

## Bulletin de l'Association des démographes du Québec



# Les migrations interrégionales au Québec : utilisation des données de la RAMQ, 1971-72

Louise Tremblay

Volume 3, numéro 1, hors-série, 1974

Année mondiale de la population

URI : <https://id.erudit.org/iderudit/305776ar>

DOI : <https://doi.org/10.7202/305776ar>

[Aller au sommaire du numéro](#)

### Éditeur(s)

Association des démographes du Québec

### ISSN

0380-1713 (imprimé)

1925-3478 (numérique)

[Découvrir la revue](#)

### Citer cet article

Tremblay, L. (1974). Les migrations interrégionales au Québec : utilisation des données de la RAMQ, 1971-72. *Bulletin de l'Association des démographes du Québec*, 3(1), 126–140. <https://doi.org/10.7202/305776ar>

## LES MIGRATIONS INTERRÉGIONALES AU QUÉBEC : UTILISATION DES DONNÉES DE LA RAMQ, 1971-72

Avant de vous indiquer le plan détaillé du sujet que nous voulons traiter aujourd'hui, j'aimerais vous signaler quelques faits. Le travail que nous effectuons présentement se situe dans le cadre du module démographie du programme MEDICS du Ministère des Affaires Sociales. Dans cette étude, nous voulons déterminer les facteurs explicatifs les plus importants de la migration interne récente au Québec, ceci dans le but de nous fournir des éléments utiles pour la prévision démographique des régions socio-sanitaires du Québec. Comme les données disponibles dans ce domaine sont limitées, nous avons pensé pouvoir exploiter le fichier de la Régie de l'Assurance Maladie du Québec. Les fichiers administratifs, cependant, ne sont pas exempts d'erreurs, celui-ci en particulier. Nous avons donc consacré une bonne partie de notre travail à la critique des données; les deuxième et troisième parties étant réservées à la construction d'une échelle de préférence régionale et à l'établissement d'un modèle explicatif utilisant la technique d'analyse de régression, nous espérons ainsi parvenir au but que nous nous sommes fixé. Nous allons maintenant exposer le contenu de chacune de ces trois sections.

## 1. LES DONNEES

### 1.1 Considérations préliminaires

Comme nous le savons, la Régie d'Assurance Maladie a été mise sur pied à la fin de 1970. Les données disponibles concernent de ce fait la période commençant en janvier 1971 et se terminant en octobre 1972. C'est à partir des réclamations des médecins que nous avons établi un processus nous permettant d'estimer l'importance des flux migratoires d'une région i vers une région j. Par conséquent, les flux que nous obtenons sont fonctions de la population qui "consomme" l'Assurance Maladie, i.e. celle qui reçoit des soins. Nous ne touchons donc qu'une partie de la population totale, échantillon que nous supposons représentatif de la population migrante du Québec. Vu les données utilisées, il nous faut prendre en considération les différentes possibilités d'erreurs qui peuvent être introduites. Elles se situent principalement à 3 niveaux: soit au niveau de la population "bénéficiaire", i.e. celle qui a une carte, soit de la population "consommatrice", i.e. celle qui reçoit des soins et utilise sa carte, ou encore du traitement que l'on fait subir aux données pour obtenir l'information voulue. Nous allons étudier une à une ces trois sources d'erreurs.

### 1.2 Le fichier des bénéficiaires

Comme nous l'avons déjà signalé, ce n'est pas toute la population du Québec qui peut bénéficier des services offerts par l'Assurance Maladie. Quelles sont les personnes concernées? Tout ceci doit être bien clair afin de ne pas interpréter faussement les résultats auxquels nous parviendrons. Si l'on se fie aux estimations de la RAMQ, la population inscrite au Régime d'Assurance Maladie du Québec compte à peu près 90% de la population totale du Québec. Or selon les chiffres

qu'on obtient en 1973 (6.5 millions pour le Québec), on constate une surestimation, non uniforme selon les groupes d'âges. Les erreurs se retrouvent surtout chez les jeunes enfants, à cause des considérations d'ordre technique entourant leur inscription au Régime et chez les vieillards à cause des décès non déclarés.

### 1.3 Le fichier des consommateurs

On a estimé aux environs de 80% la population du Québec qui consomme l'Assurance Maladie. On voit ici l'importance des personnes que nous touchons en utilisant les données de ce fichier et le fait que nous n'obtiendrons pas des flux réels mais dont l'ordre de grandeur devrait être assez près de la réalité d'où l'hypothèse que nos résultats pourront être significatifs. Cependant, nous avons pu constater par ailleurs qu'il existe des différences de consommation selon le sexe, le groupe d'âge et la région. Il nous faudra donc corriger nos flux afin de les rendre comparables. Ceci pourra être fait grâce aux informations qui nous parviendront de l'échantillon (11,200) de la RAMQ qu'exploite SMA, ceux-ci nous fournissant une estimation du pourcentage de consommateurs selon les différentes caractéristiques (sexe, âge, région). Dans ce fichier aussi, on surévalue le nombre de consommateurs et nombreuses sont les causes d'erreurs, particulièrement aux âges jeunes et élevés et chez les femmes. A quoi sont-elles dues? Principalement, il s'agit d'erreurs liées à la reconstitution de l'identification de l'individu (s'il n'a pas sa carte: cas d'un prénom composé), à l'utilisation indifférente de l'une ou l'autre carte, si on est en possession de deux cartes (cas de la femme qui se marie, divorce, devient veuve ou se remarie, carte de la mère ou de l'enfant). Pour les bébés naissants, on retrouve les mêmes problèmes que précédemment et les vieillards de plus de 100 ans retombent en enfance (101 ans = 01 an, 102 ans = 02 ans, etc.).

Suite à ces observations sur les fichiers des bénéficiaires et consommateurs et après comparaison avec les chiffres du recensement, nous avons décidé de limiter notre étude aux personnes de sexe masculin âgées de 15 à 64 ans: d'abord parce que, ceux-ci ne présentent pas d'erreurs types du genre déjà décrit ci-dessus et que dans ce cas, l'hypothèse d'erreurs aléatoires nous semble plus probable et nous permettra d'utiliser ces données de façon un peu plus fiable. Quoiqu'il en soit le comportement de la population active (pour nous 15-64 ans au lieu de 14 - 65 ans et plus) masculine devrait refléter assez bien le comportement migratoire global de la population du Québec et les résultats provenant de ce groupe seront, sans aucun doute, significatifs.

#### 1.4 Explication et critique du traitement des données

Rappelons que c'est à partir des réclamations que les médecins doivent soumettre à la Régie pour fins de remboursement que nous avons établi un programme nous permettant d'assigner une région aux individus et ensuite de voir si ils ont fait une migration à un moment donné, i.e. s'il y a eu changement de résidence au cours de la période.

L'établissement de ces flux se fait principalement en deux grandes étapes que nous décrirons ci-après mais auparavant nous devons nous interroger sur les gens qui ont droit aux services offerts par la RAMQ, i.e. les bénéficiaires qui par la suite pourront être des consommateurs. Cependant, pour des raisons pratiques, nous expliquerons d'abord notre démarche pour calculer nos flux, ceci nous permettant ensuite de réfuter les objections pouvant être soulevées relativement aux bénéficiaires de l'Assurance Maladie.

#### 1.4.1 Le traitement des données

Signalons d'abord certains faits. Vu les déficiences du système d'enregistrement au début (nouvelle organisation) et à la fin (retards) de la période considérée, nous nous servirons de toutes les informations pour le processus d'attribution des régions pour ne conserver dans une deuxième étape que les renseignements s'étalant de juillet 1971 à juillet 1972. Nous travaillerons d'abord sur la population masculine âgée de 15 à 64 ans, inclus, par groupe quinquennal. Si le temps nous le permet, nous verrons ce qui se passe pour la population féminine bien que nous doutions de la valeur des résultats que nous obtiendrions à cause des nombreuses erreurs dans les données. Nous avons d'abord pensé à travailler par région socio-sanitaire mais notre hypothèse de base (région de résidence = région de traitement) ne nous semblant pas valable pour les trois régions de Montréal, nous en sommes réduits à travailler par région administrative. Nous voulions aussi étudier le phénomène par mois, mais l'existence d'un certain problème de décalage, nous a amenés à ne conserver que les déplacements sur l'année (faisant ainsi l'hypothèse d'un seul changement de résidence par année).

##### 1.4.1.1 Première étape

Cette étape consiste à attribuer une région à un individu, région que nous qualifierons de région brute.

L'hypothèse de base qui guide toute notre étude est la suivante: l'individu qui reçoit des soins les reçoit dans sa région de résidence.

Nous ne considérons cependant que les soins offerts par les généralistes et spécialistes de lère ligne (pédiatre, gynécologue, psychiatre, interniste) leur attachant respectivement une pondération de 9 et de 1 (la plus grande pondération pouvant être attribuée à un généraliste avec une seule colonne d'ordinateur: question de coût, ne négligeant pas cependant ce type de spécialistes).

Considérons un exemple fictif: 1 individu x reçoit des soins:

1er mois:	aucun $\Rightarrow$ Région (R) = 0	R	Poids
2ième mois:	3 fois chez un spécialiste dans région	2	3
	1 fois chez un généraliste dans région	3	9
	2 fois chez un généraliste dans région	1	<u>18</u>

on lui attribue la région 1 qui a le plus grand poids. etc.

Nous aurons donc à la fin de cette première étape, les résultats suivants: en autant qu'un individu est identifié, à chaque mois où cela est possible, i.e. s'il reçoit des soins, nous lui aurons assigné une région "brute" et les résultats se présenteront sous cette forme.

No du mois	1	2	3	. . . 18	19	20	21
------------	---	---	---	----------	----	----	----

01-1971 à 10-1972

No de la région brute (R.B.)	0	1	0	2	2	2	0
---------------------------------	---	---	---	---	---	---	---

1.4.1.2 Deuxième étape

Il s'agit cette fois de ne pas se limiter à l'information du moment. Nous voulons accorder une certaine importance, pour localiser un individu un mois donné, aux régions brutes déterminées antérieurement et pour quelques mois dans le futur, afin de ne pas négliger l'effet de déplacements temporaires dans une région autre que la région de résidence. On a conservé 2 mois futurs.

Exemple fictif:

Mois	1	2	3	4	5	6	7	8	9	10	11	12	.....
R.B.	1	0	2	0	0	2	1	0	1	4	1	0	.....
Bernouille	1	0	1	0	0	1	1	0	1	1	1	0	.....
P.G. ....	$.8^8$	$.8^7$	$.8^6$	$.8^5$	$.8^4$	$.8^3$	$.8^2$	$.8$	1	$.8^{-1}$	$.8^{-2}$		
...										1.25	1.64		
	du 4 <sup>ième</sup> mois à l'infini												
	histoire passée									futur: 3 derniers mois			
R.N.									1	1	1	1	



Dans cet exemple, nous voulons assigner une région à un individu en tenant compte de son histoire passée et future. Nous qualifierons cette région de région nette R.N. Pour ce faire, nous appliquerons à une Bernouilli (X) définie comme suit:

X = 0 échec: nous n'avons pas attribué de R.B. à notre individu, un mois donné.

= 1 succès: l'individu a pu être localisé,

une progression géométrique de raison "r" définie de telle sorte que le poids des 3 derniers mois ne soit pas plus fort que l'histoire passée. Empiriquement, r varie entre .7 et .9. Nous avons conservé .8.

La région nette est alors définie comme la région différente de 0 qui obtient le maximum de la somme. Dans notre exemple,

$$\begin{array}{r}
 R1 \\
 2 \\
 4
 \end{array}
 \begin{array}{l}
 \Sigma = 3.37 \\
 \Sigma = .77 \\
 \Sigma = 1.25
 \end{array}
 \Rightarrow R.N. = 1$$

on fait de même à chaque mois. On compare avec R.B. et si on observe plus de non-correspondances que de correspondances, on rejette notre individu car il est alors trop mobile (en termes de soins). Les étapes suivantes: on remplace les 0 par la région déterminée la plus près dans le temps et on décale de deux mois.

A la fin de cette deuxième étape, nous aurons assigné une région nette à chaque mois, à chaque individu, en autant qu'il est identifié, région que nous lui attribuerons comme lieu de résidence.

Notre hypothèse de base (région de résidence = région de traitement) peut être mise en doute: il existe toute une littérature sur ce sujet portant sur différents pays. Nous avons de bonnes raisons de croire, cependant, que dans notre cas elle est justifiée. Pourquoi? Premièrement, nous ne considérons que les soins offerts par généralistes et spécialistes de lère ligne attachant une plus grande importance aux premiers. Deuxièmement, les régions administratives sont très grandes: des doutes peuvent subsister aux limites des régions. Ces problèmes sont particulièrement importants dans les régions où il existe peu de médecins et où il faut bien voir qu'une variation du nombre de médecins au cours de la période peut entraîner un certain nombre de faux-migrants. Nous examinerons donc le problème sous deux aspects. Du côté de l'offre, en voyant si il y a eu modification importante du nombre de médecins ou d'hôpitaux et leur localisation mais aussi du côté de la demande en trouvant où se situent les centres de gravité de la population, si ils sont près des frontières et si il existe des hôpitaux près des frontières des régions avoisinantes. Nous considérerons aussi la situation en examinant des zones isochrones de 60 minutes centrées sur des hôpitaux.

Nous avons ensuite regroupé pour l'année en faisant l'hypothèse d'une seule migration par année (impliquant un changement de résidence) et nous obtenons ainsi des matrices de flux, par sexe, par groupes d'âges, entre les dix régions socio-sanitaires (Montréal regroupé).

#### 1.4.2 Les bénéficiaires de l'Assurance Maladie

Signalons rapidement le fait que nous voulons étudier la migration interne donc nous ne nous intéressons pas aux immigrants et émigrants "externes", i.e. provenant ou se dirigeant hors du Québec. Le problème de ces émigrants est vite résolu (autre province, autre pays, étudiants ou fonctionnaires à l'étranger, décès): ils sont considérés soit comme non-migrants ou simplement non-inclus dans nos flux. Quant aux immigrants, la solution implique l'hypothèse que les gens qui entrent au Québec de façon permanente ont ou auront le même comportement face à la migration interne que les autres résidents du Québec.

#### 1.5 Conclusion: section données

Nous sommes maintenant conscients de toutes les implications liées à l'utilisation de ces données et nous pourrions en tenir compte lors de l'interprétation.

## 2. L'ECHELLE DE PREFERENCE REGIONALE

Nous voulons déterminer dans cette partie de notre étude, quelles sont les régions préférées des émigrants d'une région donnée et aussi celles des migrants en général, i.e. en considérant le phénomène de façon globale. Nous ferons ce travail pour chaque groupe d'âge et au total de façon à pouvoir identifier les différences qui pourraient exister selon ces groupes d'âges. Signalons

cependant la nécessité d'une étape antérieure consistant à rendre nos flux comparables, ceci, en utilisant les estimations que nous fournirait l'échantillon.

Quant au choix de la méthode, parmi les différentes méthodes possibles, nous avons retenu celles de BACHI et de TERMOTE qui consistent à comparer les flux observés à des flux "attendus". Ces derniers sont calculés à partir d'hypothèses théoriques. Pour BACHI, il nous suffit de supposer indépendance statistique dans le tableau de contingence nous donnant le nombre de migrants selon les lieux de départ et d'arrivée, signifiant ainsi la complète indépendance entre lieux d'origine et de destination.

$$\widehat{M}_{iK} = \frac{EM_i \times IM_K}{\sum_i \sum_{\substack{j \\ i \neq j}} M_{ij}}$$

$$\text{et l'indice} = 100 \frac{M_{iK} - \widehat{M}_{iK}}{EM_i - \widehat{M}_{iK}}$$

si le numérateur  $> 0 \Rightarrow$  échelle de préférence  
 $< 0 \Rightarrow$  échelle d'aversion

dans la direction i-K .

Quant à la méthode de TERMOTE, on calcule la propension nationale (i.e. pour le Québec) à immigrer en K,  $p_K$ ,

$$p_K = \frac{IM_K}{\sum_i \sum_{\substack{j \\ i \neq j}} M_{ij} - EM_K}$$

et on applique cette probabilité aux émigrants de  $i$  pour calculer un nombre théorique attendu de migrants de  $i$  vers  $K$  ( $\widehat{M}_{iK} = p_K \times EM_i$ ); l'indice utilisé est alors  $\frac{M_{iK}}{\widehat{M}_{iK}}$  qui, plus grand que 1, nous permet de construire une échelle

de préférence.

Pour toute région  $K$  de destination, on peut ensuite calculer un indice de préférence globale pour  $K$  ( $I_{pg.K}$ )

$$I_{pg.K} = \frac{\sum_i (M_{iK} - \widehat{M}_{iK})}{\sum_i \sum_{j \neq K} M_{ij} - \sum_i M_{iK}}$$

et construire ainsi une hiérarchie.

Il est clair que, se basant sur l'une ou l'autre de ces hypothèses, nous mesurons les choses différentes mais le travail étant effectué par ordinateur et par conséquent relativement court, nous pourrions ainsi constater les différences impliquées par le choix de l'hypothèse.

### 3. LE MODELE DE REGRESSION

#### 3.1 Le choix des variables

Se basant sur différentes études et sur la disponibilité des statistiques du recensement de 1971 (qui a tout de même nécessité une demande spéciale), nous avons retenu les variables suivantes par sexe, par groupe d'âge et par région:

1. Taux de chômage
2. Revenus moyens
3. % de l'emploi agricole (re au total de l'emploi dans la région)
4. % de l'emploi tertiaire (re au total de la région)
5. Education: nombre moyen d'années de scolarité
6. % de gens nés à l'extérieur de la région
7. % de propriétaires
8. Emploi par secteur d'activité économique
9. Distance entre les régions (calculée par carte routière soit entre les pôles d'attraction économique ou entre les plus grandes villes)
10. Existence d'une ville de plus de 50,000 habitants.

Nous pourrions également considérer les différences de certaines de ces caractéristiques entre les régions d'origine et de destination et l'accroissement dans le temps. Signalons ici une petite difficulté liée au type de région utilisée: pour 1966 ou 1961, il nous est impossible d'obtenir des données par région socio-sanitaire. Nous avons pensé à contourner le problème d'une certaine façon, si nous obtenons les renseignements suffisants, ces variables étant supposées très importantes dans notre modèle.

### 3.2 La méthode

Nous utiliserons le "stepwise" qui consiste à sélectionner par étape les variables que nous fournisse le  $R^2$  le plus grand après le choix de chaque variable et qui se situe finalement le plus près possible du maximum réel. Nous ferons également une régression qu'on pourrait qualifier de traditionnelle, i.e. en choisissant nous mêmes les variables que nous retiendrons et non pas en les déterminant par un processus mécanique. Il ne faut pas oublier que nous pondérerons nos flux par la population bénéficiaire des régions enlevant ainsi en partie l'effet de la taille de la population des régions qui joue sur l'ordre de grandeur des flux.

Avec les taux par âge (revenu, chômage, etc.) que nous aurons pu obtenir de Statistique Canada, nous ferons alors la régression pour le total des groupes d'âges et pour chacun d'eux. Nous pourrions alors peut être constater des différences significatives selon l'âge, ceci nous permettant d'obtenir des résultats beaucoup plus intéressants.

### Conclusion

Ayant achevé ces trois grandes parties de notre travail, il nous restera à faire tous les liens possibles entre chacune d'elles afin de pouvoir conclure de façon satisfaisante et obtenir une interprétation fiable des résultats provenant de données d'un type assez particulier et rarement utilisées dans ce genre d'études.

Nous aurions aimé pouvoir présenter quelques résultats, au moins sur la partie concernant les échelles de préférence régionale. Malheureusement, les données de l'échantillon de la RAMQ ne nous étant pas encore parvenues, cela n'a pas été possible. Quant au modèle de régression, les variables impliquées exigeant une demande spéciale du recensement de 1971, celles-ci ne seront pas disponibles avant quelque temps.

Louise TREMBLAY  
Mai 1974