Bernard AUBRY
INSEE - Alsace

23 novembre 2004

Confrontation de deux sources

Les flux résidentiels vus par les recensements et les enquêtes emploi

1982 - 1999

Cette note a deux objectifs.

En premier lieu, il s'agit de présenter une application d'un travail de fusion de deux fichiers. Un fichier extrait de Saphir (historique des recensements) et un autre extrait des enquêtes annuelles de l'emploi. L'exemple concerne les migrations résidentielles entre l'Ile-de-France et la province, mais on pourrait appliquer la même démarche à d'autres domaines, les immigrés, les catégories socio-professionnelles par exemple.

En second lieu, on voudrait, de façon élémentaire et partielle, aborder le difficile problème de la comparaison des flux migratoires quand la période de référence est variable. On s'appuie pour cela sur les travaux de Daniel Courgeau repris par Franck l'Hospital dans un rapport de stage. La question se pose depuis longtemps, chaque fois que l'on est amené à comparer les flux migratoires de recensements successifs. Elle se posera prochainement avec plus d'acuité encore puisque l'on sera tenté de comparer les flux migratoires obtenus après dépouillement du nouveau recensement et qui portent sur des durées de cinq ans, avec les flux précédents qui portaient sur des durées variables, de 6 à 9 ans.

Plan

Introduction
Première partie - **Comparaison des sources - échanges Ile-de-France province**
Les données des recensements confrontées aux enquêtes annuelles sur l'emploi (1982-90 et 1990-99)
Deuxième partie - **Examen du modèle migrants-migrations**
Les conséquences de la durée inégale des périodes de référence sur les flux
Conclusion
Annexe méthodologique - Bibliographie

Introduction

Les migrations résidentielles jouent un rôle essentiel dans la dynamique des territoires. Leur impact est d'autant plus fort que la zone est ouverte sur l'extérieur, donc, d'autant plus fort, a priori, qu'elle est géographiquement de petite taille. Il est non seulement important de connaître le solde migratoire, ce qui est l'information la plus couramment utilisée, mais il est bon de connaître les composantes qui font le solde, à savoir les flux d'entrée et les flux de sortie.

Or les informations fournies par le recensement sur le sujet sont nombreuses mais leur interprétation n'est pas toujours immédiate. Si l'on connaît bien les caractérisques des personnes qui arrivent dans un lieu donné, on ne sait pas décrire les personnes au moment de leur migration ni même à la date du précédent recensement. Or l'événement migration intervient souvent en même temps qu'un changement de statut : les personnes qui étaient étudiantes deviennent actives, d'autres qui étaient en activité partent à la retraite, d'autres qui étaient agriculteurs à la campagne deviennent ouvriers en ville, etc.
Par ailleurs, la durée inégale des périodes intercensitaires (6, 7, 8 ou 9 ans pour la période 1954-99) rend difficilement comparables des flux lorsque l'on travaille sur des séries longues.

La première partie consiste à s'interroger pour savoir si les enquêtes annuelles sur l'emploi et les recensements conduisent aux mêmes conclusions quand on les applique aux échanges entre l'Ile-de-France et la province.

Le seconde partie examine le modèle migrants-migrations développé par Daniel Courgeau que l'on essaie de valider sur ces mêmes courants d'échanges. La question est de savoir comment peut se faire la comparaison temporelle des flux de migrations résidentielles lorsque les durées de référence ont des longueurs variables, ce qui est le cas pour les six périodes intercensitaires depuis 1954 (successivement 8, 6, 7, 7, 8 et 9 ans). Ainsi un apport migratoire moyen de 1000 personnes au profit d'une région donnée n'a pas la même signification en terme d'intensité migratoire si la durée de la période intercensitaire est de 6 ou de 9 ans.

Première partie

Comparaison des échanges Ile-de-France / province

Pour mesurer les flux migratoires résidentiels on dispose de deux sources d'informations en partie comparables : les recensements de la population et les enquêtes annuelles sur l'emploi.

On se propose d'examiner dans quelle mesure les deux sources disent ou non la même chose. Pour ce faire et à titre d'exemple, on traite des seuls échanges entre l'Ile-de-France et la province. On pourrait bien entendu étudier d'autres flux, pour autant cependant que les chiffres soient significatifs. En effet, les enquêtes sur l'emploi présentent l'inconvénient de ne toucher que peu de monde chaque année, environ 140000 personnes et parmi elles seulement … % changent de région. Les incertitudes aléatoires sont donc très fortes. Elle le deviennent encore davantage quand on s'intéresse à une partie du territoire et que l'on fait une décomposition par âge. Néanmoins, quand on agrège les résultats de plusieurs enquêtes successives, on est fondé à pensé que la crédibilité des résultats s'affirme.

Les limites de la comparaison

Certes, les possiblités de comparaisons sont aussi limitées par le fait que le champ couvert par l'enquête emploi est légèrement plus étroit que celui du recensement. En effet les ménages collectifs ne sont pas interrogés à l'enquête emploi, ce qui, notamment, contribue à restreindre le nombre d'arrivées de jeunes en Ile-de-France (puisque certains étudiants vont vivre en résidence universitaire).
Autre difficulté d'interprétation qui arrive parfois : l'effet-grappes qui risque de perturber la perception des mouvements migratoires. En effet dans le cas d'une délocalisation d'établissement, il se peut que la migration, certes réelle, soit exagérément amplifiée puisque plusieurs membres d'un même ménage (ou même d'un même aire) se retrouvent parfois émigrant dans la même localité d'arrivée.

Ce qui distingue les deux sources, c'est le fait que la durée sur laquelle on s'appuie pour mesurer les flux est très inégale. A l'enquête emploi on demande aux personnes interrogées d'indiquer quelle était leur résidence antérieure quatorze mois auparavant (quand l'enquête se fait au début du mois de mars), tandis qu'au recensement on demande le lieu de résidence au premier janvier de l'année du précédent recensement. Ainsi pour la période séparant les recensements de 1990 et 1999, la durée est de 9 ans et deux mois[1]. Et si l'on veut effectuer des comparaisons sur les flux migratoires portant sur de longues périodes, on ne peut éluder la question de la durée inégale des périodes intercensitaires (de 6 à 9 ans).

Dans cette étude, on convient de ne pas s'attarder trop longtemps sur les conséquences de l'arrondi. On s'intéresse essentiellement à la méthode. Il est très facile à l'ordinateur de refaire les calculs en remplaçant 9 ans par 9,18 ans ou par tout autre valeur.

Un premier exemple : la migration des personnes nées en 1969

Comparer les résultats entre les deux sources ne se fait pas immédiatement. C'est pourquoi, on commence par un exemple. Il concerne les flux résidentiels entre l'Ile-de-France et la province pour la période 1990-99 chez les personnes ayant effectué leur migration autour de 25 ans.

D'après le recensement.

La période de migration couvre les années 1990, 1991… jusqu'à 1998, si l'on fait abstraction des deux mois de janvier et de février de 1999. Ces personnes ont donc réalisé leur migration au cours des années où elles ont atteint les âges de 21, 22, … 29 ans.

En moyenne, elles ont migré à l'âge de 25 ans.

Au cours d'une aussi longue période, il est arrivé à de nombreux migrants d'effectuer un second déplacement, voire un troisième ou plusieurs autres encore. Si le migrant qui a quitté l'Ile-de-France revient dans la région-capitale, le recensement ignorera les déplacements intermédiaires.

D'après les enquêtes emploi.

Chaque enquête annuelle donne également le nombre des arrivées et des départs d'Ile-de-France. L'enquête permet donc connaître le *nombre de mouvements aux différents âges*. Ainsi à l'enquête emploi effectuée en mars de l'année 1999, les personnes nées en 1969 ont migré en moyenne à l'âge de 29 ans (elles ont 30 ans au moment de l'enquête).

Si l'on veut comparer les résultats des enquêtes emploi avec les recensements, il faut donc sommer l'ensemble des migrations des personnes nées en 1969 aux enquêtes de 1991, 1992, …1999 (21, 22, … 29 ans).

[1] *Dans cette étude, on convient de ne pas s'attarder trop longtemps sur les conséquences de l'arrondi. On s'intéresse essentiellement à la méthode. Il est très facile à l'ordinateur de refaire les calculs en remplaçant 9 ans par* 9,18 ans ou par tout autre valeur.

Les résultats comparés apparaissent dans le tableau qui suit.

Tableau 1 - **la génération née en 1969**

Sources : les enquêtes emploi et le RP99

Année de l'enquête	Année de naissance	Age à la migration	arrivées	départs	solde
1991	1969	21	7063	5309	1754
1992	1969	22	7106	6029	1077
1993	1969	23	9147	9243	-96
1994	1969	24	6616	6822	-206
1995	1969	25	7389	6382	1007
1996	1969	26	6063	6737	-674
1997	1969	27	7934	4926	3008
1998	1969	28	6882	6052	830
1999	1969	29	3819	5808	-1989
91-99 (1)	**1969**	**25**	**62019**	**57308**	**4711**
RP 99 (2)	**1969***	**25**	**38949**	**24982**	**13967**
Ratio (1/2)	.	.	*1,69*	*2,29*	.

** âge au recensement : 30 ans*

En agrégeant les 9 enquêtes annuelles sur l'emploi, il apparaît que 62019 personnes, nées en 1969, ont déclaré *arriver* en Ile-de-France. Ces personnes ont migré à des âges différents (en moyenne à 25 ans). Selon le recensement de 1999 le nombre des personnes nées en 1969 venues en Ile-de-France s'élève à 38949. Ce chiffre est tout naturellement inférieur à celui des enquêtes emploi car nombre de jeunes venus s'installer dans la capitale pour le temps de leurs études ou pour leur premier emploi sont repartis. Entre les deux chiffres, le rapport est de 1 à 1,69.

Pour les *départs* de l'Ile-de-France, également plus élevés selon les enquêtes emploi, le rapport est beaucoup plus important encore, de 1 à 2,29.

En principe, le solde migratoire ne devrait pas être grandement affecté par les différences de mesure, puisqu'il est censé témoigner du *bilan des échanges*. Il est toutefois notablement différent : un solde en faveur de l'Ile-de-France de 4711 selon les enquêtes emploi, de 13967 selon le recensement.

Rappelons que les jeunes - c'est le cas notamment pour les étudiants - qui viennent pour quelques années s'installer dans la capitale, sont souvent logés dans un ménage collectif (cité universitaire, foyer…). N'étant pas concernés par l'enquête emploi, ils ne figurent pas dans les arrivées (leur nombre n'est sans doute pas négligeable au titre des lignes 21, 22 et 23 ans du tableau) . En revanche ils figurent dans les départs, s'ils vivent après leur migration dans une résidence principale comme c'est le cas le plus probable. On peut trouver là une explication à la différence observée sur le solde migratoire, plus faible selon l'enquête emploi.

Nonobstant ce point relatif aux arrivées, les flux observés aux recensements doivent logiquement être plus faibles que ceux observés de façon cumulée aux différentes enquêtes annuelles successives. On trouvera plus loin les évolutions des flux (graphiques A et B).

Un second exemple : la migration des personnes âgées 25 ans

On constate donc que le recensement ne permet pas de mesurer l'événement migration *à un âge donné*. L'enquête annuelle sur l'emploi en revanche offre cette possiblité. On peut admettre en première approximation que l'année de la migration est décalée d'un an par rapport à l'âge observé lors de l'enquête. De ce fait les personnes âgées de 26 ans à la date d'une enquête ont fait en moyenne leur migration à 25 ans. D'où un second tableau.

Tableau 2

Les personnes ayant migré à l'âge de 25 ans

Sources : les enquêtes emploi et le RP99

Année de l'enquête	Année de naissance	Age à la migration	arrivées	départs	solde
1991	1965	25	5713	7585	-1872
1992	1966	25	4471	7023	-2552
1993	1967	25	5516	4777	739
1994	1968	25	6282	8752	-2470
1995	1969	25	7389	6382	1007
1996	1970	25	6748	7093	-345
1997	1971	25	8971	5777	2894
1998	1972	25	15204	8293	6911
1999	1973	25	13062	6364	6698
91-99	**1969**	**25**	**73056**	**62046**	**11010**
RP99	**1969**	**25**	**38949**	**24982**	**13967**

La ligne inférieure (RP99) ainsi que la ;ligne « 25 ans » sont les mêmes que dans le tableau 1

Ce second tableau illustre les variations annuelles qui apparaissent d'année en année, à un âge donné de la migration, à savoir 25 ans. Comme dans le cas précédent, il faut tenir compte des incertitudes aléatoires, considérables. Ainsi, comment savoir si l'augmentation des arrivées aux titre des deux dernières années (1998 et 1999) s'explique par une attraction croissante de la capitale ou par les simples fluctuations d'échantillon (voire par des anomalies de colecte ou des effets de grappe) .

L'agrégation des résultats sur 9 ans permet cependant d'établir un indicateur caractérisant la propension de la région à migrer (graphiques C).

Pour différentes raisons (incertitudes liées à l'échantillon, flou dans les dates de migration, etc.) on procédera par la suite à des comparaisons entre les sources faites sur des âges quinquennaux.

Les grandes lignes de la confrontation

On présente dans les pages qui suivent une série de graphiques.

A - **Comparaisons entre les sources (enquêtes emploi et recensement)**
période-1990-99 - arrivées, départs, soldes - par âge (A3M)

B- **Comparaisons entre périodes (1982-90 et 1990-99)**
enquêtes emploi et recensements - arrivées, départs - par âge (A3M).

C - **Enquête emploi (1990-99) par âge détaillé**
Arrivées - départs

A3M signifie une moyenne sur trois âges (âge de la migration) centrée sur l'année indiquée sur l'axe horizontal
Les résultats concernent les migrations sur la période (huit ans pour la période 1982-90, 9 ans pour la période 190-99).
La première valeur de la série est 24 ans (23, 24 et 25 ans). Les âges inférieurs ont été omis, d'une part parce que les enfants dans les enquêtes emploi font l'objet d'un fichier particlier (que l'on n'a pris en compte), et aussi parce que les arrivées à Paris de la population en âge d'être étudiant sont probablement sous-estimés.

Il ressort d'un examen rapide des graphiques une convergence d'ensemble entre les deux enquêtes (compte tenu du fait que les arrivées des jeunes dans la capitale sont sous-estimées chez les jeunes, on fait commencer la série à 24 ans, du moins pour les graphiques A et B).

Toutefois, certaines divergences significatives apparaissent, notamment au titre de la population jeune qui quitte la capitale. Les deux courbes sont nettement divergentes dans leur évolution (graphique A).

En revanche les deux enquêtes semblent dire à peu près la même chose quand on se réfère au seul solde migratoire. Une exception cependant pour les jeunes, du fait probablement, que ceux-ci quand ils arrivent en Ile-de-France, re rentrent pas toujours dans le champ des enquêtes emploi.

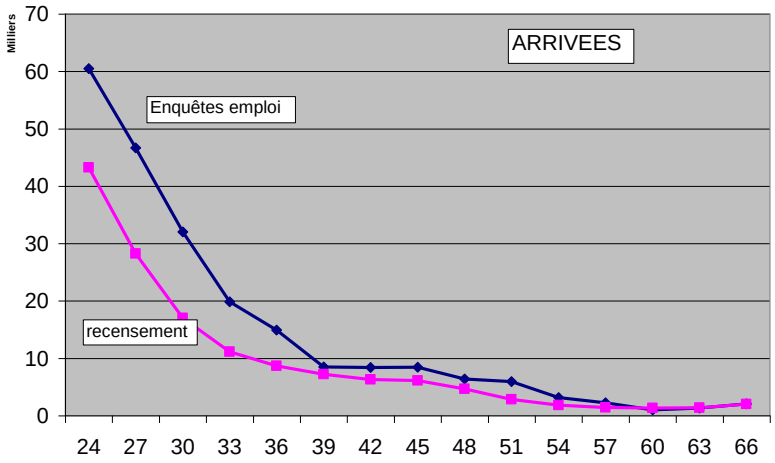
Notons que le ratio (rapport entre les flux de l'enquête emploi et du recensement), décroît nettement avec l'âge. Il est de 1,7 chez les jeunes pour les arrivées, supérieur à 2 pour les départs. Le ratio serait même parfois même inférieur à 1

après 55 ans, ce qui voudrait dire que le recensement enregistre davantage de migrations que les enquêtes emploi cumulées.

Si on s'intéresse à la question de l'évolution des migrations au cours des deux périodes considérées (graphiques B), il semble que l'on puisse conclure à une bonne concordance entre les deux sources. Selon l'une et l'autre enquête, les arrivées dans la capitale diminuent légèrement, les départs progressent fortement.

A priori, c'est le recensement qui doit être la source privilégiée, car il permet de procéder à des croisements de variables. En revanche l'enquête emploi apporte des éclairages complémentaires qu'il ne faut pas négliger. Elle permet de disposer d'indicateurs pertinents par âge détaillé et aussi de suivre les tendances s dans le temps de quelques indicateurs de base, à condition de lisser les séries annuelles, en moyenne mobile sur quelques années)

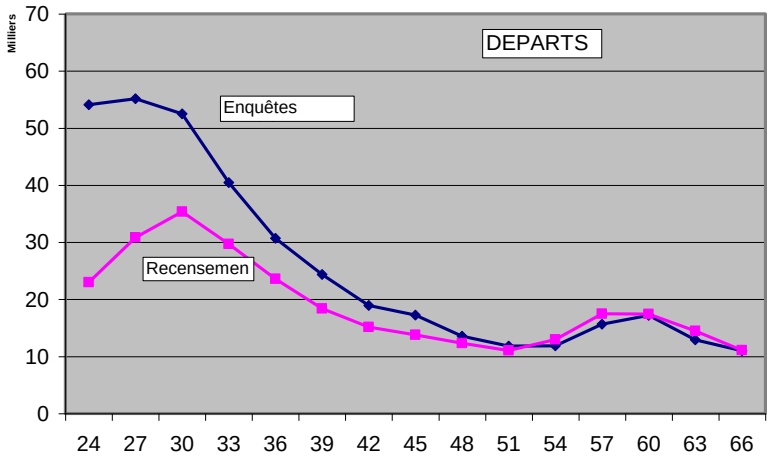
A - **Comparaisons entre les sources (enquêtes emploi et recensement)**
1990-99 - arrivées, départs, soldes - âge A3M



ARRIVEES

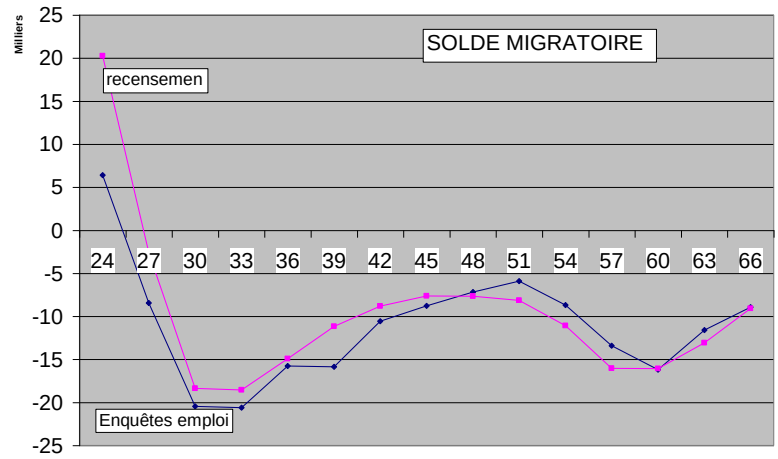
Les deux enquêtes convergent pour dire que maximum des arrivées correspond à la tranche 23-25 ans. Le chiffre de l'enquête emploi serait plus élevé si l'enquête prenait en compte les ménages collectifs (cités universitaires….

L'écart relatif entre les deux sources (rapport entre les flux) est plus fort chez les jeunes (moins de 35 ans) que chez les autres. Autour de 60 ans, il n'est que de 10% environ.



DEPARTS

Les deux sources divergent : l'enquête emploi fait apparaître un palier couvrant la tranche 23-31 ans tandis que le recensement. Marque une valeur maximale peu avant 30ans.
Chez les plus jeunes le ratio est très élevé, nettement supérieur à 2 Au-delà de 50 ans, il est voisin de 1 (il semble même que les flux cumulés des enquêtes emploi soient inférieurs aux flux du recensement.

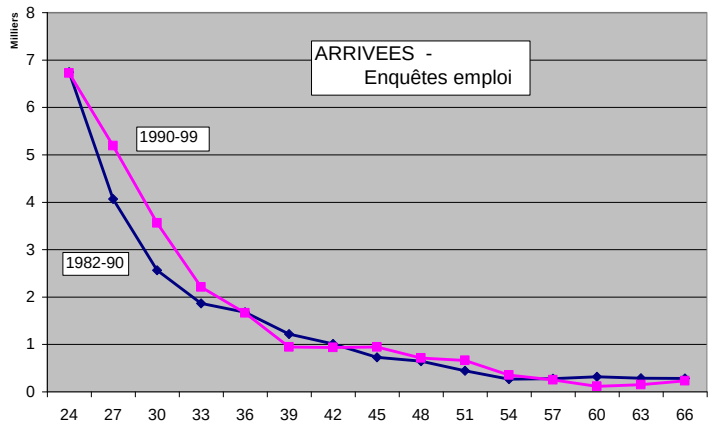


SOLDE MIGRATOIRE

En principe, les modalités de collecte ne devraient pas créer d'écarts significatifs sur les soldes (contrairement à ce qui se passe sur les flux), sauf chez les plus jeunes (étudiants…).
De fait à l'exception des moins de 30 ans, les deux courbes sont assez voisines. A noter cependant que les deux courbes se croisent autour de 45 ans : les soldes, négatifs, sont00 plus prononcé au recensement.

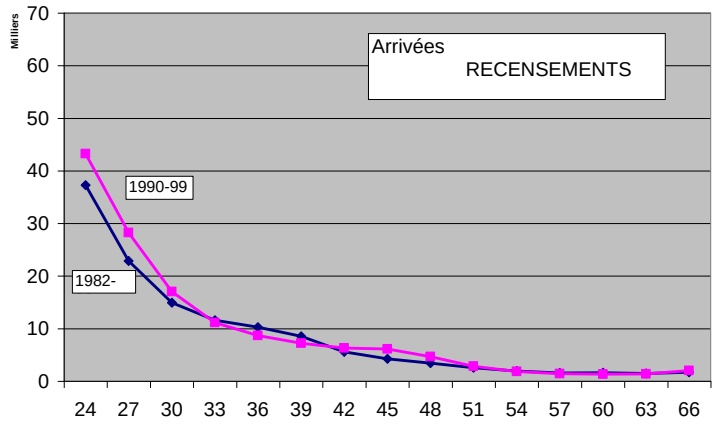
B - comparaisons entre périodes (1982-90 et 1990-99)

enquêtes emploi et recensements - arrivées, départs - âge A 3M



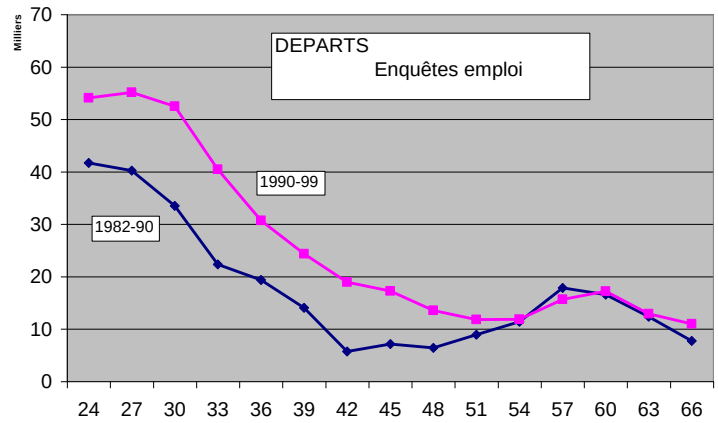
ARRIVEES selon l'Enquête emploi

Ce qui frappe surtout, c'est l'augmentation des arrivées qui touche les personnes de 25 à 34 ans et, plus modestement, les personnes de 42 à 54 ans.



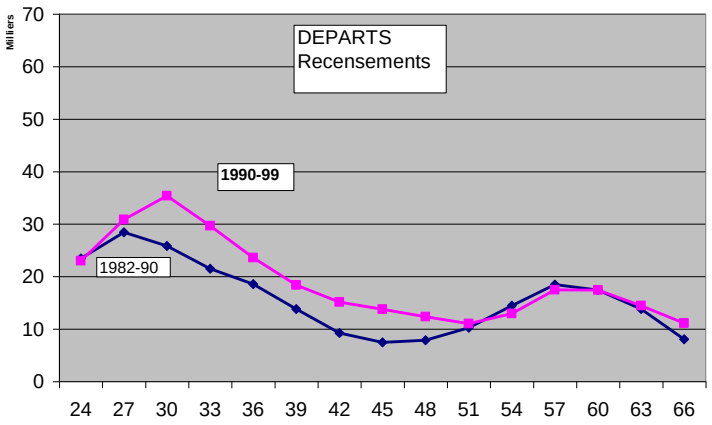
ARRIVEES - Recensements

L'attraction parisienne se serait également légèrement accentuée, mais les deux courbes sont plus proches l'une de l'autre que ne le sont les courbes des enquêtes sur l'emploi.
A noter que les deux enquêtes disent bien la même chose quand il s'agit d'un léger affaissement de l'attraction chez les 30-35 ans et à l'inverse une légère accentuation entre 40 et 50 ans.



DEPARTS - Enquête emploi

Le verdict de l'enquête emploi est très net : la propension à quitter l'Ile-de-France a très fortement augmenté : les départs sont nettement plus nombreux à tous les âges jusqu'à 55 ans. En revanche, il y a quasiment stabilité..

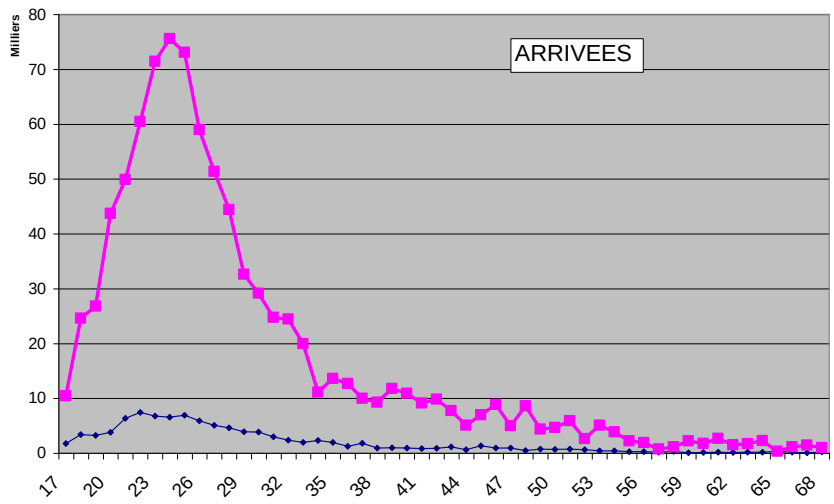


DEPARTS selon le recensement

Selon les recensements, la propension au départ des plus jeunes (23 à 15 ans) n'aurait pas changé, ce qui diffère des enquêtes sur l'emploi.
En revanche les départs augmentent sensiblement de 25 à 50 ans, pour être quasiment stables aux âges supérieurs. Les deux sources convergent bien pour affirmer la reprise des départs après 50 ans (et autour 66 ans, les deux sources disent l'inverse.

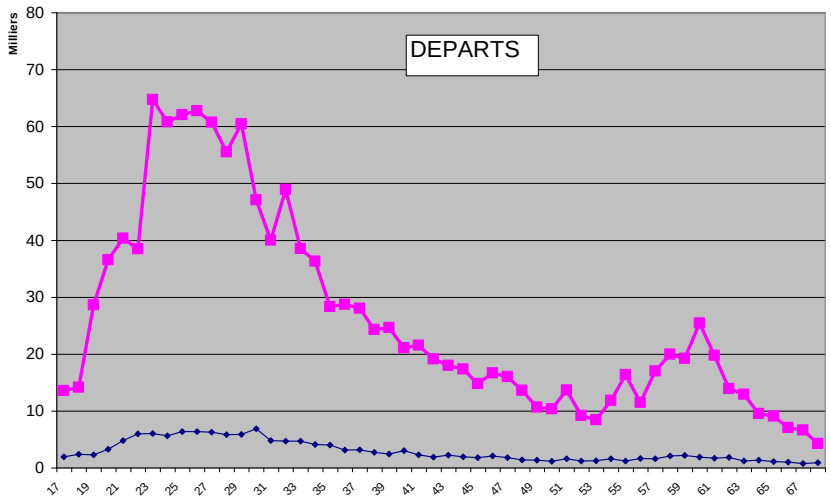
C - **Enquête emploi (1990-99) par âge détaillé**

Arrivées - départs



ARRIVEES en Ile-de-France

Il s'agit des arrivées à un âge donné pour la période 1990-99 (cumul/moyenne de 9 observations). La courbe est remarquable de régularité . Le flux est maximal à l'âge de 24 ans, On voit nettement que la plage des âges est très restreinte.



DEPARTS d'Ile-de-France

La courbe des départs est beaucoup moins régulière. Par ailleurs, il n'y a pas de valeur modale précise. On observe seulement un palier qui touche sept valeurs, correspondant à la tranche 24-30 ans.
Ensuite, malgré une rupture vers 35 ans, la baisse apparaît assez régulière.
La propension aux départs reprend à 54 ans avant d'atteindre un maximum à 61 ans.

Références

Les résultats chiffrés présentés dans cette première partie proviennent de deux fichiers historiques disponibles à l'Insee sur le site du CNIO d'Orléans.

- un fichier historique des recensements pour la période 1962-99 (Saphir)
  SP72.l00.SAPHIR.RP6299 / AA
  60 860 360 enregistrements, 37 variables

- un fichier historique des enquêtes annuelles sur l'emploi (1982-2001)
  SP72.L00.ENQEMP (2 842 018 / 34 variables)

Les variables (une sélection de variables proches ou complémentaires de celles du recensement pour les enquêtes emploi) ont été harmonisées de façon à ce que la confrontation des fichiers soit quasi-immédiate.

Ces deux fichiers, surtout le fichier historique des recensements, sont volumineux. Mais quand on s'intéresse aux migrations, on est amené à faire des extractions qui permettent de n'avoir à traiter que des fichiers de petite taille, immédiatement utilisables puisque les variables ont été harmonisées.

En prolongement du travail que l'on a effectué et qui ne porte que sur les échanges entre l'Ile-de-France et le reste de l'hexagone, on aurait pu sortir des tableaux croisant les catégories de population suivante. La confrontation des sources pourrait être effectuée poules variables suivantes :

Région
Diplôme
Sexe et âge
Pour l'emploi : statut et catégorie socio-professionnelle

En introduisant les données de l'enquête emploi, on « complète » en quelque sorte l'information du recensement. On peut ainsi étudier la corrélation entre migration et changement de statut des personnes (ce que ne permet pas le recensement, puisqu'on ne demande pas les caractéristiques de statut et d'emploi des personnes à la date du précédent recensement.

Où l'on retrouve l'objectif initial : comment faire parler simultanément plusieurs sources sans avoir au préalable à effectuer des opérations lourdes, fastidieuses et donc coûteuses.

Deuxième partie

Les conséquences sur les flux migratoires de la durée inégale des périodes intercensitaires

Examen du modèle migrants-migrations

Le problème abordé ans cette note est celui de la comparaison temporelle des flux de migrations résidentielles, lorsque les durées de référence ont des longueurs variables, ce qui est le cas pour les périodes intercensitaires depuis 1954 (successivement 8, 6, 7, 7, 8 et 9 ans). Ainsi un apport migratoire moyen de 1000 personnes au profit d'une région A en provenance d'une région B n'a pas la même signification en terme d'intensité migratoire si la durée de la période intercensitaire est de 6 ou 9 ans.

Comment faire ?

Le problème n'est pas nouveau, mais il n'a été semble-t-il relativement peu traité jusqu'à présent, du moins à l'Insee. Le numéro 624 d'octobre 2001 de la collection RAR : « *Le calcul du taux de mobilité à l'aide des données censitaires : le modèle migrants-migrations* » Franck L'Hospital) avait précisément pour but de proposer plusieurs formules permettant de ramener les flux observés sur des longues périodes à des flux annuels. Les formules sont empruntées à Daniel Courgeau. Les hypothèses qui ont conduit à leur expression sont simples. L'idée première est qu'une fraction de la population qui vient s'établir dans une zone géographique repart, soit dans sa zone d'origine, soit ailleurs en métropole. Le nombre des départs est proportionnel à la population soumise au risque, ce qui conduit à un processus exponentiel. Au moment où s'achève la période intercensitaire, une partie seulement de cette population

soumise au risque est repartie. Par conséquent, plus la durée de la période intercensitaire est longue et plus grande est la proportion de ceux qui repartent.

Le RAR qui présente cette formule donne des paramètres pour les deux seules périodes intercensitaires, 1982-1990 et 1990 -1999.

Or la mesure des migrations est indispensable si l'on veut percevoir correctement la dynamique des territoires, et celle-ci ne prend tout son sens que si l'on dispose de séries temporelles longues. C'est pourquoi, on aurait aimé que les paramètres calculés pour la période 1982-99 puissent être également disponibles pour les quatre périodes intercensitaires précédentes (de 1954 à 1982). On pourrait ainsi, grâce au fichier Saphir, disposer de séries couvrant 45 ans.

C'est que la durée des périodes intercensitaires est inégale. Elle s'échelonne de 6 à 9 ans. Aussi, diviser les flux par le nombre d'années séparant deux recensements successifs, ce que l'on fait souvent, conduit à la production d'indicateurs biaisés.

Mais au-delà du calcul des coefficients de redressement, il faut s'interroger plus généralement sur l'usage approprié que l'on peut faire de tels indicateurs, non seulement quand on suit les migrations dans le temps , mais même quand on se limite à l'examen d'une seule période intercensitaire.

la formule

Le RAR propose quatre formules de base, d'allure assez voisine. On s'en tient à l'une d'entre elles, celle qui concerne les arrivées dans une zone.. L'écriture de la formule est la suivante, pour une unité de population :

$$I = it - K(i + jl)(t - \frac{1 - e^{-kt}}{k}) \qquad (1)$$

avec les coefficients suivants, valables pour une région et pour la période 1990-1999 :

$K = 0{,}448$ (fraction de la population effectuant une nouvelle migration)
$k = 0{,}262$ (probabilité instantanée de migrer, pour cette fraction)
$l = 0{,}351$ (proportion de migrants effectuant un retour)
t : nombre d'années

I et j sont les coefficients instantanés de migration, pour les immigrants (I) et les émigrants (J) observés au titre de la période intercensitaire.

Le taux instantané d'immigration s'écrit : $\lim_{\Delta t \to 0} \frac{\Delta I}{\Delta t}$.

Plutôt que de s'intéresser à l'expression de *I*, on s'intéresse à l'écriture d'un coefficient correcteur $\rho(t)$. Dans le cas où l'émigration est de même importance que l'immigration (c'est souvent vrai en première approximation pour les régions), on écrit une formule qui s'est affranchie de *j*.

$$\rho(t) = \frac{I}{it} = \left[1 - K(1+l)(1 - \frac{1-e^{-kt}}{kt} \right]$$

Pour mieux en saisir le contenu, on se propose de poser :

$$A = e^{-kt} \qquad B = 1 - \frac{1-e^{-kt}}{kt}$$

Le tableau donne les valeurs du coefficient de correction, ainsi que de ses deux composantes A et B dans quatre cas de figure correspondant à durées de référence et pour deux hypothèses sur la proportion de migrants effectuant un retour *: l = 0,351* et *l = 0*.

Nombre d'années t	$A = e^{ikt}$	$B = 1 - \frac{1-e^{-kt}}{kt}$	$\rho = \frac{1}{it} = 1 - K(1+l)(1 - \frac{1-e^{-kt}}{kt})$		
			l=0,351		
			Base ann. 0	*Base ann. 1*	*l = 0*
(a)	*(b)*	*(c)*	*(d)*	*(e)*	*(f)*
0	0	0	1	-	1
1	0,770	0,120	0,927	1	0,946
5	0,270	0,443	0,732	0,790	0,830
9	0,095	0,616	0,627	0,676	0,724

Rappel : on considère j = i (le taux d'arrivées est égal au taux de départs)

La colonne e diffère de la ligne d par le fait que l'on change l'année de base pour le calcul du coefficient de correction.

Application numérique à pour une période intercensitaire de 9 ans

Si l'on prend pour taux de retour la valeur *l=0,351*, le coefficient de correction est égal à 0,627 (col d), du moins si l'on prend comme base le taux instantané (t=0) .

Si on fait l'hypothèse qu'il n'y a pas de retour (l=0) le coefficient est alors plus petit (0,724 , col f).

La colonne e correspond au cas où l'on prend comme référence non pas le taux instantané (qui ne parle guère, puisque c'est une abstraction), mais le taux ramené à 1 an (celui que l'on mesure d'après l'enquête emploi).

exemple chiffré

° *Soit un flux d'immigration de 9000 personnes pour une période intercensitaire de 9 ans. Si comme on le fait souvent on ramène ce taux à un an, cela fait une migration moyenne de 1000 personnes.*
Si en revanche on applique le coefficient du modèle migrant-migrations (0,627) dont l'inverse est 1,595) on est amené à fixer à 1595 le flux instantané. Autrement dit, pour une période de 9 ans, aux 9000 immigrants correspond en fait 15950 immigrations.

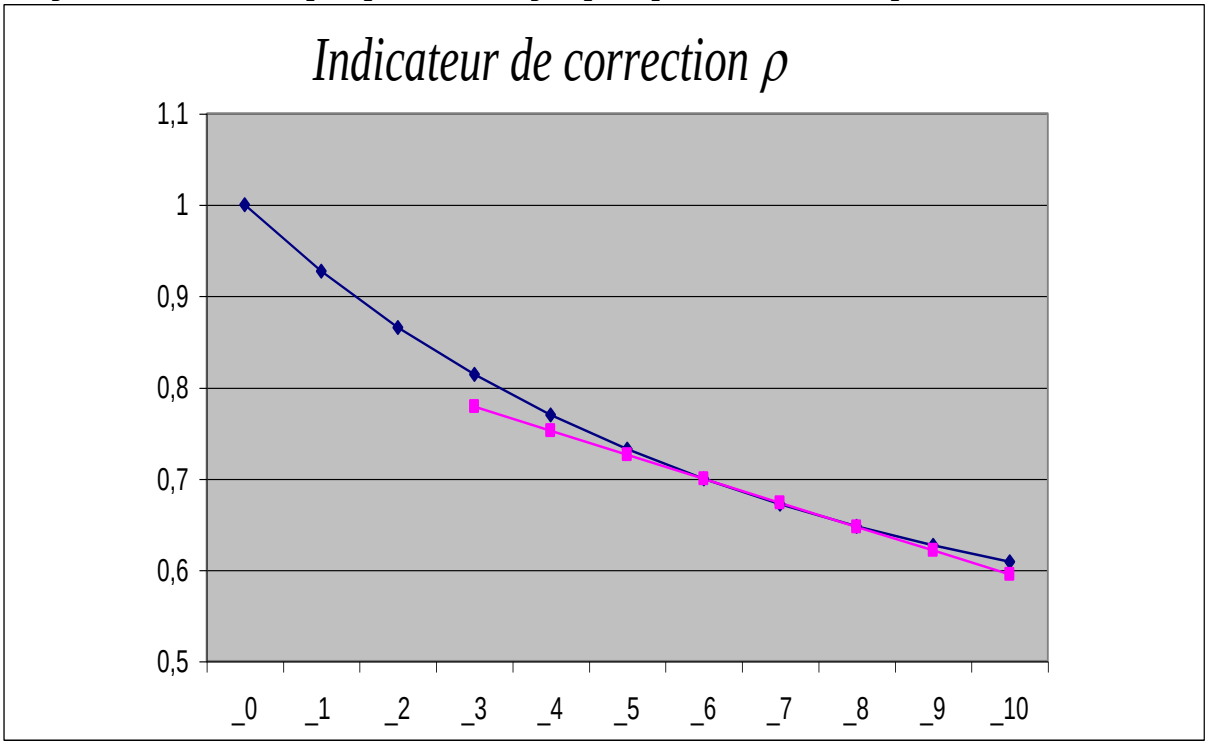
En fait, la notion de taux instantané d'immigration est une abstraction. Pour rendre plus concret, donc plus compréhensible la correction, on peut préférer prendre comme référence le nombre d'immigrants pour une période de 1 an. En ce cas les résultats doivent être corrigés en s'appuyant sur une série qui prend la valeur 1 pour 1 an (col e). Ce ne sont pas 9000 immigrants, mais 14800 immigrants.

C'est sans doute ce résultat qu'il faut retenir :

Un flux de 1000 personnes au recensement de 1990-99 équivaut en intensité migratoire

- ***à un flux de 1500 migrations s'il est comparé à un flux de 1 an***
- ***à un flux de 1170 personnes s'il est comparé à un flux de 5 ans.***

Un graphique
Pour visualiser les effets du modèle-migrants - migrations, on propose le graphique suivant qui donne la



Bien entendu les calculs peuvent être refaits pour des valeurs plus précises, par exemple 9,18 années pour la dernière période intercensitaire et 1,2 années pour l'enquête emploi. Mais ce point est secondaire dans la démonstration.

Comment appliquer le modèle

L'acceptation ou la remise en cause de la formule du modèle relève d'une démarche scientifique. On pose quelques hypothèse et l'on en déduit une expression méthématique. Il en est un peu de la formule du modèle migrants-migrations comme des modèles économiques (par exemple l'oscillateur de Samuelson). Ce sont des outils utiles à la compréhension des phénomènes, mais leur aptitude à rendre compte de la réalité est loin de s'imposer. En revanche, ce sont des instruments pédagogiques et ils aident bien à mettre en garde contre la tendance à ignorer ou à éluder certains problèmes.
Pour revenir au problème migratoire, il ne s'agit ni de condamner la méthode un peu simpliste qui consiste à diviser un flux migratoire par le nombre des années de la période intercensitaire, ni bien évidemment de condamner le modèle migrants-migrations, il s'agit de rendre sensible les utilisateurs sur un problème difficle qui n'a donc pas de réponse simple.

C'est précisément l'objet de cette investigation que de s'efforcer de « médiatiser » des formules mathématiques dont il faut bien reconnaître qu'elle ont la propriété d'effaroucher nombre de ceux qui pourraient en tirer profit dans leurs travaux.

L'application à l'analyse locale

Si l'on admet le bien-fondé de la question, on est appelé à chercher des réponses appropriées. En dehors bien sûr des travaux de spécialistes, de chercheurs qui pourront donner des résultats sur des domaines spécifiques au terme d'analyses fouillées (réponse à des question du type : qu'en est-il de la propension moyenne de la population à changer de commune, de département, de région ?), c'est aux applications locales - qui exigent forcément des simplifications conceptuelles - que l'on pense.

Car précisément, et bien que le niveau de formation moyen de nos interlocuteurs qui travaillent sur le terrain soit d'année en année plus élevé, il reste que la plupart d'entre eux sont allergiques à l'emploi des formules mathématique, nous nous devons de diffuser des

données claires, accessibles au plus grand nombre.

Or parmi ces informations, les indicateurs de flux sont utiles à l'élu et à ses collaborateurs qui préparent des dossiers. Pour comprendre le fonctionnement d'un territoire, la perspective historique s'impose. En l'occurrence il s'agit de savoir si l'attaction (resp. la répulsion) de la zone d'étude s'accentue ou non. Cette considération vaut aussi bien quand la migration concerne des petits territoires que de plus grands (migrations interrégionales, Paris-province, intérieur littoral, centre périphérie des villes, etc.).

Le nouveau recensement : des périodes de référence de 5 ans

Que faire en particulier lorsque seront diffusées les flux migratoires issus du nouveau recensement ? Il faudra bien comparer des flux sur *cinq* ans avec des flux sur *neuf* ans.

Deux façons de faire extrêmes sont possibles :

° soit appliquer les formules du modèle migrant-migrations à tout résultat en sortie de l'ordinateur (voire même programmer les formules dans les phases de traitement, quitte à diffuser deux chiffres, un chiffre brut et un chiffre corrigé) . Il faudrait alors pour cela que soient validés les paramètres. Cela impliquerait aussi que l'on ait la certitude que les paramètres ne sont pas trop différents en fonction des cas de figure[2] (dans quelle mesure les paramètres sont sensibles aux structures des zones d'origine et d'accueil structures par sexe âge, diplôme…).
Cette solution ne serait réaliste que si la communauté des statisticiens intégre bien le fonctionnement du modèle et qu'elle soit en mesure de l'expliquer clairement sur le terrain. Il faudrait donc au préalable que soient réalisés de nombreux travaux de validation. C'est pourquoi on donne en annexe une bibliographie (fournie par Daniel COURGEAU) qui permettra peut-être de susciter des vocations.

° soit se résigner à une procédure « simpliste » : diviser le nombre de migrants par le nombre d'années de

[2] *Or, si l'on en juge d'après les résultats obtenus pour les échanges Paris-province, il semble qe l'effet âge soit déterminant : des ratios parfois supérieurs à 2 chez les jeunes, voisns de 1 chez les plus de 60 ans. Ces écarts se comprenent : chez les jeunes, les multi-migrations sont très nombreuses. S'agissant des échanges Paris-province, il ne saurait y avoir d'autre migrations supplémentaires que celles de retour. Cela n'est pas le cas lorsquel'on considére des échanges entre deux villes (ou régions) : pour sa deuxième migration la personne peut certes revenir dans sa ville (ou région), soit se rendre dans une ville (ou région) tierce. Il faut donc examiner la question de la sensibilité des indicateurs à ces problèmes de taille et de nature des territoires d'origine et d'accueil.*

la période de référence. En ce cas il serait très osuhaitable d’assortir la diffusion des données par quelques indications fournis par le modèle migrants-migrations.

Une suggestion

Bien que l'on ne soit pas dans le même cas de figure, on pourrait par analogie s'inspirer de ce qui a été fait au moment de la diffusion des résultats des recensements issus des sondages. Les « fascicules verts » donnent une formule approchée qui offre la possibilité de savoir approximativement quelle est l'incertitude aléatoire des résultats diffusés. Cela permet aux utilisateurs à l'esprit rigoureux de rappeler en marge des documents qu'ils rédigent les limites des résultats, notamment quand ils traitent à des niveaux géographiques ou structurels fins.

On pourrait donc suggérer d'associer à la diffusion des données un texte simple qui fournirait les ordres de grandeur des erreurs systématiques qui se créent lorsque l'on fait cette correction « brutale ». Ces ordres de grandeur pourraient même être accompagnées le cas échéant d'une présentation des hypothèses du modèle ainsi que d'une abaque.

Cette façon de procéder permettrait de fixer une « ligne de conduite » à suivre pour la comparaison des flux d'une période à l'autre (notamment pour comparer les flux du nouveau recensement à ceux de la période intercensitaire précédente).

En s'appuyant sur le modèle migrants-migrants (dont les grandes lignes seraient reprises dans le commentaires), on aurait donc un commentaire de ce type (texte extrait d'un message que m'a envoyé Daniel COURGEAU :

On constate d'abord que le nombre de migrations annuelles, pendant une période de quelques années, est approximativement constant : il doit donc être possible d'estimer un quotient instantané de migration valable pendant cette période. Cette hypothèse n'est d'ailleurs pas indispensable et on peut envisager une variation linéaire de ce quotient au cours du temps. Observant maintenant à l'aide d'enquêtes longitudinales les migrations de divers rangs, on constate que parmi les individus ayant fait une migration de rang n, seule une partie d'entre eux fera une migration de rang n+1, K, proportion qui dépend peu du rang de la migration. Ensuite, en travaillant sur cette part de population, on vérifie que la probabilité de faire une nouvelle migration, k, dépend peu de la durée de séjour, du rang de la migration et du découpage géographique sur lequel la migration est mesurée. Enfin, la proportion de retours parmi les individus effectuant une migration multiple, l, en constitue une proportion qui est indépendante de la durée de séjour.

A partir de toutes ces vérifications, on peut alors construire un modèle mathématique donnant l'effectif de « migrants » observé en fonction de l'effectif de « migrations » et des paramètres K, k, l, et de la durée de séjour. Nous renvoyons le lecteur intéressé

aux articles et travaux qui les présentent de façon plus détaillée (Courgeau, 1973, 1979 a, 1979 b, 1982, 1983, 1986 ; Ledent, 1982, Courgeau et Lefebvre, 1982).

Texte auquel on ajouterait quelques indications chiffrées :

> *Si l'on veut comparer les flux observés au nouveau recensement (qui portent sur 5 ans) avec ceux de la période intercensitaire précédente (qui portent sur 9 ans), les flux migratoires doivent être divisés par 1,167.*
>
> *Par conséquent un flux de 1000 personnes observé entre 2000 et 2005 est équivalent, en première approximation, à un flux de 857 observé entre 1990 et 1999.*

Annexe méthodologique

Modèle migrants-migrations : démonstration partielle de la formule

On ne cherche pas à faire une démonstration rigoureuse des formules du modèle migrants-migrations (on pourra se référer à la bibliographie présentée à la fin du document). On se propose seulement de poser quelques hypothèses et de montrer comment on introduit, logiquement, l'expression exponentielle qui contribue à donner une apparence de complexité aux formules telles qu'elles sont publiées dans le RAR, formules que nous avons prises comme référence pour l'écriture de ce chapitre.

La formule à démontrer relative à l'**immigration** est, dans sa forme élémentaire la suivante :

$$I = iT - Ki(T - \frac{1 - e^{-kT}}{k})$$

Cette expression peut être vue comme la valeur d'une intégrale correspondant à une fonction d'immigration $f(t)$.

On écrit donc d'une façon générale : $I = \int_0^T f(t)dt$

I représente le nombre d'immigrants dans une zone au cours d'une période intercensitaire, i le taux instantané d'immigration (supposé constant au cours de la période de durée T). Les données se rapportent à une unité de population.
Par rapport à la formule de référence on fait deux hypothèses, à savoir : $j = i$ (le taux instantané d'émigration est égal au taux instantané d'immigration), et $l=0$ (la proportion de retours est nulle).

Cas particulier :

Le taux d'immigration i est constant et le paramètre K est nul (personne parmi les immigrants n'effectue une seconde migration). En ce cas l'intégration est simple car $f(t)$ est constant et l'on écrit : $I = iT$.

Ce sont les effets des migrations multiples qui compliquent les choses et obligent à retours et la rédaction de formules plus compliquées.

On pose trois hypothèses :

Première hypothèse : on admet que le paramètre *i* est constant au cours de la période.

Deuxième hypothèse : parmi les personnes qui arrivent dans la zone de référence (au cours d'un instant infiniment petit *dt*, il arrive *idt* personnes), une fraction (notée $iKdt$) supposée constante) effectue une nouvelle migration.

Troisième hypothèse : la loi de probabilité correspondant aux départs (cf hypothèse précédente) s'écrit $p(t)=ke^{-kt}$ où *k* est un paramètre. Cela signifie que plus on reste longtemp dans dans la zone d'accueil, plus faible est la probabilité de la quitter.

Conséquence de la troisième hypothèse :

On cherche à déterminer la proportion de la population qui repart avant la date du recensement qui suit celle de l'arrivée dsans la zone de référence.

On note N= K idt le nombre de personnes qui ont immigré au cours d'un instant dt très court et qui sont soumis au risque d'une nouvelle migration.
A l'infini toutes personnes qui doivent quitter la zone sont parties. Mais ce qui nous intéresse c'est la proportion de ceux qui sont repartis avant la date du recensement qui suit la date de l'installation.
Pour connaître cette probabilité d'être reparti, il suffit d'intégrer la loi de probabilité de la valeur 0 à la valeur t, de sorte que l'on écrit :

$$P(t)=\int_0^t e^{-k\tau}dt$$

La primitive de la fonction est $\left[-e^{-k\tau}\right]$ *et la solution, compte tenu des bornes est la suivante.* $P(t)=1-e^{-kt}$
A l'infini, on écrit, comme le paramètre k est positif , $P=1$ *: CQFD : tous ceux qui devaient*

partir ont quitté la zone.

Remarquons d'abord que la formule de la loi de probabilité qui régit les départs $p(t)=ke^{-kt}$ *n'est autre elle-même que la solution d'une équation différentielle élémentaire du type :* $dN=-kN(t)$. *Cette relation signifie que le nombre de départs pour un temps donné est constant en valeur relative, mais diminue en valeur absolue, au fur et à mesure que s'éteint la population susceptible de partir*

Sur la base de ces hypothèses, la fonction d'immigration est la suivante :

$$f(t) = i(1 - K(1 - e^{-kt})$$

L'intégration de cette fonction ne pose pas de problème. On constate qu'elle est la somme de deux éléments. Le premier n'est autre que la constante i (voir plus haut), l'autre étant une fonction exponentielle liée à l'effet des retours.

Pour couvrir tout la période intercensitaire de durée T :

$$I(t) = i\int_0^T \left[1 - K(1 - e^{-kt}\right]dt \qquad = I = iT - Ki(T - \frac{1 - e^{-kT}}{k})$$

linéarisation de la formule

On simplifie beaucoup l'écriture en remplaçant l'espression obtenue par une forme linéaire. Deux situations sont envisagées

* Dans le premier cas, si la durée de la période intercensitaire est fixée, le temps devient un paramètre et, moyennant quelques développements de la formule, on parvient à une forme linéarisée.

* Dans le second cas, on considère des durées de référence assez voisines (le terme « voisin » étant très relatif). On constate alors que l'expression mathématique se linéarise au prix d'une très légère erreur.

Cette simplification se justifie d'autant plus que l'on ne saurait connaître avec une grande précision des informations qui font appel à la mémoire des personnes dont on sait qu'elle n'est pas toujours fiable surtout quand il s'agit de périodes aussi longues

Bien noter que se référer à la propriété linéaire de la fonction n'enlève en rien à l'intérêt qu'il y a donner la formule compliquée qui seule permet d'appréhender le fonctionnement du système migratoire.

Dans les deux cas de figure, on pose $e = (1+\varepsilon)i$. Quand les taux d'émigration et d'immigration sont égaux, ce qui est le cas assez fréquemment, du moins en première approximation, on a alors $e = i$. Dans les autres cas, le terme ε est l'écart relatif qui dsépare le taux d'émigration du taux d'immigration.

A - La durée de la période intercensitaire est fixée (T constant).

On fixe T, durée de la période intercensitaire. Ainsi pour les travaux portant sur la seule période 1990-990, on a T=9,18

le RAR propose une formule simplifiée, à savoir :

$$i = \frac{I-(I-E)X_1}{Y_1}$$

où X_1 et Y_1 résultent de la transformation de la fonction (en prenant en compte également la formule relative àç l'émigration) sont respectivement 0,098 et 5,727).

L'examen de cette formule montre qu'effectivement, comme le paramètre X_1 est petit (environ 0,1) et qu'il s'applique au solde migratoire, l'effet correctif est relativement faible quand E est voisin de I.

Aussi transforme-t-on cette formule[3] et l'on écrit :

$$I = \frac{Y_1}{1+\varepsilon X_1} i \approx (1-\varepsilon X_1) Y_i i\,, \qquad (8)$$

La formule est linéarisée : I et i sont dans un rapport constant. On retrouve en effet la formule étudiée : $\rho = \frac{I}{iT} = \frac{Y_1}{iT} = \frac{5{,}727}{9{,}18} = 0{,}636$ (voir gr 1)

Application numérique : le tableau donne les valeurs du coefficient correctif.

Nombre d'années T	formule de base (2)	formule linéarisée (3)
9,18	5,725	5,741
8,18	5,268	5,265
7,18	4,792	4,789
6,18	4,292	4,313

La formule simplifiée pour l'émigration est analogue. Cela permet de conclure que le solde migratoire n'est pas affecté par la durée de la période intercensitaire : le solde est divisible par le nombre d'années de la période intercensitaire.

[3] On sait en effet que $\frac{1}{1-\alpha} \approx 1+\alpha$

b- Les périodes sont de durée assez voisine

L'examen de la courbe figurant les effets du modèle (cf graphique I) fait ressortir une évolution quasi-linéaire dès lors que l'on considère des durées voisines. Même quand on considère des durées s'étendant de 5 à 9 ans (5 ans est la durée de référence pour le nouveau recensement, 9 ans est la plus longue des durées intercensitaires pour la période 1954-62), la linéarisation est possible sans risque de biais important.

De fait si on superpose une droite d'équation

$$\rho(t) = -0{,}262t + 0{,}857$$

on constate que les deux courbes sont très proches l'une de l'autre sur le domaine considéré.

Pour autant que la linéarisation de la fonction exponentielle se justifie , on peut généraliser la formule au cas où les taux d'émigration ne sontr pas identiques aux taux d'immigration. En ce cas, on pourraut écrire :

:

$\rho = t - cy(t) - c'\,y(t)$, expression où $t - cf\,(t)$ représente la contribution de la correction dans le cas particulier où immigrants et émigrants sont en nombre égal. $c'\,f\,(t)$ est un terme complémentaire.

Dans cette expression, $c = \frac{K}{k}(1+l)$ et $c = \frac{K}{k}l$

Références bibliographiques communiquées par Daniel Courgeau

Baccaïni B., Courgeau D. et Desplanques G. (1993) – Les migrations internes en France de 1982 à 1990 : Comparaison avec les périodes antérieures, *Population*, 48 (6), pp. 1771-1790.

Baccaïni B. (1999) - Analyse des migrations internes et estimation du solde migratoire externe au niveau local à l'aide des données censitaires . *Population*, n°4-5, pp. 801-815.

Calot G. (1975) – Note méthodologique sur les migrations, in *Migrations intérieures. Méthodes d'observation et d'analyse*, CNRS, Paris, 47-71.

CNIS (2002) – *Rapport de la commission spéciale mise en place en application de l'article 158 de la loi n° 2002-276 relative à la démocratie de proximité,* n° 77.

Collomb P. (1984) – *La mort de l'orme séculaire. Crise agricole et migrations dans l'Ouest audois au cours des années cinquante*, Cahier Travaux et Documents, n° 105, Ined / Puf, 466 p.

Courgeau D. (1973 a) - Migrants et migrations, *Population*, 28 (1), pp. 95-129.

Courgeau D. (1973 b) – Migrations et découpage du territoire, *Population*, 28 (3), pp.511-538.

Courgeau D. (1979) – Migrants and migrations, *Population, Selected papers*, 35 p.

Courgeau D. (1979) – Migration and demographic phenomena in France, in *The urban impact of internal migration*, J. W. White ed., Institute for research in social sciences, The University of North Carolina at Chapel Hill, USA, pp. 1-32.

Courgeau D. (1980) – *Analyse quantitative des migrations humaines*, Masson, Paris, 226 p.

Courgeau D. (1982) – Methods of linking migration statistics collected from national survey with those of population censuses, in *National Migration Surveys, X, Guide-lines for analysis, United Nations*, ESCAP, pp. 312-330.

Courgeau D. (1983) – Measuring flows and stocks of internal migrants: Selected statistical issues, in *Actes de la 44é Session de l'Institut International de Statistique, Madrid, vol. 2*, pp. 1208-1224. (repris en 1993, in *Readings in population research methodology*, vol. 4, Bogue, Arriaga, Anderson eds, United Fund for Population Activities, pp. 50-65).

Courgeau D. (1986) – *Utilisation des données de l'enquête sur l'emploi sur les migrations annuelles en France*, **Rapport à la Délégation à l'Aménagement du Territoire et à l'Action Régionale**, 91 p.

Courgeau D. (1988) – *Méthodes de mesure de la mobilité spatiale. Migrations internes, mobilité temporaire, navettes*, Editions de l'Ined, Paris, 304 p.

Courgeau D. et Lefebvre M. (1982) – Les migrations internes en France de 1954 à 1975. II.- Migrations et urbanisation, *Population*, 37 (2), pp. 341-370.

Courgeau D. (2001) – Mobilité et hétérogénéité spatiale, in *Démographie : analyse et synthèse. I La dynamique des populations*, Caselli G., Vallin J. et Wunsch G. ed, Editions de l'Ined, pp. 479-502.

Courgeau D. et Lelièvre E. (2003) – Motifs individuels et sociaux des migrations, in *Démographie : Analyse et synthèse. Vol ; IV Les facteurs du mouvement : les déterminants de la migration*, .Caselli G., Vallin J. et Wunsch G. ed, Editions de l'Ined, à paraître.

Ledent J. (with the collaboration of Courgeau D.) (1982) – *Migration and settlement : 15. France*, International Institute for Applied system analysis, Laxenburg, Austria, 200 p.

Le Jeannic T. (1997) - Trente ans de périurbanisation : extension et dilution des villes, *Economie et Statistiques*, 307, pp. 21-42.

L'Hospital F. (2001) – *Les migrations internes en France. Estimation des paramètres du modèle "migrants - migrations" de Daniel Courgeau*, Service des Etudes Diffusion - Lyon, INSEE Rhône Alpes, 77 p.

Long L. (1976) – The geographical mobility of Americans. An international comparison, *Current Population Report*, Series P-23, n° 64, 46 p.

Long L. (1988) - *Migration and residential mobility in the United States*, Russell Sage Foundation, New York.

Nations Unies (1971) – *Méthodes de mesure de la migration interne*, Etudes démographiques, n° 47, 85 p.

Nations Unies (1980) – *Principes et recommandations concernant les recensements de la population et de l'habitation*, Etudes Statistiques, Série M, n° 67, 374 p.

Nicholson B. (1990) – The hidden component in census-derived migration data: assessing its size and distribution, *Demography*, 27 (1), 111-119.

Poulain M. (1985) – La migration, concept et méthodes de mesure, in *Migrations internes. Collecte des données et méthodes d'analyse*, Chaire Quetelet '83, Jezierski ed., Louvain-la-Neuve, pp. 7-38.

Rees P. and Kupiszewski M. (1999) – *Internal migration and regional population dynamics in Europe: a synthesis*, Population Studies, 32,113 p.

Tugault Y. (1973) – La mesure de la mobilité. Cinq études sur les migrations internes, Cahier Travaux et Documents, n° 67, INED / PUF, 226 p.

Willekens F. (1985) – Comparability of migration data. Utopia or reality?, in *Migrations internes. Collecte des données et méthodes d'analyse*, Chaire Quetelet '83, Jezierski ed., Louvain-la-Neuve, pp. 409-441.