrant de retour si on a émigré et revient dans son pays d'origine (la définition de l'origine restant à préciser, mais nous avons retenu ici le lieu de naissance). Cela suppose qu'on puisse calculer des taux de retour, risque pour les migrants de revenir dans leur pays d'origine. La définition des migrants (de retour) par les recensements facilite cette étape, car il suffit de connaître ou d'estimer l'effectif des migrants hors du pays à la date de référence (t) de la résidence antérieure, en utilisant les données des pays d'accueil. Le taux de retour est alors le rapport des migrants de retour recensés au stock de la population émigrée t années avant le recensement. - On peut aussi calculer la proportion de migrants de retour dans la population du pays à la date du recensement, mais il ne s'agit pas d'un taux. - Cependant, il intervient une ré-émigration entre le retour et le recensement et comme à l'habitude le recensement ne donne qu'une sorte de solde : personnes revenues non encore reparties à la date du recensement. Une enquête basée sur un échantillon issu du recensement permet de

calculer des taux de ré-émigration entre l'enquête et le recensement, migrants de retour tirés dans l'échantillon non trouvés en raison d'une ré-émigration, la destination de cette nouvelle migration peut être fournie par des personnes de l'entourage : ré-émigration vers l'étranger ou migration interne. Cependant il s'agit alors de la ré-émigration entre le recensement et l'enquête et non entre le retour et le recensement ou l'enquête. Ces données seraient fournies par une enquête sur les migrants dans les pays d'accueil en les interrogeant sur leurs retours antérieurs. Les taux de retour et de ré-émigration calculés sont sujets aux problèmes des petits nombres en ce qui concerne la variabilité des résultats. Ainsi on aura intérêt à considérer des grands groupes d'âges plutôt que des âges détaillés ou des groupes d'âge quinquennaux. Si des données sont disponibles à différentes dates, il faudra certes considérer la significativité des résultats mais surtout interpréter les changements en fonction d'événements particuliers pouvant expliquer les variations importantes : retours massifs liés à des fins de contrats de travail en raison de conditions particulières (fermeture de certaines entreprises, travail saisonnier), croissance du chômage, expulsions massives ou politique de régularisations dans les pays d'accueil, etc.

1. - L'approche par les recensements

Outre les flux qui conditionnent son évolution, il faut connaître les principales caractéristiques de la population étudiée. Le but de l'étude était, à cette étape de l'utilisation de données de recensement, de répondre à un certain nombre de questions simples de manière fiable et représentative malgré les petits nombres. Ces questions étaient d'ordre géographique: le pays d'où revenaient les migrants, se réinstallaient-ils dans leur île ou village de départ; démographique: sexe, âge ; et socio-économique: niveau d'études, activité et profession. Cela nécessitait dans certains cas, une comparaison à une population témoin ; nous avons retenu la population native totale, plutôt que les seuls natifs non migrants. Ce choix s'explique par le fait que la population native ayant répondu résider dans le pays à la question sur la résidence antérieure ne représente pas stricto sensu des non migrants, mais seulement dans le sens du recensement, puisque certains ont pu migrer avant la date de la résidence antérieure ou entre cette date et le recensement. Il n'était donc guère possible de les considérer comme non migrants pour leur comparer les migrants de retour. Cependant, il aurait été plus facile de faire ressortir des différences significatives en utilisant cette population « non migrante » plutôt que la population totale. Cependant, il vaut mieux ne pas perdre de vue le sens de ce que l'on fait que de chercher à faire ressortir coûte que coûte des différences significatives.

Nous allons exposer la démarche que nous avons suivie en présentant quelques résultats concernant le niveau d'études, les taux d'activité (actifs employés) et la distribution par profession dans les trois pays étudiés. Nous essaierons de faire apparaître, au-delà du flou de la multiplicité des situations engendré par les petits nombres, différents sous-groupes, sinon significativement différents selon les tests statistiques, du moins nettement différenciés. Le but de l'analyse apparaît donc être, comme dans n'importe quelle étude démographique d'une population de grande taille, de définir des sous groupes ayant des caractéristiques et des comportements particuliers.

Les niveaux d'études

La comparaison des niveaux d'études des migrants de retour et de la population totale s'est faite à différents niveaux d'études selon les pays (tableau 1). Si des différences intéressantes apparaissent au niveau de « form 4 » à Kiribati¹, il n'en est pas de même aux Îles Cook où pratiquement tout le monde atteint ce niveau. Il faut donc choisir un autre seuil pour

¹ Les données publiées pour Tuvalu étaient regroupées et c'est au niveau de «form 5» que des différences apparaissent.

faire apparaître des différences, en l'occurrence 'form 6'. - Il n'existait pas jusque récemment d'université dans les pays insulaires et il n'est donc pas intéressant de considérer des niveaux d'études supérieures car seuls des migrants peuvent avoir ce niveau. - Cependant, les données de recensement ne fournissent pas d'information sur le niveau d'étude atteint avant la migration, ce qui serait nécessaire pour mesurer l'effet sélectif du niveau d'étude sur la migration et l'acquis pendant la migration. Il faut recourir pour cela à des enquêtes. Malgré les petits nombres, le choix judicieux des niveaux d'études considérés permet de maximiser les différences. Un choix judicieux des variables s'applique d'ailleurs aussi bien à des populations de taille importante qu'à de petites populations. On remarque à cette étape de l'analyse que les migrants de retour ont un niveau d'éducation plus élevé que l'ensemble de la population. Cependant ce n'est pas le cas des migrants revenus d'une destination particulière (Nauru). Nous reviendrons sur ce point ci-dessous.

TABLEAU 1.- NIVEAU D'ÉTUDES DES MIGRANTS DE RETOUR ET DE LA POPULATION TOTALE

Rarotonga, proportion (%) de personnes ayant le niveau fin d'études ou plus						
	hommes			femmes		
	migrants de retour		population totale.	migrants de retour		population totale.
20-29	30		20	32		21
30-39	25		19	21		15
40-49	16		12	6		6
20 +	20		14	18		12
20 + (N)	254		2513	250		2426
Kiribati, proportion (%) de personnes ayant le niveau de « form 4 » ou plus						
	tous migrants de retour	de Nauru	population totale	tous migrants de retour	de Nauru	population totale
20-29	46	10	12	41	15	12
20 +	28	9	10	25	9	12
20 + (N)	586	254	16275	399	261	17238
Tuvalu, proportion (%) de personnes ayant le niveau de « form 5 » ou plus						
20-29	23	18	20	67	50	19
20 +	19	11	17	31	18	10
20 + (N)	213	27	2268	67	22	2838

Source : Exploitations spéciales des recensements de Kiribati en 1990, Tuvalu et des Îles Cook en 1991.

Lire : 30 % des migrants de retour âgés de 20-29 ont atteint la classe « form 6 » ou au-delà.

L'activité

Les taux d'activité font également apparaître des différences intéressantes entre les migrants de retour et la population totale. Il est évident, que la prise en compte du chômage n'est pas possible dans les économies où il n'existe pas d'assistance aux chômeurs et où les critères pour classer les actifs sans emploi en chômeurs ou en inactifs sont peu précis. Par ailleurs, les nombres de chômeurs seraient encore plus petits et plus aléatoires que ceux des actifs et il vaut donc mieux considérer ces derniers. Il y a aussi intérêt à considérer les seuls actifs employés dans le secteur formel, pour faire ressortir une différence d'accès à l'emploi rémunéré des migrants par rapport à la population totale, la plupart des hommes ayant au moins une activité d'agriculteurs ou de pêcheurs de subsistance. A nouveau, à l'opposé de ce qu'on pourrait attendre dans le cas de

petites populations, nous avons divisé la population étudiée en sous-groupes plus petits mais plus homogènes, faisant apparaître des différences plus nettes qu'entre l'ensemble des migrants de retour et la population totale. Nous avons ainsi considéré des groupes d'âges particuliers : les jeunes adultes, (aux âges plus élevés, la prédominance des activités de subsistance, coexistant ou non avec un emploi formel, rend la comparaison difficile), distingué les milieux : retour dans l'île principale ou dans les îles éloignées, où les activités de subsistance sont plus fréquentes, et à nouveau l'origine des migrants pour Kiribati et Tuvalu (tableau 2).

TABLEAU 2 : PROPORTIONS (%) DE PERSONNES EMPLOYÉES DANS LE SECTEUR FORMEL, SELON L'ORIGINE ET DES MIGRANTS DE RETOUR ET LA POPULATION TOTALE

Îles Cook	H	F	Kiribati	H	F	Tuvalu	H	F
Rarotonga			South Tarawa			total		
migrants '90' *	76	67	tous migrants	67	46	tous migrants	19	20
migrants '86' **	81	63	de Nauru	52	13	de Nauru	22	9
			de Fidji	71	75	de Fidji	55	21
			de « autres »	71	65	Austr., NZ, etc.	39	36
						de « autres »	4	
population totale	91	72		58	28		38	17
ÎLES EXTÉRIEURES			ÎLES EXT.					
migrants '90' *	36	11	tous migrants	49	26			
migrants '86' **	54	30	de Nauru	37	12			
			de Fidji	78	75			
			de « autres »	60	38			
population totale	70	30		43	15			
* migrants de retour, à l'étranger en 1990.								
** migrants de retour, à l'étranger en 1986.								
Îles Cook (25-34 ans), proportion de personnes employées, à leur compte ou dans la population active à plein temps ou temps partiel.								
Kiribati (25-34 ans), proportions de personnes dans le «cash sector ».								
Tuvalu (20+ ans), proportion de personnes qui ont travaillé pour un salaire, un bénéfice ou dans une entreprise familiale la semaine précédant le recensement.								
lire : 76% des migrants (absents en 1990) de retour à Rarotonga étaient actifs employés au recensement de 1991.								

Finalement, il apparaît que les migrants de retour à Kiribati et Tuvalu présentent des taux d'activité (actifs employés) inférieurs ou supérieurs à ceux de l'ensemble de la population selon le pays d'où ils reviennent et que l'importance de l'écart varie selon qu'on considère l'île principale et les îles extérieures. Cette méthode fait donc ressortir des différences relatives parfois plus importantes que celles entre tous les migrants de retour et la population totale, montrant qu'on peut, et doit, caractériser des sous-populations parmi les petites populations.

Professions

Vu les petits nombres, il est exclu d'utiliser des distributions fines des professions, un classement en 10 groupes est le maximum et aussi le minimum pour donner une image assez précise. - Cependant, il peut encore être nécessaire de regrouper les classes pour calculer des tests statistiques.

TABLEAU 3 : DISTRIBUTION DES MIGRANTS DE RETOUR ET DE LA POPULATION TOTALE PAR PROFESSION (%).

RAROTONGA	Hommes		Femmes	
	Migrants de retour	Population totale	Migrants de retour	Population totale
patrons, administrateurs	13	11	8	8
professions libérales	12	9	20	17
techniciens	7	9	4	5
personnel de bureau	7	8	27	27
employés des services	9	9	23	21
agriculteurs, pêcheurs	9	7	2	3
ouvriers de production	45	48	16	18
total	100	100	100	100
(N)	200	2460	157	1659

KIRIBATI SOUTH TARAWA	Hommes					Femmes	
	Tous migrants de retour	De Nauru	De « autre »	De Fidji	Population totale	Migrants de retour	Population totale
patrons, administrateurs	4	0	2	6	3	7	1
professions lib. techn.	37	14	39	58	17	41	27
personnel de bureau	6	3	3	18	11	23	29
personnels de commerce	2	0	1	3	3	4	8
employés des services	4	11	1	6	12	5	11
agriculteurs, pêcheurs	5	6	5	0	10	4	2
ouvriers de la production	40	64	47	6	42	14	19
autres	2	3	0	3	1	1	0
total	100	100	100	100	100	100	100
(N)	180	36	94	37	3221	75	1648

TUVALU	Hommes		Femmes	
	Migr. de retour	Pop. totale	Migr. de retour	Pop. totale
patrons, administrateurs	5	9	0	3
professions lib. techn.	23	18	71	27
personnel de bureau	10	9	21	28
personnels de commerce	8	6	0	12
employés des services	0	9	0	10
agriculteurs, pêcheurs	15	8	0	2
ouvriers de production	40	40	7	17
total	100	100	100	100
(N)	40	929	14	539

Les distributions professionnelles des migrants de retour et de la population totale ne présentent pas de différences marquées à Rarotonga (Îles Cook) (tableau 3). Cependant, les migrants de retour sont plus souvent dans les professions élevées (25% contre 20% quand on regroupe les deux premiers groupes de professions). On remarque que ceci est cohérent avec un niveau d'études plus élevé des migrants de retour. Les situations sont plus variées à Kiribati où les migrants de retour de Fidji sont beaucoup plus souvent dans les professions libérales ou techniciens. A l'opposé les migrants de retour de Nauru sont plus souvent ouvriers. Les migrants de la catégorie « autre » ont une distribution à deux pôles, correspondant aux migrants de retour d'Australie et de Nouvelle Zélande qui ont des professions élevées et aux marins qui sont souvent ouvriers après leur retour.

Les différences observées vis à vis de l'éducation, l'activité et la profession ne sont pas toujours statistiquement significatives. Il faut alors se demander ce que reflète les résultats. On observe alors une certaine logique ou cohérence qui permet de définir très clairement des types de migrants à Kiribati et Tuvalu. Les migrants de retour des pays développés de la ceinture pacifique ou de Fidji (où se trouve l'Université du Pacifique Sud) ont un niveau d'études secondaires plus élevé (certains ont migré pour faire des études supérieures), ont fréquemment un travail et sont dans les professions élevées. D'un autre côté, les migrants revenant de Nauru ont un niveau d'études faible et trouvent difficilement un emploi dans le secteur formel (ils ont travaillé comme ouvriers, souvent non qualifiés, à la mine de phosphates et ne trouvent pas de travail équivalent à Kiribati ou Tuvalu), cependant, ils sont souvent indépendants travaillant dans le bâtiment, la mécanique, les transports ou le petit commerce. Ceux qui ont migré comme marins sur des bateaux allemands ou japonais ont aussi un niveau d'études faible, et ils ne cherchent guère à travailler au retour, attendant seulement un nouveau contrat, tout au moins pour les plus jeunes. Ainsi, malgré des différences non significatives vis à vis d'une variable particulière, on peut définir des sous-populations parmi les migrants de retour, des types de migrants, le recoupement de quelques caractères pertinemment choisis finissant par établir un faisceau de présomptions traduisant l'existence de sous-populations aux comportements différents et ayant leur logique propre.

Il est alors remarquable que malgré des effectifs encore beaucoup plus petits, Tuvalu présente les même types de différences entre migrants de retour et population totale vis à vis des trois phénomènes considérés, montrant qu'une réalité se cache derrière ces petits nombres, même s'il est difficile ou impossible de la rendre évidente par des tests statistiques, notamment dans le cas de Tuvalu. Ajouter les données de Kiribati, Tuvalu, et aussi des Îles Cook, pour améliorer quelques tests n'est alors pas incorrect.. Il faut alors mettre l'accent sur le fait que nous sommes avec les recensements en présence de données exhaustives et non d'échantillons plus ou moins biaisés. Dans le cadre d'une enquête, la légitimité de cette approche est plus contestable.

Les petits nombres imposent cependant des limites. Il n'a pas été possible de considérer le rôle des durées sur les taux d'activité. Certes, en l'absence de l'année du retour, les données nécessaires n'étaient pas disponibles. Si elles l'avaient été, on aurait pu faire des regroupements (retours de moins de 2 ans, de 2 ans ou plus) mais il aurait sans doute fallu renoncer à considérer cet aspect dans les plus petites populations. - Cependant les données selon la résidence 1 an antérieurement au recensement aux Îles Cook, montrent sans doute un rôle de la durée sur l'obtention d'un emploi (tableau 2).

Des études plus approfondies nécessiteraient le croisement de variables : diplôme et activité ou profession. Il n'est pas possible de descendre à ce niveau avec d'aussi petits effectifs. La profession est très corrélée avec le diplôme et cela est vrai pour les migrants aussi bien que pour l'ensemble de la population. Il ne semble donc pas possible de faire ressortir un avantage des migrants à ce niveau et il faut s'en tenir à considérer que leur meilleure situation professionnelle est liée à une qualification plus élevée, sans pouvoir faire apparaître, à

qualification égale, un effet spécifique de la migration qui resterait à interpréter : préférence des employeurs ou rôle d'une meilleure expérience professionnelle des migrants.

2. - L'approche par une micro-enquête.

Si les données de recensement, malgré les petits effectifs, ont pu faire apparaître des spécificités et même des types variés de migrants de retour, alors une enquête devrait permettre de préciser les points que le recensement laisse dans l'ombre, comme les itinéraires migratoires, professionnels, le bénéfice de la migration. C'est dans ce but que nous avons réalisé une enquête aux Îles Cook. Le choix de ce pays, plus petit que Kiribati, tenait au fait que ce type de migration est plus représentatif des migrations dans le Pacifique que la migration par contrats de travail qui prédomine à Kiribati et Tuvalu.

Le choix du type d'enquête

Le but de l'enquête était donc, en plus de connaître la composition de cette population, déjà en partie fournie par le recensement, de définir et d'analyser les itinéraires et les processus qui amènent les migrants de retour à être différents de l'ensemble de la population. La diversité des situations peut être importante dans une petite population, comme l'ont déjà montré les recensements, et il en est de même des itinéraires des migrants. Il est donc nécessaire d'en rendre compte et c'est pour cela que nous avons choisi d'effectuer une enquête dans l'ensemble des migrants de retour plutôt que d'isoler une catégorie particulière, par exemple une profession, pour une analyse dite « qualitative » ou « approfondie ».

L'étude d'une seule catégorie de migrants n'aurait fourni ni une image de la diversité, ni des résultats applicables à l'ensemble des migrants de retour, même si elle semblait plus simple à réaliser et pouvoir fournir des résultats plus cohérents car les individus seraient moins dispersés. En fait il faut choisir ce que l'on veut étudier, les migrants de retour dans leur ensemble, avec la variété des situations et des problèmes qui les caractérisent, ou une catégorie particulière, les cadres par exemple. Mais sous prétexte que les petits nombres rendent l'étude de l'ensemble difficile, il ne faut pas essayer d'en donner une image biaisée à partir d'une sous-population particulière et dont on sait qu'elle n'est pas représentative de l'ensemble. Cela va à l'encontre de l'exigence fondamentale de représentativité, base des statistiques et de la démographie. Pourtant, c'est souvent ce type d'étude qui est fait. On évite alors les problèmes d'échantillonnage et on interroge un certain nombre de cadres, censés représenter une forte proportion de cette profession. Cependant, la sélection est très biaisée, du point de vue même de la seule population des cadres, - des médecins ou des infirmières pour prendre d'autres exemples. Ainsi, dans ce type d'études, on choisit habituellement les enquêtés dans un fichier d'adresses des personnes en activité, un annuaire professionnel, voire téléphonique, ou pire, on contactera des personnes dont l'adresse aura été communiquée de bouche à oreille - une méthode répandue dans les études des populations insulaires. Avec les fichiers professionnels, listes téléphoniques, on perd les anciens cadres ou infirmières qui ne travaillent plus ou ont changé de profession, biaisant les résultats sur la mobilité professionnelle, corollaire de la mobilité géographique à étudier pourtant en prédilection. Dans le cas du bouche à oreille, pratiquement toutes les personnes en échec professionnel ou social se trouvent éliminées car on ne mentionne que les cas de réussite. Ces méthodes d'enquête ont donné lieu à de nombreux travaux d'étudiants et aussi à des articles de chercheurs, principalement remarquables, quand elle ne reflètent pas des biais systématiques, par leurs résultats contradictoires.

Sur de petites populations, les biais peuvent devenir rapidement importants et rédhibitoires pour la représentativité de l'étude. La réflexion sur le choix de la population à étudier, sur le tirage de l'échantillon est donc primordiale pour éviter tout biais systématique.

Compte tenu des contraintes budgétaires, nous avons choisi d'enquêter 100 migrants de retour. Parmi les 401 migrants de retour âgés de 25-49 ans à Rarotonga, nous avons tiré un échantillon aléatoire de 100 personnes et une liste complémentaire de 50 noms pour remplacer les ré-émigrants, les refus et les personnes hors champs². L'échantillon tiré était représentatif selon le sexe et l'âge, mais celui des enquêtés effectivement interviewés l'était un peu moins selon l'âge, les jeunes ayant ré-émigré plus fréquemment que les personnes âgées de plus de 40 ans.

TABLEAU 4 : RÉPARTITION (%) DES MIGRANTS PAR PROFESSION EN NOUVELLE ZÉLANDE ET APRÈS LE RETOUR, À L'ENQUÊTE ET AUX RECENSEMENTS.

profession	1er emploi en NZ		Dernier emploi en NZ		Nés aux Îles Cook		1er emploi après le retour		Migrants de retour à Rarotonga	
	(enquête)		(enquête)		(recensement de NZ, 1986)		(enquête)		(recensement de 1991)	
	H	F	H	F	H	F	H	F	H	F
cadres, patrons	0	0	3	0	1	0	19	2	13	8
prof. libérales	8	8	5	10	3	5	5	23	12	20
techniciens	5	10	8	12	(1)		19	13	7	4
secrét. bureau	5	20	5	20	4	14	5	23	7	27
services, commerce	3	10	5	12	7	27	2	29	9	23
agri. pêche	0	2	3	0	3	2	9	0	9	2
ouvriers	79	49	72	47	82	51	40	10	45	16
total (N)	38	49	39	49	4857	2847	42	52	200	157

(1) inclus avec les professions libérales.

Les itinéraires professionnels

Il est remarquable de constater que l'échantillon est assez représentatif de l'ensemble des migrants de retour recensés en ce qui concerne la distribution par profession lors du recensement, après le retour (tableau 4)³, et que la distribution des migrants par profession pendant la migration est aussi assez semblable à celle des nés aux Îles Cook recensés en Nouvelle Zélande en 19864. Ainsi, l'échantillon donne sans doute une bonne image de la mobilité professionnelle autour du retour. - Il est déjà remarquable que malgré des distributions non significativement différentes des premiers et derniers emplois à l'étranger, il apparaisse une évolution cohérente avec l'accès des migrants à des emplois plus élevés avec le temps au cours de la migration.

L'approche la plus simple des changements d'emplois autour du retour par une matrice croisée des professions, en 10 classes, pendant la migration et après le retour est difficile à analyser en raison des petits effectifs et a dû être abandonnée. Un regroupement en trois

² Personnes ayant déclaré résider à l'étranger à la question sur la résidence antérieure mais n'y étant qu'en court séjour ou en vacances, ou personnes s'étant déclarées résidant des Îles Cook au recensement, mais n'étant à l'époque que de passage, ces dernières sont rentrées comme résidents ensuite.

³ Il en est de même pour les taux et les secteurs d'activité, montrant que l'échantillon est représentatif de l'ensemble des migrants de retour malgré sa petite taille ; il représente certes un fort taux de sondage.

⁴ Cependant, il peut y avoir eu une surévaluation de leur profession pendant la migration par les migrants, résultant dans des proportions plus élevées de professions libérales, assistants de ces professions et techniciens, alors que ces certains pouvaient n'être qu'employés de bureau ou de commerce.

classes aurait donné des effectifs plus importants dans chaque case du tableau, mais les passages d'une grande classe à l'autre auraient été peu fréquents et on aurait beaucoup perdu en finesse de l'analyse. Pour analyser la mobilité professionnelle, nous avons retenu un classement des professions en 10 classes hiérarchisées et calculé la différence entre les professions pendant la migration et après le retour et mesuré ainsi une montée ou une baisse dans l'échelle des professions. - Les mêmes calculs ont été effectués pour la profession avant la migration (pour ceux qui avaient travaillé avant leur départ) et la profession à l'étranger. Ainsi, nous avons considéré des itinéraires professionnels en essayant d'en découvrir les dynamiques.

TABLEAU 5 : TYPES D'ÉVOLUTIONS PROFESSIONNELLES AVANT, PENDANT ET APRÈS LA MIGRATION À L'ÉTRANGER.

3 OBSERVATIONS (AVANT, PENDANT, APRÈS)

	Hommes	Femmes	Total	%
/ /	0	0	0	0
\ /	2	8	10	21
\ _	3	2	5	10
_ /	5	3	8	17
_ _	8	11	19	40
- \	1	0	1	2
/ -	1	1	2	4
/ \	1	1	2	4
\ \	0	1	1	2
total	21	27	48	100

2 OBSERVATIONS (PENDANT, APRÈS)

/	8	9	17	45
_	9	9	18	47
\	2	1	3	8
total	19	19	38	100

2 OU 3 OBSERVATIONS (PENDANT, APRÈS)

Profession après le retour	Supérieures	Intermédiaires	Ouvriers	Total %
/	20	16	0	42
_	13	11	21	52
\	0	2	3	6
total	33	29	24	100

/ ascension professionnelle, _ stabilité, \ baisse

On observe une fréquence élevée (42%) de montées dans l'échelle professionnelle après le retour (tableau 5 - 3ème partie). Pour l'expliquer plus en détail, nous sommes passés alors - seulement alors - à une analyse qualitative, relisant les questionnaires des personnes présentant une ascension professionnelle. Il apparaît deux raisons principales, un certain nombre d'ouvriers se mettent à leur compte, deviennent indépendants ou chef d'entreprise s'ils ont des ouvriers, dans leur ancien domaine : bâtiment, mécanique, ou après une reconversion dans les services liés au tourisme ; c'est aussi parfois le cas de personnes travaillant antérieurement dans le service public. Une autre manière est de rester salarié, mais d'obtenir une position plus importante, plus élevée ou de supervision par exemple, principalement en raison d'études ou

de formations suivies à l'étranger mais parfois simplement du fait de l'expérience acquise dans un milieu professionnel plus exigeant pendant la migration. Il faudrait des effectifs plus importants pour considérer d'autres facteurs expliquant l'ascension professionnelle des migrants de retour. On peut néanmoins observer que seulement un sur deux des entrepreneurs a suivi des études à l'étranger et un sur deux également avait travaillé avant le départ à l'étranger, mais il n'est pas possible de voir l'impact de ces situations qui se recouvrent en partie. Cependant cette analyse qualitative n'empêche pas de garder un aspect statistique en classant les cas d'ascensions professionnelles selon les types définis, pour voir la variété des processus : devenir entrepreneur avec ou sans formation à l'étranger, monter dans la hiérarchie principalement mais pas exclusivement grâce à des études à l'étranger.

Pour l'ensemble des migrants, l'ascension professionnelle est le cas de 38% des migrants ayant travaillé avant, pendant et après la migration (3 observations), 45% pour ceux qui n'avaient pas travaillé avant la migration (2 observations) et 42% pour l'ensemble. Mais c'est 60% (20/33) des personnes dans le groupe de professions les plus élevées et 55% dans le groupe intermédiaire qui y sont venues à la suite d'une ascension professionnelle après leur retour. Ceci montre comment une étude limitée à une classe sociale particulière peut donner des résultats différents et induire en erreur si on l'appliquait à l'ensemble des migrants de retour. On remarque aussi pour l'ensemble des migrants une grande variété d'itinéraires avec des fréquences très variables (cf. première partie du tableau 5).

Ainsi, le recours à l'analyse qualitative ou la limitation de l'étude à une catégorie particulière dans le but de fournir plus de détails, une description plus fidèle, apparaissent finalement donner des résultats biaisés ou non représentatifs et être, de plus, réductrice de la diversité, ce qui est toujours le cas lorsqu'on se limite à un sous-groupe particulier ou présente un nombre limité d'histoires de vie. Il n'y a rien de plus ennuyeux, et non représentatif, que les études qualitatives qui résument trois ou quatre histoires de vie, toutes semblables dans leurs grandes lignes, oubliant que d'autres situations sont tout aussi fréquentes sinon plus.

Conclusion

Lorsqu'on travaille sur de petites populations, il faut distinguer selon qu'il s'agit de données exhaustives ou d'enquête. Les secondes sont affectées par des biais de sondage alors que les premières représentent l'ensemble des individus. Des différences, même non statistiquement significatives, peuvent être intéressantes à noter et s'organiser pour identifier des sous-populations présentant un comportement spécifique et cohérent vis à vis de diverses variables. Cela représente sans aucun doute une avancée dans la connaissance des populations, permettant des actions et des politiques visant des groupes spécifiques. - Cette étude a fait reconsidérer aux pouvoirs publics de Kiribati et Tuvalu, les problèmes du prochain retour de l'ensemble des migrants à Nauru, en raison de la fin prévue de l'exploitation des phosphates dans cette île.

Lorsqu'on travaille sur des échantillons, il faut d'abord évidemment qu'ils soient représentatifs. Toute étude dite qualitative digne de ce nom ne peut ignorer cette exigence statistique, mais on doit aussi veiller au fait que les biais de sélection peuvent devenir importants sur de petites populations. La présentation des résultats doit aussi refléter la diversité du réel et non privilégier quelques types, parfois un seul type, de comportements particuliers, à moins que ce ne soit le champs spécifique de l'étude, mais alors il faut mettre en garde contre toute extension abusive des résultats. L'analyse qualitative semble pouvoir n'être qu'un complément à l'analyse statistique, celle-ci restant possible même sur de petits nombres. Il apparaît alors que considérer des dynamiques, des itinéraires, donne de meilleurs résultats que le morcellement du réel en croisant l'ensemble des modalités. Le but est cependant le même que pour l'étude de grandes populations: le découpage du réel en sous-ensembles présentant des comportements particuliers.

La démographie appliquée aux populations de petite taille. Le point de vue du statisticien-démographe

Philippe WANNER¹

Office Fédéral de la Statistique, Berne²

1. Introduction

Le statisticien - ou démographe - qui produit des informations chiffrées sur l'état et les phénomènes associés à l'évolution de la population est mal à l'aise lorsqu'il s'agit de fournir au public ou administrations locales une "information" (sous forme d'un tableau, d'une série de chiffres ou d'un indicateur) portant sur des petites populations ou des phénomènes rares et soumis à de fortes fluctuations aléatoires. En effet, le rôle du statisticien ne se limite pas à la production et la diffusion de données. Il lui appartient aussi de s'assurer de l'utilisation correcte de ces données et notamment d'éviter l'emploi abusif ou erroné des informations publiées.

Ces précautions sont indispensables lorsque le statisticien-démographe met en place un système d'informations qui sert directement à la gestion de petites populations, et plus encore lorsque ces petites populations évoluent rapidement, nécessitant ainsi des adaptations de la planification scolaire, sanitaire, ou de la prise en charge des personnes âgées, par exemple.

Or, il faut reconnaître qu'en matière de démographie des petites populations, la littérature n'offre pas beaucoup de recommandations à l'usage du producteur de statistiques (voir cependant Lapham et Spencer, 1982 ; Ruzicka, 1994 pour les problèmes liés à l'estimation de la mortalité dans les petites populations), alors même que l'usage des données démographiques se rapportant à de petites entités géographiques paraît se multiplier (Russell, 1984) et se multipliera probablement encore.

Dans la présente contribution, nous adoptons une approche de type illustrative et nous nous mettons à la place d'un statisticien chargé de répondre aux besoins d'un gestionnaire local (il peut s'agir du maire d'une localité ou d'un chef d'administration par exemple). La situation serait identique dans le cas d'un statisticien engagé par une entreprise, une école, une association ou une corporation de métiers, secteurs dans lesquels l'information démographique de type statistique peut être très pertinente (Hugo, 1991)³.

Notre contribution s'articule autour de quatre parties. Dans une première partie, nous présentons quelques objectifs d'un système d'information démographique dans les populations de taille limitée, et les principaux problèmes qui peuvent y être associés. Dans une seconde partie, nous énumérons quelques outils permettant de résoudre en partie les problèmes d'analyse des petites populations, et de limiter l'utilisation abusive de données non représentatives. Nous présentons ensuite un exemple tiré de la réalité, montrant jusqu'où un système d'informations locales peut être utilisé et quels sont les dangers d'interprétations erronées. Finalement, nous proposons quelques recommandations dans un domaine où, il faut bien en convenir, le bon sens l'emporte souvent sur l'application de méthodes rigides.

¹ E-mail Philippe.Wanner@bfs.admin.ch

² L'auteur remercie MM Stéphane Cotter et Alexis Gabadinho de l'Office fédéral de la statistique pour la pertinence de leurs remarques à propos de ce texte

³ Notamment dans le cadre des études de marché. On peut aussi citer la «démographie médicale» qui en Europe est une discipline à part entière, développée par la corporation des médecins et responsables de la santé.

Nous nous limitons ici au contexte européen, caractérisé par un système d'informations démographiques de type exhaustif (par l'état civil ou les recensements) et jugé le plus souvent de bonne qualité. La situation est plus complexe dans les pays où le système d'information statistique est déficient, et dans lesquels le problème de la qualité des données doit être pris en considération.

2. Les objectifs d'un système d'information démographique local et les problèmes qui y sont associés

Le statisticien étudiant la population ou son évolution, qu'il appartienne à un office régional ou local de statistique, à une administration, à un hôpital ou à une entreprise privée, poursuit l'objectif d'apporter une information claire au gestionnaire. Cette information peut être classée en deux catégories, selon l'usage auquel elle est destinée.

- *Une information en nombres absolus, dans le contexte d'une action déjà décidée* : par exemple, l'effectif de la population d'une commune, en vue de la redistribution de revenus fiscaux ; l'effectif du personnel d'une entreprise, pour la planification de la production et des coûts ; l'effectif des élèves d'une école, pour répondre aux besoins en personnel d'enseignement ou en matériel scolaire. Dans ce cas, l'information livrée ne souffre d'aucune discussion. Elle peut être fournie sans traitement statistique particulier et peut être utilisée telle quelle.
- *Une information « contrôlée », dans un contexte de planification, de compréhension des phénomènes sous-jacents à l'évolution démographique, de prise de décisions, ou de dépistage des problèmes prioritaires propres à déclencher une intervention spécifique* : C'est le cas qui nous intéresse ici. Les phénomènes mesurés sont le plus souvent exprimés sous forme de tendances, de taux ou d'indicateurs, afin de pouvoir être interprétés dans de bonnes conditions, après « contrôle » de l'âge notamment. On effectuera par exemple des projections de population locale afin de planifier les revenus fiscaux, une étude de la mobilité d'une entreprise afin de juger des conditions de travail, une estimation des mouvements migratoires futurs des couples avec des enfants afin de prévoir les besoins en équipements scolaires, ou un dénombrement des cas de maladies infectieuses (SIDA, rougeole) afin de contrôler le développement ou l'évolution d'épidémies⁴. Dans ces cas, il convient de « valider » les informations transmises, c'est-à-dire de s'assurer de leur signification, et de veiller à ce qu'elles ne conduisent pas à une interprétation erronée des phénomènes étudiés (lire à ce propos Volle, 1984).

Nous distinguerons au moins trois principales sous-catégories d'informations, en fonction des données produites et de leur type de recueil :

- Une information *traduisant directement le phénomène étudié* : par exemple un taux d'émigration pour l'étude de la mobilité intercommunale.
- Un indicateur *traduisant indirectement une dimension étudiée*. Il peut s'agir par exemple du nombre de condamnations pour usage de stupéfiants comme mesure de la consommation de drogues illégales à un niveau local, de l'absentéisme dans une entreprise comme mesure de la qualité des conditions de travail.
- Une information *collectée par une enquête non exhaustive*, par exemple une étude sur les conditions de vie dans une commune, les désirs des personnes résidentes sur les

⁴ Il faut noter qu'en Suisse, ce type d'informations à usage local est le plus souvent totalement occulté, notamment dans les petites localités. Ainsi, une nouvelle école sera souvent trop petite pour répondre aux besoins communaux avant même la fin de sa construction ; un bâtiment nouvellement construit sera inoccupé en raison de tendances migratoires qui n'ont pas été prises en compte. A priori, le secteur privé est plus sensible à l'évolution de la situation locale que le secteur publique.

équipements et programmes de loisirs ou encore les souhaits des citoyens en matière d'urbanisme.

Dans le chapitre suivant, nous présenterons quelques outils utiles pour présenter une information locale correcte et efficace. Nous classifions ces outils en fonction des possibilités d'utilisation selon les trois sous-catégories définies.

3. Quelques outils pour la validation d'une information locale

Le statisticien dispose de méthodes adaptées aux petites populations ou permettant de limiter les risques - ou au moins de les mesurer - d'utiliser une information qui ne serait pas pertinente. Ces méthodes sont les suivantes :

- *Le calcul d'un degré de signification ou d'un intervalle de confiance, à partir d'une loi de distribution statistique.* Dans le cas de données portant sur des populations exhaustives, il est plutôt discutable de calculer des tests de signification. Cependant, l'application de tels tests peut apporter une information utile sur les possibilités d'appréhender, en toute sécurité, la valeur obtenue (Duchêne et Thiltgès, 1993 ; Robitaille 1980a ; Robitaille, 1980b). Encore faut-il choisir une *loi de distribution correcte* (mais existe-t-il toujours une loi statistique qui puisse refléter la complexité des phénomènes démographiques ?) et être conscient que le degré de signification d'un indicateur n'est pas une information toujours pertinente⁵.
- Différentes méthodes statistiques ont été développées pour répondre aux problèmes touchant l'analyse des petites populations ou l'analyse de phénomènes rares. Sans vouloir entrer dans les détails méthodologiques, signalons quelques travaux. Anson (1988) a relevé l'avantage des indices comparatifs de mortalité (SMR - Standardized mortality ratios) pour l'étude de la mortalité des petites populations. Gage et al. (1986) ont proposé une méthode d'analyse de la mortalité, pour des populations inférieures à 1000 individus, basée sur les données de deux recensements. Étudiant la mortalité par cancer dans des petites populations, Manton et al. (1981 et 1987) ont développé des gradients spatio-temporels ou utilisé des facteurs de pondération tirés des valeurs nationales. Dans le domaine des projections, Tayman (1994) a discuté les méthodes applicables aux petites populations. Le recours à ces méthodes peut être, dans certains cas, très utile.
- Dans le même ordre d'idées, relevons *l'utilisation de méthodes d'estimation indirectes « classiques »*, développées pour les données déficientes, par exemple des tables-types pour l'estimation de l'espérance de vie d'une commune⁶. Plutôt que de calculer l'indicateur conjoncturel de fécondité à l'aide des taux de fécondité par année d'âge qui peuvent être soumis à de fortes variations, on pourra aussi recourir à des méthodes de type de celles de Brass, par exemple sur les enfants déjà nés. Il faut cependant tenir compte des biais et limites de chacune de ces méthodes indirectes.
- Le phénomène peut se poursuivre dans le temps ou concerner plusieurs entités géographiques, par exemple une région et pas seulement une commune. Suivant la nature du phénomène en question, le statisticien proposera *d'adopter plutôt une approche portant sur plusieurs années ou sur une région plus large*, afin d'augmenter la taille de la population étudiée et de valider une information qui, limitée à une année ou une localité, ne serait pas significative (voir Congdon, 1990).
- Un *regroupement*, non plus des périodes ou des communes voisines, mais *des populations présentant des caractéristiques socio-économiques identiques*, peut être envisagé. Un

⁵ Ainsi, si on parle d'une augmentation d'un taux de fécondité *significative à 95%*, on suppose que dans 95 cas sur 100, cette augmentation traduit un phénomène « réel » alors que dans les 5 cas restant, elle est le résultat du « hasard ».

⁶ Utilisé par exemple en Autriche pour l'établissement de perspectives régionales

gestionnaire qui voudrait comprendre la croissance naturelle et migratoire de sa ville pourrait ainsi se voir proposer l'étude de l'évolution démographique de plusieurs localités présentant les mêmes caractéristiques en terme de lieu d'habitation, de taille de la commune ou de structure par âge et catégorie socio-économique de la population résidante. Le recours à cet artifice aura pour conséquence de diminuer les fluctuations aléatoires, mais elle risque en même temps de limiter la compréhension de mécanismes spécifiques à la commune étudiée.

- Le recours à des données exhaustives, par rapport à des enquêtes présentant plus de problèmes méthodologiques mais parfois aussi des avantages (notamment de coût) est à encourager, principalement pour l'étude de phénomènes dont l'incidence est rare.
- Tonini (1988) propose des modèles complexes de l'économétrie (modèles ARIMA, approche de Box-Jenkins), qui pourraient dans certains cas être adaptés à notre problématique, mais qui, d'une part, doivent être utilisées avec une certaine prudence, d'autre part, nécessitent des connaissances statistiques poussées.
- Le démographe sous-estime généralement la possibilité d'utiliser des données auxiliaires, soit pour identifier un problème, soit pour le valider. Lorsque l'on se limite à une population de petite taille, de telles données sont généralement disponibles. Il s'agirait par exemple de renoncer à un taux de mortalité infantile pour s'intéresser à un indicateur de morbidité infantile (nombre de consultations d'un service de pédiatrie). Dans le même ordre d'idées, Westerling (1995) a proposé de ne pas se limiter à la cause principale de décès, mais l'ensemble des causes précisées dans le certificat de décès. Dans le cas de l'étude de la mobilité spatiale d'une commune, les tendances du marché immobilier pourraient être utiles comme indicateur de l'attractivité ou de la répulsivité qu'exerce une commune sur la population, en complément aux données sur la migration.
- Enfin et surtout le bon sens, la connaissance du contexte et de son évolution, apportent des éléments essentiels, qui ne peuvent être négligés par le statisticien. Ainsi, le plus souvent, une diminution spectaculaire de la proportion d'étrangers dans une localité s'explique par la fermeture d'une école privée internationale. Une augmentation de la mortalité trouve sa source dans l'ouverture d'un établissement médico-social. La présence d'un centre d'accueil des malades du SIDA explique une augmentation communale de la mortalité pour cette cause. Plus l'entité géographique étudiée est petite, plus ce type de phénomènes peut se produire. De même, des problèmes de recueil de l'information, qu'il faut prendre en compte, agiront de manière plus sensible dans les petites localités, alors qu'à moins d'un biais systématique, ils s'annuleront dans de grandes communes.

Ces exemples d'« outils » de travail ne sont pas exhaustifs. L'objectif est surtout de s'adapter à la situation prévalant étudiée et d'être capable de développer ses propres outils en fonction de ses besoins

Le tableau 1 récapitule les outils qui s'appliquent aux différentes situations présentées dans le chapitre précédant. L'essentiel du problème réside cependant dans le phénomène étudié, plutôt que dans la situation en matière d'informations disponibles. Ainsi, un exemple, adapté de la réalité helvétique, illustre la complexité des situations et l'impossibilité d'une approche systématique du problème.

TABLEAU 1 : RÉCAPITULATIF DES PRINCIPAUX OUTILS POUR L'ÉLABORATION D'UN SYSTÈME D'INFORMATIONS LOCALES, ET RECOMMANDATIONS D'UTILISATION

« Outil »	Situation ¹		
	information obtenue directement	indicateur approximatif	information non exhaustive
Calcul de niveaux de signification et d'intervalles de confiance	(X) ²	(X)	X
Utiliser des méthodes spécifiques aux petites tailles	X	X	
Utiliser des méthodes indirectes d'estimation	X	X	X
Étendre l'analyse à plusieurs années ou à une région	X	X	
Étendre l'analyse à d'autres localités présentant les mêmes caractéristiques	X	(X)	
Utiliser d'autres données exhaustives, disponibles notamment par des fichiers administratifs		X	X
Utiliser des méthodes statistiques complexes	(X)	(X)	(X)
Utiliser et comparer des indicateurs auxiliaires qui sont censés refléter le même phénomène		X	X
Tenir compte du contexte local et s'y insérer	X	X	X

¹ Se référer au chapitre 2 pour une définition précise de ces situations.
² X à conseiller, (X) à conseiller éventuellement.

4. Un exemple pratique

Une ville de taille moyenne (100 000 habitants) a vécu une diminution de sa population résidente. Celle-ci, entre le 1er janvier et le 31 décembre a diminué de 4,2%. Ces chiffres, publiés par l'Office national de statistique, inquiètent le statisticien chargé des problèmes de population. Celui-ci pourrait se poser les questions suivantes avant d'intervenir auprès des gestionnaires locaux.

1. Les informations sont-elles correctes et la tendance à la diminution est-elle réelle ou artéfactuelle ? L'explication de cette tendance peut-elle trouver sa source dans la définition de la statistique ? Par exemple :
 - une modification de la définition de la population résidente, par exemple la prise en compte ou la suppression d'une catégorie de résidents (fonctionnaires internationaux, étrangers dont le permis est à durée limitée, stagiaires professionnels) ; autre exemple, le passage d'une statistique basée sur le domicile *légal* (commune dans laquelle l'individu paie des impôts) à une statistique basée sur le domicile *économique* (comme dans laquelle l'individu passe le plus clair de son temps).
 - une révision du système national de la statistique de la population, qui tient compte de sources d'information remises à jour (un recensement récent, de nouvelles données administratives).
2. Si ces cas sont à écarter, le statisticien passe à une seconde étape. L'évolution observée est-elle significative et est-elle susceptible de refléter un phénomène plus profond, qu'il convient de comprendre sinon de maîtriser ? Le Tableau 1 présente quelques outils qui peuvent être utilisés pour notre situation, dans laquelle l'information traduit directement le phénomène étudié.

- Un calcul du degré de signification du taux de décroissance permet d'écarter l'hypothèse d'une fluctuation aléatoire.
 - Les données sur la population au cours des 20 dernières années indiquent que, non seulement la diminution de la population montre une nette cassure dans l'évolution démographique de la commune, mais qu'en plus, elle a été précédée par une diminution du rythme de croissance au cours des 5 années précédentes.
 - Il semblerait que dans la même année, d'autres localités de même taille aient connu une pareille évolution. Cette observation défend l'idée d'une tendance significative exprimant un phénomène réel qu'il convient de comprendre.
 - En revanche, les communes à la périphérie se sont caractérisées par une augmentation significative de leur population. Cette réalité pourrait soutenir la thèse d'une mobilité entre le centre et la périphérie.
3. Dans de nombreux cas, une tendance observée est soit artéfactuelle, soit non significative. Il convient ainsi d'apporter un soin particulier à ces deux premières étapes de *validation*, afin de ne pas s'engager dans un processus *explicatif* souvent coûteux, qui constitue l'étape suivante. *Quels sont les facteurs de cette soudaine diminution de la population ?* Plusieurs outils d'analyse permettent d'aborder cette étape.
- Une étude de l'évolution démographique des différents quartiers permet d'estimer la contribution respective de chaque quartier à cette diminution de la population. On observe alors deux phénomènes : d'une part les quartiers où la population est de type âgé se caractérisent soudainement par un solde naturel négatif non compensé par le solde migratoire, ceci en raison d'une augmentation du nombre de décès ; d'autre part, les quartiers les plus favorisés en terme de répartition socio-économique se caractérisent par un solde migratoire fortement négatif.
 - Une analyse d'un échantillon de personnes présentes en début d'année, tiré du fichier des habitants montre que la probabilité d'émigrer est la plus élevée chez les personnes dont le profil socio-économique est plutôt favorable, mais qui sont sans emploi, ainsi que chez les personnes nouvellement mariées. Les données de l'Office du chômage indiquent que, dans la région, le chômage est de type conjoncturel plutôt que structurel et la durée moyenne passée comme demandeur d'emploi ne dépasse généralement pas six mois. Les résultats d'une enquête sur la famille, menée dans une autre région ayant connu le même phénomène migratoire, ont été consultés. Il semblerait qu'une proportion élevée de couples vivant en milieu urbain et désireux d'avoir des enfants, recherchent un nouvel appartement plus spacieux, dans des endroits où le loyer est moins élevé et où le trafic automobile est limité.
 - Le calcul d'indices comparatifs de mortalité par quartier, pour les dix dernières années, montre que l'augmentation observée du nombre de décès est explicable - de manière significative - par la grippe, particulièrement virulente en début d'année. En revanche, les risques pour les autres causes de décès ne semblent pas avoir augmenté durant l'année étudiée.
4. Les facteurs responsables de la diminution de la population ayant été dépistés, il appartient au démographe d'attirer l'attention des gestionnaires locaux sur l'évolution observée et sur les conséquences de cette évolution, si elle venait à se poursuivre. Cette concertation entre démographe et politiciens pourrait aboutir aux résultats suivants :
- Le départ pour des communes voisines de jeunes personnes, nouvellement mariées ou temporairement sans emploi, entraîne différentes décisions : la création d'un groupe de travail chargé d'envisager des mesures d'aides au logement pour ces jeunes couples ; la mise sur pieds d'un programme de soutien aux jeunes chômeurs ; une réflexion sur la possibilité de créer quelques quartiers résidentiels à trafic limité ; une révision du

budget communal, les revenus fiscaux étant revus à la baisse, alors que les dépenses sociales exprimées en francs par habitant sont corrigées à la hausse ; enfin, la révision des projections démographiques et scolaires pour les dix prochaines années.

- Les conséquences sur la mortalité d'épidémies de grippe sont à la fois bien documentées et difficilement évitables. Néanmoins, on insistera, au cours des prochaines années, sur la promotion des campagnes de vaccination de la grippe dans la population à risques élevés, principalement les personnes âgées.
- Enfin, le maire chargera son office de statistiques de mettre en place un « tableau de bord démographique » permettant d'une part de prévoir de futures modifications de tendances à l'aide de *leadings indicators* et d'autre part de prévoir des « niveaux d'alerte » qui imposeraient la prise rapide de mesures en cas de problèmes associés à l'évolution de la population. Il est notamment décidé de recueillir régulièrement des données des différentes polycliniques sur la vaccination contre la grippe, des principales gérances sur le nombre d'appartements qui se libèrent ou qui se libéreront prochainement, ainsi que des données sur le taux de chômage. Lorsque l'indicateur atteint son niveau d'alerte, on proposera respectivement un communiqué de presse sur les avantages de la vaccination contre la grippe, l'entrée en vigueur d'un programme de subvention de loyer pour les jeunes couples avec enfants ou l'envoi d'une information complète sur l'aide offerte par la commune aux personnes sans emploi.

Le présent exemple illustre, de manière très simplifiée, les validations et analyses qui, d'une observation démographique, conduisent à une décision locale. Il fait référence à plusieurs outils de validation, d'analyse, et au recours à différentes sources auxiliaires. Le statisticien devrait pouvoir choisir les outils pertinents dans le contexte étudié, et ceux qui ne le sont pas. Dans notre cas, l'indicateur qui provoque une analyse statistique - le taux de croissance - traduit directement le phénomène étudié. La situation est beaucoup plus complexe lorsqu'un statisticien est chargé d'expliquer une évolution significative d'un indicateur *approximatif* (par exemple le nombre de condamnations pour usages de drogues illégales comme indicateur de la consommation de stupéfiants) ou un phénomène observé au cours d'une enquête (pour rester dans le domaine de la santé publique, une proportion plus élevée de personnes consommant des médicaments dans la ville, par rapport aux localités rurales) qui nécessitent une validation plus poussée.

5. Quelques recommandations

Afin d'apporter une information précise et correctement interprétable, le statisticien s'occupant d'un système d'informations locales devrait suivre les règles essentielles résumées ci-dessous.

1. Faire une investigation complémentaire à chaque fois qu'un phénomène démographique inattendu est observé
2. Donner l'information la plus transparente et la plus lisible possible : il ne s'agit pas « d'étouffer » l'utilisateur d'une multitude de chiffres, mais de proposer quelques indicateurs sensibles et spécifiques, validés, comparables dans le temps ou dans l'espace, formant un ensemble cohérent. Il convient également, dans un souci d'honnêteté, de préciser le nombre de cas sur lequel repose le calcul de chacun des indicateurs. Enfin, il paraît pertinent de décrire l'évolution et l'état démographique de manière plutôt générale, et de ne recourir à une analyse plus fine qu'en cas de nécessité.
3. Présenter des informations cohérentes avec la taille de la population et éviter de présenter celles qui ne sont pas applicables à cette taille. On ne calculera pas une table de mortalité pour une petite commune, mais plutôt des indices standardisés de mortalité. Autre exemple, on ne présentera pas des taux de fécondité par âge mais plutôt par groupe d'âges.
4. *Être honnête et neutre*, c'est-à-dire de laisser s'exprimer les chiffres sans vouloir les faire parler. Le statisticien est au service des chiffres et non l'inverse. Dans le cas des petites

populations, pour lesquelles les phénomènes démographiques sont bien moins connus que pour l'ensemble d'un pays, cette règle est cruciale.

5. *Préciser clairement les biais possibles et les limites des résultats obtenus.* Une information statistique ne se résume pas à un chiffre ou une série de chiffres, mais doit aussi comprendre des indications claires sur le processus de production de l'information et les limites de la signification des informations produites selon une certaine problématique.

BIBLIOGRAPHIE

- ANSON J. (1988), Mortality and living conditions : relative mortality levels and their relation to the physical quality of life in urban populations, *Social Science and Medicine*, 27(9), 901-910.
- CONGDON P. (1990), Issues in the analysis of small area mortality, *Urban Studies*, 27(4), 519-536.
- DUCHÊNE J., THILTGÈS E. (1993), La mortalité des plus de 15 ans en Belgique : les disparités régionales en 1985-1987, *Espace, Populations, Sociétés*, 4, 73-85.
- GAGE T.B., DYKE B., MACCLUER J.W. (1986), Estimating mortality level for small populations : an evaluation of a pair of two-census methods, *Population Studies*, 40(2), 263-273.
- HUGO G. (1991), What population studies can do for bussiness, *Journal of the Australian Population Association*, 8(1), 1-22.
- LAPHAM R.J., SPENCER B.D. (1982), Small-area analysis, In : J.A. Ross, *International encyclopedia of population*, Vol. 2, New York, Free Press, pp. 607-614.
- MANTON K.G., STALLARD E. (1981), Methods for the analysis of mortality risk across heterogeneous small populations : examination of space-time gradients in cancer mortality in North Carolina counties 1970-1975, *Demography*, 18(2), 217-230.
- MANTON K.G., STALLARD E., WOODBURY M.A., RIGGAN W.B., CREASON J.P. MASON T.J. (1987), Statistically adjusted estimates of geographic mortality profiles, *Journal of the National Cancer Institute*, 78 (5), pp 805-815.
- ROBITAILLE N. (1980a), Application des tests statistiques à des populations exhaustives : l'exemple de la mortalité, *Cahiers Québécois de Démographie*, 9(3), p. 105-116.
- ROBITAILLE N. (1980b) Commentaire relatif à l'article « application des tests statistiques à des populations exhaustives » : l'exemple de la mortalité, *Cahiers Québécois de Démographie*, 9(3), p. 131-132.
- RUSSEL C. (1984), The business of demographics, *Population Bulletin*, 39(3), 1-40.
- RUZICKA L.T. (1994), Progress in demographic methodology, *Journal of the Australian Population Association*, 11(1), 21-31.
- TAYMAN J. (1994), Small area demographic forecasts, *Applied Demography*, 9(1), 2-4.
- TONINI G (1988), Analisi e previsione delle nascite mensili con modelli della classe ARIMA, *Genus*, 44(3-4), 37-82.
- VOLLE M. (1984), Le métier de statisticien, Economica, Paris.
- WESTERLING R. (1995) Small-area variation in multiple causes of death in Sweden - a comparison with underlying causes of death, *International Journal of Epidemiology*, 24(3), 552-558.