

Chapitre 3

L'explication du niveau de consommation

Élaboration d'un modèle théorique

« Tout code de comportement fondé sur des lois biologiques se doit de reconnaître que le travail est une nécessité fondamentale de l'être humain, et tout particulièrement le travail dont on peut récolter les fruits. »

HANS SELYE

« Le travail constitue le lien essentiel entre l'individu et la société. »

HENRI LABORIT

L'EXPLICATION DES NIVEAUX RÉGIONAUX DE CONSOMMATION DE SERVICES : ÉLABORATION D'UN MODÈLE THÉORIQUE

Le chapitre traite des phénomènes qui peuvent, *en théorie et logiquement*, expliquer le niveau de consommation des populations. On y distingue deux parties. Dans la première, à l'aide des connaissances provenant de disciplines pertinentes, il s'agit de *comprendre* – par opposition à expliquer – *la nature de la relation déjà vérifiée entre le niveau de non-emploi des populations et l'incidence de la morbidité caractérisant ces populations*; morbidité reflétée, à titre d'hypothèse, par leur niveau de consommation respective.

La seconde partie, débutant sous la rubrique « Un second déterminant... », identifie et analyse les éléments permettant de comprendre et de situer, les uns par rapport aux autres, des phénomènes de divers ordres indissociables du niveau de consommation de services, phénomènes néanmoins essentiels à la compréhension d'ensemble qui lie cette consommation à d'éventuels facteurs causaux. Cette partie traitera notamment :

- ◆ de l'existence d'une *autre variable sociale* nécessaire pour rendre compte du niveau de consommation des populations. Cette variable est elle-même, en bonne partie, expliquée par le niveau de non-emploi des populations.
- ◆ de l'influence qu'exercent les fonctions du système sanitaire, notamment les décisions médicales d'hospitaliser, la fréquence de réhospitalisation des personnes et les durées de séjour.
- ◆ des facteurs, autres que sociaux, qui contribuent à expliquer les écarts interrégionaux de consommation de services, soit les pratiques professionnelles, l'organisation des services et la « performance » des établissements.

Cette seconde partie arrive, par procédé déductif, à isoler les concepts aptes à mesurer et à expliquer entièrement le niveau de consommation de services des populations régionales.

UNE PRÉSUMPTION DE LIEN CAUSAL

Le discours scientifique, comme celui de la pratique, regorge d'énoncés clamant l'existence d'un lien évident entre la pauvreté et la maladie. Il est néanmoins difficile de trouver une seule vérification systématique et causale sur la base de données primitives, complètes, diachroniques et généralisables à des populations entières, en appui à ces énoncés couramment admis. Curieusement d'ailleurs, ce défaut de vérification n'est pas imputable à l'inexistence de telles données.

À cet égard, le chapitre 1 nous laissait sur des présomptions surprenantes. On note d'abord une concomitance très élevée entre deux variables apparemment fort éloignées l'une de l'autre : le non-emploi des populations et l'incidence des jours d'hospitalisation consommés en santé physique. On relève, par ailleurs, un ensemble de manifestations permettant de supposer l'existence d'un lien causal entre ces variables. Il ne peut, à l'évidence, s'agir d'un lien direct mais plutôt, le cas échéant, d'une chaîne causale liant un phénomène social – mesuré par le taux d'inoccupation – à l'expression de la maladie mesurée à l'échelle des populations, telle qu'elle peut s'observer à travers la fenêtre des hospitalisations.

Les manifestations qui fondent l'intérêt de l'hypothèse sont les suivantes :

- ◆ l'accroissement des corrélations à mesure qu'on s'éloigne de l'année de prise du Tl : l'indicateur se comporte comme un déterminant, au plan statistique;
- ◆ l'indicateur rend compte du niveau relatif de la consommation régionale de chacun des grands groupes d'âge, les 0-15 ans, les 15-64 ans, les 65 ans et plus; de même, pour tous les groupes d'âge quinquennaux, des 0-4 ans jusqu'aux 75 ans et plus;
- ◆ pour un ensemble de maladies représentant 85 % de tous les jours d'hospitalisation, le Tl rend compte des niveaux régionaux de consommation de chacun des groupes diagnostiques.

Ce dernier constat laisse voir qu'une population régionale présentant un fort manque d'emplois cumulera, pour la plupart des maladies considérées indépendamment les unes des autres, un taux concomitant et élevé de recours aux services. Comme si une population plus affectée l'était pour l'ensemble des maladies, en même temps. Or, une manifestation aussi structurée à l'échelle de grandes populations n'est pas perceptible à l'échelle des individus, ni même à celle de petites populations. En effet,

on n'observe pas, dans la réalité, d'individus pour la quasi-totalité des maladies organiques au cours d'une année.

Ces constats ouvrent sur deux questions distinctes mais interreliées : Est-il possible qu'il faille voir comme deux notions distinctes « l'état de santé des individus » et « l'état de santé des populations »? Secondement : Comment les problèmes de santé à l'échelle des populations peuvent-ils varier ensemble, d'un seul tenant, sans qu'il soit nécessaire d'envisager qu'un déterminant, qu'une cause commune, les fassent réagir tous en même temps et dans la même « direction »?

LE CONCEPT « D'ÉTAT DE SANTÉ DES POPULATIONS »

Les faits nouveaux et structurés constatés précédemment interrogent les notions généralement véhiculées sur l'état de santé des populations. On se rend d'abord compte que la notion, généralement véhiculée par le concept de « morbidité » notamment, renvoie implicitement soit à la sommation des caractéristiques de santé des individus ou à la simple transposition, à l'échelle des populations, de la notion d'état de santé des individus. Pourtant, contrairement à ce qui se vérifie à l'échelle individuelle, on constate que certaines populations *cumulent en même temps un ensemble cohérent de maladies* qui, à l'échelle individuelle, disposeraient illico de la vie de personnes affectées de tous ces maux. Il est donc possible que puisse être mesurée, à l'échelle d'une population, une expression de la maladie qui soit particulière aux populations et que celle-ci soit différente de la simple projection, sur de grands ensembles des états de santé individuels et singuliers.

En poussant plus avant cette réflexion, on remarque qu'indépendamment du facteur numérique qui les différencie par nature, les caractéristiques propres aux populations sont, à l'évidence, très différentes de celles des individus qui les composent, et cela, spécialement en regard de la santé. En effet, contrairement aux individus, les populations n'ont pas d'organes : elles ne peuvent donc être affligées de maladies organiques. De plus, contrairement aux individus, les populations survivent au décès de chacun des individus qui en font partie à un moment donné car elles ont, comme les espèces et les lignées, la capacité de se reproduire pour se maintenir dans le temps, comme c'est le cas pour les cellules du corps humain, à une moindre échelle. Ces évidences mettent en lumière le fait que les populations et les personnes qu'elles regroupent constituent des entités, des objets fondamentalement distincts par nature, tel que le propose la définition du mot « social ». Cela suggère donc le questionnement suivant : de quoi parle-t-on au juste lorsqu'on utilise l'expression

« état de santé des populations »? D'un phénomène social précis et mesurable ou de la simple juxtaposition d'une pluralité de phénomènes biologiques, organiques et individuels?

COMMENT POSER LE PROBLÈME DU LIEN ENTRE LE NON-EMPLOI ET LA CONSOMMATION DE SERVICES?

Le caractère systématique et l'ampleur des constats réalisés dans le temps long à l'échelle des populations régionales invitent à une double démarche. Il faut :

- ◆ d'une part, s'expliquer comment de grands nombres d'individus biologiquement distincts – mais réunis au sein d'une même collectivité – peuvent, de façon concomitante et simultanée, voir leur santé affectée en regard de la plupart des problèmes organiques, ce qui suppose minimalement qu'un facteur ou qu'un tronc commun de facteurs déterminants agissent comme déclencheurs de la maladie chez les personnes;
- ◆ d'autre part, voir comment ce déclencheur ou ce tronc commun de déclencheurs individuels peut être activé par un processus social fortement associé à la plus ou moins grande carence d'emplois prévalant dans les populations régionales.

Ne disposant pas, a priori, des compétences biomédicales permettant d'élucider le premier terme de la démarche, ont été programmés et réalisés, en 1996 et 1997, une quinzaine de séminaires scientifiques mettant à contribution des praticiens et des spécialistes¹⁹ de disciplines telles la génétique, la biologie, la psychiatrie, la médecine de santé publique, la psychologie et l'écologie. La question générale qui se posait consistait à identifier, le cas échéant, le déterminant qui permette de résoudre le paradoxe apparent qui existe entre l'expression individuelle de la maladie et le caractère tout à fait différent et systématique de son expression à l'échelle des populations régionales.

De façon plus spécifique, la question première qui se posait alors consistait à identifier, à l'échelle des phénomènes biologiques, un *déterminant commun* à l'ensemble des maladies organiques dont souffrent les personnes. L'existence d'un tel déterminant étant vue comme une condition préalable pour comprendre pourquoi certaines populations manifestent une propension à produire, en grande quantité, une diversité de

¹⁹ On trouvera, en annexe 4, une liste de personnes rencontrées en expertise et des groupes consultés en validation.

maladies organiques, comparativement à d'autres populations. En second lieu, serait posée la question relative à la nature de la relation existant entre cette ou ces causes éventuelles et le taux d'inoccupation.

LES DÉTERMINANTS BIOLOGIQUES DES MALADIES ORGANIQUES : L'ÉCLAIRAGE DE LA THÉORIE DES MALADIES MULTIFACTORIELLES

La théorie des maladies multifactorielles

À l'évidence, les maladies organiques dont souffrent les personnes sont d'abord déterminées par des facteurs biologiques – endogènes aux individus – et sur lesquels ces derniers n'ont pas d'emprise.

À ce sujet, la théorie des maladies multifactorielles²⁰ soutenue par les généticiens²¹ propose une explication au fait que des groupes d'individus exposés à des contraintes similaires verront émerger, en leur sein, une recrudescence anormale de problèmes organiques, tant au plan de l'incidence que de la diversité. D'après ce point de vue, il existe un processus causal et commun à la grande majorité des maladies organiques dont souffrent les personnes.

Selon cette théorie, et de façon évidente, le premier déterminant des maladies organiques est biologique. Il s'inscrit dans le patrimoine génétique dont chaque individu hérite de ses parents lors de la conception. Les « forces et les faiblesses » ou les prédispositions qui y sont inscrites agissent comme des conditions nécessaires *mais non suffisantes* à l'émergence des maladies organiques. C'est pourquoi on dit ces maladies *multifactorielles*, car il faut au moins deux facteurs pour les déclencher, dont l'un est nécessairement d'origine biologique.

Dans les grandes populations – celles des régions, par exemple – la répartition géographique des individus porteurs de ces prédispositions spécifiques est globalement considérée aléatoire²². Par exemple, les phénomènes de consanguinité, qui peuvent éventuellement favoriser l'émergence de maladies spécifiques dans certaines populations restreintes, sont noyés dans la masse à l'échelle de la morbidité générale des grandes populations régionales.

Les « faiblesses » inscrites dans le patrimoine génétique, dont chaque personne est porteuse, *prédisposent* certaines d'entre elles à diverses

²⁰ Voir, à ce sujet, Porter (1986, pages 1427-1472).

²¹ Parmi les généticiens rencontrés, le D^r Marc DeBraekeleer a le premier amené le rapport entre les faits présentés au chapitre 2 et l'application de la théorie des maladies multifactorielles.

²² Voir, à ce sujet, Bouchard (1994, pages 39-40).

maladies spécifiques, tout en épargnant d'autres personnes appartenant aux mêmes populations. Il faut cependant noter que les personnes prédisposées et les autres sont globalement soumises à des déterminants sociaux similaires, selon le territoire où elles habitent.

Les déclencheurs des maladies multifactorielles

D'après les spécialistes consultés sur le sujet, il existe, dans les limites des connaissances actuelles, trois déclencheurs principaux ayant la propriété d'activer les « faiblesses » chez les personnes prédisposées en raison de leur patrimoine génétique. L'exposition *chronique*, voire cumulative, des personnes à des facteurs de stress, de malnutrition – nutrition inadéquate ou insuffisante – et à des conditions environnementales défavorables – insalubrité et pollution – est reconnue comme une condition suffisante pour provoquer l'émergence de maladies organiques chez l'individu déjà prédisposé biologiquement à celles-ci.

Or, selon certains spécialistes rencontrés, les types de maladies organiques, dont la distribution régionale varie en relation avec le taux d'inoccupation, comptent parmi les maladies organiques susceptibles d'être activées par les trois déclencheurs énumérés.

Envisagé sous cet éclairage, le rapport entre l'incidence de la consommation de services et le taux d'inoccupation des régions s'expliquerait par l'influence du TI sur la *présence* plus ou moins importante, selon les populations, des déclencheurs des maladies multifactorielles.

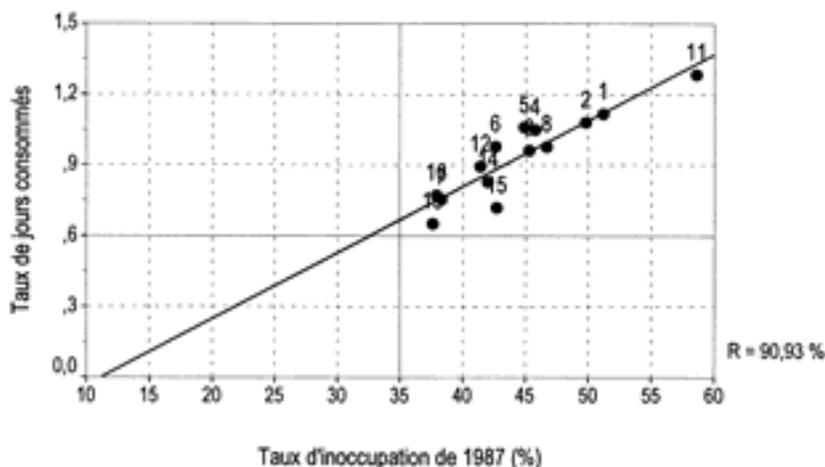
La portée explicative de cette théorie sur les faits mis en évidence

Contrairement à ce qu'on observe à l'échelle des personnes, certaines populations régionales cumulent systématiquement les plus fortes incidences de consommation pour la plupart des maladies organiques, en même temps. À l'opposé, d'autres populations en sont davantage épargnées sous tous les aspects. Cette cohérence territoriale de la consommation se vérifie même chez les représentants de tous les groupes d'âge : les enfants, les aînés et les adultes.

Or, à l'échelle du système organique des individus, il n'existe pas de déterminant qui puisse provoquer, chez une même personne, l'émergence simultanée de toutes les maladies organiques, contrairement à ce qui se vérifie à l'échelle des populations. Dans ce contexte, l'expression « état de santé des populations » apparaît fondée, dans la réalité, sur des caractéristiques spécifiques, *distinctes de la somme des caractéristiques propres aux individus qui en font partie*. À cet égard, comme sur d'autres aspects déjà évoqués, une population présente des caractéristiques qui en

GRAPHIQUE 10

Relation entre le taux d'inoccupation de 1987 et le niveau de consommation des régions en santé physique pour 1993



font un objet concret qui ne peut être confondu avec la simple collection statistique des individus qui composent cette population.

De plus, les faits mis en évidence dans le deuxième chapitre corroborent la pertinence des propos des experts consultés en montrant l'existence d'un seuil approximatif dans les taux régionaux de consommation (graphique 10).

On peut donc présumer de l'existence d'un seuil, d'une limite en deçà de laquelle le niveau de consommation des populations régionales soit attribuable à des facteurs biologiques aléatoirement distribués entre individus, sans égard à leur région d'appartenance. Dans le graphique 10, une ligne arbitraire tracée sous la région 13 illustre cette démarcation. Au-dessus de ce seuil cependant, d'autres facteurs propres aux régions d'appartenance – des facteurs sociaux – interviendraient pour déterminer le niveau de consommation spécifique à chaque population régionale. Là où s'arrêterait l'influence des facteurs strictement biologiques, on retrouve les populations dont le niveau de consommation est le plus faible, Laval dans l'exemple étudié. Là où s'arrêterait l'influence des facteurs sociaux, on retrouve celles dont le niveau est le plus élevé (la Gaspésie), le double des premières.

Sur la base des populations régionales du Québec, on ne peut établir, de façon absolue, le seuil biologique défini; on ne peut, non plus, établir les variations maximales imputables à des facteurs sociaux. Quant au seuil biologique, l'interprétation proposée tient compte du fait qu'il est difficile d'identifier des sociétés où le Tl soit très inférieur à 30 %. Cela signifierait pratiquement que 70 % et plus des individus de 15 ans et plus – incluant les personnes âgées – soient sur le marché du travail. La position occupée par Laval, en regard des autres régions, apparaît correspondre, dans ce contexte, à un seuil approximatif.

LES DÉTERMINANTS SOCIAUX AGISSANT SUR LES DÉCLENCHEURS DES MALADIES MULTIFACTORIELLES

Mieux éclairé sur ce qui, au plan biologique, pouvait activer en même temps nombre de maladies organiques, il importait ensuite de cerner les faits structurés qui, en lien avec le manque relatif d'emplois, expliquent et modulent l'intensité des principaux déclencheurs connus des maladies multifactorielles – présence chronique de stress, de mauvaise nutrition et d'environnements défavorables. Les travaux antérieurs menés sur les caractéristiques sociales des populations rendaient possible l'identification de cette partie du processus causal.

Le manque d'emploi et la faiblesse du revenu moyen

Le chapitre 1 a montré que le niveau de non-emploi des populations explique *directement et indirectement* leur niveau de pauvreté économique, mesuré par le revenu moyen des 15 ans et plus. Il l'explique directement, car l'emploi constitue, de loin, la principale source de revenus des populations; il l'explique indirectement lorsque la rareté de l'emploi contraint les personnes à recourir aux paiements de transferts publics (aide sociale, chômage, etc.) comme source de revenus exclusive. Les sources de revenu des populations *autres* que l'emploi et les transferts apparaissent négligeables et plus uniformément réparties entre les populations.

La faiblesse du revenu moyen et l'incidence des personnes financièrement contraintes

Le revenu moyen des populations mesure le rapport entre les masses monétaires disponibles dans les populations, et le volume de personnes qui doivent s'y ajuster pour se procurer les biens et services de toute nature – en particulier les biens et services de première nécessité²³ – pour

²³ Parmi les biens de première nécessité : le logement, la nourriture, le chauffage, le vêtement, le transport, le loisir.

assumer tous les impératifs et les aléas de la vie courante : les enfants aux études, les réparations et l'entretien du domicile, du véhicule, etc.

Ainsi et à l'*évidence*, une population « pauvre » se distingue d'une population « riche » par une plus forte incidence et donc une plus grande proportion de personnes financièrement limitées et même contraintes face aux nécessités de la vie quotidienne. Cette proportion de personnes contraintes peut même s'amplifier si les revenus disponibles sont, en plus, mal redistribués entre les individus et les familles²⁴ appartenant aux mêmes populations.

Les personnes sous contraintes financières exposées aux déclencheurs des maladies

La proportion de personnes limitées ou contraintes financièrement – qui distingue les populations les une des autres – mesure l'incidence des personnes exposées aux situations propres à activer les trois déclencheurs des maladies multifactorielles : la chronicité d'exposition au stress, à des problèmes de nutrition et d'environnement. De façon complémentaire, il faut noter que les régions où le TI est plus élevé montrent de plus fortes concentrations de travailleurs œuvrant dans l'industrie primaire et secondaire; ces personnes sont susceptibles d'être plus exposées, directement, à des conditions environnementales pathogènes.

Comme les facteurs biologiques, les facteurs sociaux relatifs aux populations échappent à l'emprise des personnes. Contrairement aux déterminants biologiques – endogènes aux individus – ces facteurs sociaux agissent de façon exogène sur la santé des personnes. Ainsi, selon leur population d'appartenance, les personnes déjà prédisposées biologiquement sont exposées, selon des niveaux différents d'intensité, à des conditions faisant en sorte qu'elles deviennent à la fois « pauvres et malades » plutôt que « riches et en santé ». Ainsi, dans les régions où l'emploi est rare, on trouve systématiquement une plus forte proportion de personnes pauvres et également une plus forte propension de la population à devoir recourir aux services hospitaliers lourds²⁵.

²⁴ À cet égard, l'indicateur le plus approprié du niveau de pauvreté des populations est fondé sur le revenu médian, les seuils de faibles revenus étant totalement contre-indiqués à l'échelle des régions pour mesurer leur niveau de pauvreté respective.

²⁵ Il faut, à cet égard, remarquer que les personnes ne s'hospitalisent pas elles-mêmes : la « consommation hospitalière » ne peut être confondue avec l'acquisition courante d'un bien ou d'un service. En effet, ce geste dépend d'une évaluation de situation par des professionnels de la santé et de l'arbitrage nécessaire des priorités d'hospitalisation par les établissements de traitement qui sont souvent hors de la région de provenance des malades.

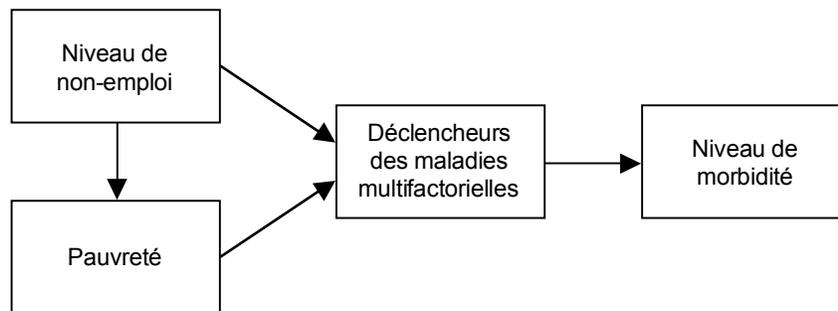
LA CHAÎNE CAUSALE ENTRE LE NIVEAU DE NON-EMPLOI ET LA MORBIDITÉ

Dans le contexte qui vient d'être décrit, il apparaît logique et évident que le processus qui active l'incidence de la maladie dans les populations – la morbidité – coïncide avec celui qui révèle l'expression même de la maladie à l'échelle des populations, soit leur niveau de consommation de services hospitaliers. Ainsi, les populations régionales dont le niveau de consommation est plus élevé sont des populations que des facteurs sociaux, déterminés en première instance par la rareté d'emploi, rendent plus malades.

Plus qu'un indicateur du niveau de consommation de services des populations régionales, le TI se pose alors comme la *mesure exacte de son principal déterminant*. Au plan théorique, le niveau de non-emploi explique, en première instance, le niveau de pauvreté des populations. Ce dernier, en seconde instance, mesure le degré d'exposition des populations aux trois déclencheurs majeurs des maladies multifactorielles (stress, malnutrition et environnements défavorables). Ces déclencheurs sont responsables, en troisième instance, de l'émergence des affections organiques de diverses natures chez les personnes déjà biologiquement prédisposées. Ce qui se manifeste, en dernière instance, par une consommation de services hospitaliers (diagramme 2).

DIAGRAMME 2

La chaîne causale entre le niveau de non-emploi et le niveau de morbidité des populations



Dans ce contexte, le niveau de non-emploi – mesuré par le TI – constituerait le *principal facteur de morbidité* qui différencie les populations régionales les unes des autres : ce que reflète leur niveau de consommation respectif. Dès lors, *sous réserve des analyses qui suivent*,

il apparaît pertinent d'avancer, conformément à une définition courante du concept de « morbidité » – soit l'ensemble des causes qui peuvent produire des maladies (v. Le Petit Robert) – que la corrélation entre le taux d'inoccupation et le niveau de consommation mesure, au-delà de l'association, une relation causale. La contribution prédictive du TI, parmi l'ensemble des déterminants possibles du niveau de consommation des populations régionales, est d'ailleurs égale, pour l'année 1993-1994, à 82,7 % de la variance du niveau de consommation de services (R^2)²⁶.

UN SECOND DÉTERMINANT DU NIVEAU DE CONSOMMATION DE SERVICES DES POPULATIONS RÉGIONALES : LA STRUCTURE D'ÂGE DES POPULATIONS

LES ÉLÉMENTS DE PROBLÉMATIQUE ET LES HYPOTHÈSES

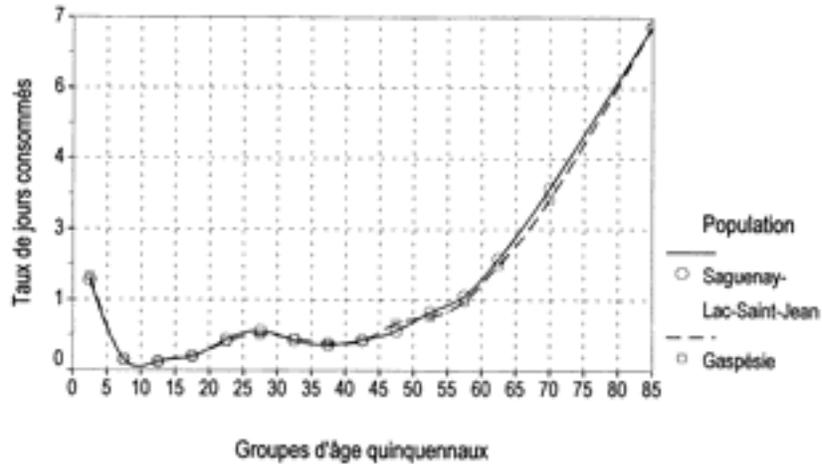
Il est peu vraisemblable, a priori, que le niveau de non-emploi des populations soit l'unique facteur social de leur morbidité. De même, il apparaît, à première vue, prématuré d'établir une relation d'équivalence entre « morbidité » et « consommation hospitalière », a fortiori, si le niveau de consommation est mesuré en taux de jours de courte durée. En effet, les taux de jours consommés pourraient varier d'une région à l'autre pour d'autres motifs que ceux reliés à la santé des personnes. Des raisons d'ordre organisationnel pourraient être en cause : notamment, des pratiques professionnelles et médicales différentielles; l'organisation des services ou la performance des établissements pourraient aussi influencer différemment le taux de jours consommés selon les régions.

Ces deux objections pertinentes doivent trouver réponse dans cet ouvrage. Par contre, on ne peut, à cette étape, s'attaquer aux raisons

²⁶ Tel qu'exprimé, le lien causal entre le TI et la morbidité, passant par l'intermédiaire de la pauvreté, représente un point de vue *minimaliste*. Certains auteurs, en effet, n'hésitent pas à établir un lien causal direct entre le manque d'emploi et le stress, identifiant même l'emploi comme une nécessité biologique de l'homme au même titre que « l'air qu'on respire, etc. ». C'est notamment le point de vue exprimé explicitement par le Dr Hans Selye (1974). Par ailleurs, d'autres travaux cités par le même auteur (Selye, 1974, p. 155) établissent un lien direct entre le stress et un très grand nombre de maladies organiques telles : « les migraines, les désordres de croissance, de développement et de métabolisme, les maladies des yeux, des voies respiratoires, de l'appareil gastro-intestinal, de l'appareil locomoteur, du système cardio-vasculaire, de la peau et des organes génitaux (*Association for research in nervous and mental disease*) ». Le point de vue minimaliste est ici retenu, car la réalisation de travaux ayant pour objet d'associer des affections organiques spécifiques à tel ou tel déclencheur particulier relève davantage des spécialistes du domaine biomédical.

GRAPHIQUE 11

Profil de consommation de services en courte durée pour deux populations selon la moyenne quinquennale de 1989-1993



d'ordre organisationnel avant d'avoir répondu à la première objection. Ainsi, la ligne directrice que nous suivrons pour aborder ces matières consistera, *en premier lieu*, à disposer des facteurs explicatifs et sociaux de la consommation, puis, *en second lieu*, à traiter de la consommation, comme telle, et des « paramètres » qui la composent et la mesurent.

Des indices révélant l'existence d'autres facteurs explicatifs

Même si le TI explique 82,7 % de la variance des niveaux régionaux de consommation, plusieurs indices permettent de poser l'hypothèse de l'existence d'au moins un autre facteur explicatif de la morbidité des populations et, par conséquent, de leur niveau de consommation de services. Au-delà de la variance de la consommation expliquée par le TI, il reste un résidu correspondant à 17,3 % de la variance, à mettre sur le compte d'autres facteurs explicatifs.

Par ailleurs, un indice trace la voie vers un autre élément de réponse. On a vu, au deuxième chapitre, que le TI mesure, pour chaque région, l'importance des taux de consommation de chaque groupe d'âge : la distribution graphique des taux de consommation ventilés par groupe d'âge trace d'ailleurs *le profil de consommation* des régions; un profil semblable pour toutes les populations, mais de hauteur inégale selon les

régions de provenance. D'ailleurs, le Tl mesure la hauteur relative des profils régionaux de consommation.

Dans l'éventualité où le profil par groupe d'âge de chaque région rende exactement compte du taux de jours global consommé par sa population totale, on devrait alors conclure que le Tl est, à l'échelle des populations, le seul facteur explicatif de leur niveau de morbidité. Par conséquent, les 17 % de la variance encore inexpliquée devraient être attribuables à d'autres facteurs, tels les facteurs de type organisationnel énumérés précédemment. Dans le cas contraire, la « différence » entre le taux global et la somme des taux par groupe d'âge laisse, comme seule autre hypothèse explicative plausible, *la structure d'âge des populations*.

Ce facteur pourrait notamment expliquer pourquoi deux régions, la Gaspésie et le Saguenay–Lac-Saint-Jean, ayant deux profils de consommation identiques et presque « parfaitement » égaux l'un à l'autre, groupe d'âge par groupe d'âge, ont, par contre, un taux de consommation global différent (graphique 11). En effet, la population de la Gaspésie a un taux de consommation global plus élevé que celui du Saguenay–Lac-Saint-Jean.

L'INFLUENCE DE LA STRUCTURE D'ÂGE : VÉRIFICATION DE L'HYPOTHÈSE

À l'évidence, pour chaque population, le nombre de jours d'hospitalisation (NJ) consommés par la population d'un groupe d'âge (P_i) est égal au taux de jours consommés par la population de ce groupe d'âge, multiplié par les effectifs de la population du groupe d'âge :

$$NJ_i = \frac{NJ_i}{P_i} \times P_i$$

En découlant, le nombre total de jours consommés par la population d'une région sera égal à la somme des jours consommés par la population de tous les groupes d'âge (16 groupes d'âge, en l'occurrence) :

$$NJ = \sum_{i=1}^{16} \left(\frac{NJ_i}{P_i} \times P_i \right)$$

Or, dans cette formule, P_i peut être remplacé par le produit de la population totale régionale (P) multipliée par la proportion que représentent les effectifs du groupe d'âge sur la population totale (pr_i). Alors :

$$P_i = P \times pr_i \text{ et :}$$

$$NJ = P \sum_{i=1}^{16} \left(\frac{NJ_i}{P_i} \times pr_i \right)$$

En divisant les deux membres de l'équation par la population totale (P), on constate alors que le taux de jours (NJ/P) global est égal pour quelque population que ce soit – régionale ou autre – au produit « membre à membre » des taux de jours (expliqués par le TI) multipliés par la proportion que représente chaque groupe d'âge sur la population totale (pr_i), soit *la structure d'âge de la population* :

$$\frac{NJ}{P} = \sum_{i=1}^{16} \left(\frac{NJ_i}{P_i} \times pr_i \right)$$

Il faut donc conclure qu'une autre caractéristique des populations régionales contribue à l'explication de leur niveau de consommation : la structure d'âge de chaque population à un moment donné.

LE CHOIX D'UN INDICATEUR DE LA STRUCTURE D'ÂGE DES POPULATIONS

La représentation de la structure d'âge d'une population par une seule valeur numérique pose une difficulté. En effet, par définition, la somme des proportions que constituent les effectifs de chaque groupe d'âge dans une population totale est toujours égale à l'unité : elle ne permet donc pas de discriminer les populations plus vieilles des populations plus jeunes. À cette fin, au moins trois indicateurs pouvaient être utilisés, soit l'âge moyen des populations, leur âge médian ou simplement *la proportion de personnes âgées de 65 ans et plus* parmi la population totale (PRA), pour l'année de consommation étudiée.

La troisième option a été retenue à partir de trois critères :

- ◆ L'existence d'un haut niveau de corrélation (95 %) entre la proportion de 65 ans et plus et la moyenne d'âge des populations régionales, d'après les données du Recensement de 1991.
- ◆ Le fait que les 65 ans et plus – environ 11 % de la population du Québec en 1991 – et les moins de 65 ans consomment presque le même pourcentage de l'ensemble des jours de courte durée consommés en santé physique; 45 % pour les 65 ans et plus contre 55 % pour les moins de 65 ans²⁷.

²⁷ Cette situation réelle est très voisine de celle qui mesure la variance maximale des proportions, [(0,45 X 0,55) ≡ (0,5 X 0,5)]; soit une situation conservatrice et idéale pour l'utilisation des méthodes statistiques sans risque d'erreur d'interprétation; voir à ce

- ◆ Le fait que la « proportion de personnes âgées de 65 ans et plus » soit déjà utilisée par nombre d'organismes comme indicateur de la structure d'âge des populations.

LA PORTÉE EXPLICATIVE DE LA STRUCTURE D'ÂGE DES POPULATIONS SUR LE NIVEAU DE CONSOMMATION DE SERVICES DES POPULATIONS RÉGIONALES

Les résultats et les objections pertinentes

En 1993, le taux d'inoccupation expliquait 82,7 % de la variance du niveau de consommation régionale. Or, la structure d'âge des populations (PRA) ajoute 7,3 % à l'explication, pour un total de 90,03 % de la variance du niveau de consommation des populations régionales expliqué par le TI et PRA, conjointement.

Une *première critique* peut être posée : si la proportion de personnes âgées (PRA) était introduite dans le modèle *avant* le TI plutôt qu'après, les résultats énoncés ci-haut seraient alors différents : 36,3 % de la variance s'expliquant par PRA, au lieu de 7 %, et 53,7 % par le TI, au lieu de 83 %, pour un total équivalent de 90 %²⁸.

La réponse à cette critique : il s'agit d'une objection fondée en théorie, *sauf s'il existe un lien explicatif entre le niveau de non-emploi et la structure d'âge* (TI et PRA). Le cas échéant, dépendamment de la variable qui influence l'autre, il est possible de déterminer lequel des deux résultats énoncés est conforme à la réalité. D'ailleurs, ce problème et sa solution sont traités dans les pages qui suivent.

La *seconde objection* repose sur le fait que, par définition, le taux d'inoccupation englobe la quasi-totalité des personnes âgées de 65 ans et plus. Dans ce contexte, ces deux mesures sont théoriquement redondantes et cette redondance devrait être mesurée par la corrélation existant entre les deux variables.

En réponse à cette critique : à l'évidence le TI de 1987 ne contient pas intégralement les mêmes personnes âgées que celles qu'on retrouve

sujet les techniques de mesure des intervalles de confiance des échantillons statistiques.

²⁸ Ces résultats ont été obtenus en utilisant la technique de la régression multiple. En outre, les réponses apportées à ces questionnements ne sont pertinentes que dans le cas où les variables sont impliquées dans un rapport de causalité. Si tel n'est pas le cas, alors les proportions de la variance mesurées par les variables indépendantes sont établies par une autre procédure : $R^2_{yy'} = (\beta_{1y} R_{1y}) + (\beta_{2y} R_{2y})$.

dans les populations régionales en 1993, car la structure d'âge a changé entre les deux moments. Ainsi, même si la critique est théoriquement fondée, les deux facteurs explicatifs sont, dans une large mesure, mutuellement exclusifs dans la réalité. De surcroît, l'influence que les deux indicateurs exercent respectivement sur le niveau de consommation de services n'est pas de même nature : le TI détermine spécifiquement le niveau de consommation régional pour tous les groupes d'âge – les personnes âgées incluses – alors que PRA mesure le poids qu'il faut donner au TI pour obtenir la meilleure estimation du niveau de consommation des régions, compte tenu de leur structure d'âge respective.

LES FACTEURS SOCIAUX RESPONSABLES DE 90 % DE LA VARIANCE DU NIVEAU DE CONSOMMATION DES POPULATIONS RÉGIONALES

Sachant que la proportion de personnes âgées et le taux d'inoccupation rendent respectivement compte de l'importance relative de la *structure d'âge* et du *profil de consommation par groupe d'âge*, sachant aussi que ces deux dernières composantes, ensemble, mesurent algébriquement la totalité du taux de jours consommés par chaque population, alors il ne peut exister d'autres facteurs liés à des caractéristiques de population pour expliquer le niveau de consommation des populations régionales²⁹. Dans ce contexte, les 10 % de la variance des taux de jours non encore expliqués, par le TI et PRA, ne peuvent être attribuables qu'à des facteurs explicatifs d'ordre organisationnel.

L'INFLUENCE DU NIVEAU DE NON-EMPLOI SUR LA STRUCTURE D'ÂGE DES POPULATIONS RÉGIONALES

LE NIVEAU DE NON-EMPLOI ET LES MIGRATIONS INTERRÉGIONALES

Les éléments de théorie

Plusieurs travaux de recherche réalisés au Québec³⁰ à partir des années 1980 ont mis en lumière la *relation explicative* entre la disponibilité d'emploi, qui distingue les régions entre elles, et la propension systématique des résidents des régions où le niveau d'emploi est chroniquement faible à émigrer vers les régions où le niveau d'emploi est plus favorable. Au plan théorique, il s'agit là de confirmations d'un phéno-

²⁹ Compte tenu également des résultats présentés au deuxième chapitre sur les indicateurs, et du rapport de causalité entre le TI et les déclencheurs des maladies multifactorielles.

³⁰ Voir, entre autres, Côté (1982).

mène qui se vérifie non seulement dans toutes les sociétés humaines, mais aussi chez les espèces animales dans leur rapport aux ressources vitales.

Tel que démontré au deuxième chapitre, le niveau d'emploi détermine le niveau de revenu moyen des populations et, ce faisant, la capacité « moyenne » des individus qui en font partie d'acquérir les biens matériels et services nécessaires à l'existence, conformément à la définition du mot *besoin*. Ainsi, et comme cela se constate chez les espèces animales, lorsqu'un milieu n'offre pas les conditions nécessaires à leur existence, les individus émigrent vers d'autres milieux où ces exigences vitales ont plus de chance d'être satisfaites.

Les bilans migratoires interrégionaux : un déterminant de la structure d'âge des populations

Les études réalisées sur le sujet ont également démontré que cette pension des individus à émigrer d'une région à l'autre n'est pas équivalente pour tous les groupes d'âge. En effet, les tendances systématiques observées partagent les populations en deux groupes principaux :

- ◆ celui formé des personnes âgées de moins de 15 ans et des 35 ans et plus : ce groupe est davantage captif ou enraciné dans le milieu, donc moins mobile;
- ◆ le groupe des 15-34 ans : ce contingent parvenu à la période où l'on accède aux études supérieures et au marché du travail s'avère beaucoup plus mobile que le premier.

La mobilité géographique des 15-34 ans constitue le principal facteur direct du changement de la structure d'âge d'une population entre deux recensements. Cette mobilité a aussi un effet indirect : en effet, puisque 90 % des bébés naissent d'une mère âgée entre 20 et 34 ans, l'exode des personnes en âge de procréer et de se reproduire contribue radicalement à la chute des naissances et ce, sans égard à la fécondité des candidates à la maternité.

Ainsi, l'exode des jeunes, conjugué à la chute des naissances en découlant, déterminent ensemble la vitesse de changement de la structure d'âge des populations – la vitesse de leur vieillissement – mesurable par l'évolution de la proportion de 65 ans et plus (PRA) dans les régions. Dans le temps long, ces facteurs déterminent aussi la vitesse à laquelle ces populations se désintègrent démographiquement, par la perte progressive de leur capacité de remplacement naturel (Côté, 1991).

LA MESURE DU BILAN MIGRATOIRE DES PERSONNES EN ÂGE DE SE REPRODUIRE

Dans le groupe des 15-34 ans, la mortalité contribue de façon négligeable à l'explication du changement observable des effectifs démographiques des populations régionales, entre deux recensements. De surcroît, ce facteur de changement – la mortalité des 15-34 ans – est distribué entre les régions de façon relativement uniforme, de telle sorte que les variations des effectifs régionaux sont essentiellement imputables à leurs bilans migratoires et ce, avec une marge de sécurité de l'ordre de 97 %.

Pour mesurer l'importance des bilans migratoires des jeunes âgés de 15-34 ans, nous avons comparé les effectifs des 20-34 ans sur quatre recensements consécutifs aux effectifs des mêmes cohortes, soit celles des 15-29 ans des recensements tenus cinq ans auparavant. Un bilan positif mesure le volume net de jeunes qui se sont ajoutés à certaines populations régionales au cours de l'intervalle intercensitaire; un bilan négatif mesure le nombre net de ceux qui ont quitté les autres populations au cours de cette période quinquennale.

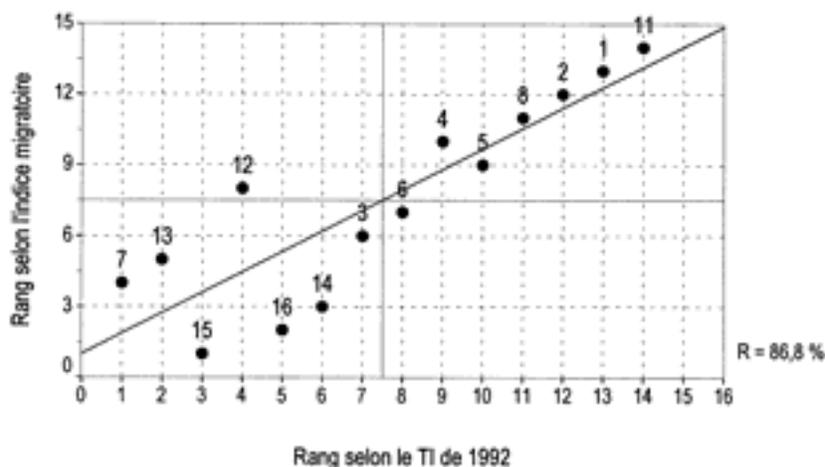
Pour chaque région, a été dressée la somme des bilans – positifs ou négatifs – des quatre périodes intercensitaires s'échelonnant entre 1971 et 1991. Cette somme correspond aux effectifs nets de personnes différentes, en âge de se reproduire, qui ont quitté une région ou qui s'y sont ajoutées, avant d'avoir atteint l'âge de 35 ans. Ce bilan est conservateur, car il ne tient évidemment pas compte de tous ceux qui auraient quitté leur région d'origine en compagnie de leurs parents, avant d'avoir atteint l'âge de 15 ans.

Cette procédure a permis de confectionner un indice pour mesurer la vitesse de changement de la structure d'âge des régions non imputable au vieillissement naturel et attribuable en majeure partie³¹ à la migration interrégionale. Pour chaque région, cet indice est égal au rapport entre la somme des bilans migratoires intercensitaires entre 1971 et 1991, et les effectifs de référence, soit pour chaque région, les personnes de 15 à 29 ans présentes en 1971. Ainsi mesuré, l'indice a la signification suivante :

³¹ Il faut noter que le bilan tient nécessairement compte des personnes entrées dans les régions en provenance de l'extérieur du Québec alors qu'il ne peut, à l'évidence, mesurer la part d'exode des jeunes vers l'extérieur du Québec. Les régions qui perdent leurs jeunes attirent peu l'immigration étrangère, comme le manifeste le bilan.

GRAPHIQUE 12

Corrélation par rang entre le TI de 1992 et l'indice de la vitesse d'exode hors-région des 15-29 ans entre 1971 et 1991



il est égal au pourcentage de ceux qui, parmi ce groupe, auraient quitté leur région entre 1971 et 1991 dans l'hypothèse *volontairement fictive* où tous ces jeunes auraient cessé de vieillir en 1971, le temps s'étant arrêté de couler jusqu'en 1991. Cet indice, *l'indice migratoire*, permet ainsi d'établir globalement l'ampleur des pertes (émigration ou exode) ou des gains régionaux (immigration), en regard d'une population de référence statique.

LE NIVEAU DE NON-EMPLOI ET LA VITESSE DE CHANGEMENT DE LA STRUCTURE D'ÂGE DES POPULATIONS RÉGIONALES

Le coefficient de corrélation par rang (*rho* de Spearman) entre l'indice mesurant la vitesse d'exode des jeunes et le taux d'inoccupation de 1992 est égal à 86,8 % : globalement, plus le niveau d'emploi est chroniquement faible dans les régions – donc, plus leur population est en besoin – plus importante sera la proportion des jeunes en âge de se reproduire qui quitteront la région comme cela se passe à l'échelle écologique pour les autres espèces animales (graphique 12). Par conséquent, plus la structure d'âge de ces populations changera rapidement dans la direction d'un vieillissement accéléré et, éventuellement, d'une extinction démographique.

On pourrait ici s'interroger sur l'utilisation de corrélations par rang plutôt que des corrélations linéaires pour mesurer la relation entre les deux variables : vitesse d'exode et taux d'inoccupation. Une raison d'ordre théorique et méthodologique l'explique. En effet, les résultats obtenus graphiquement illustrent une particularité des processus migratoires en tous points conforme aux théories explicatives sur le sujet, particularité dont les corrélations linéaires, fondées sur les modèles paramétriques, ne rendent pas compte de façon adéquate : les « *push and pull factors* »³².

La théorie des *push factors* et des *pull factors*

Pour tout individu qui prend la décision d'émigrer, le processus migratoire s'effectue sur deux ordres de critères consécutifs. Les premiers sont établis à partir des raisons qui l'incitent à quitter un endroit : les *push factors*; les seconds reposent sur les motifs d'aller s'établir dans un endroit plutôt que dans un autre : les *pull factors*. Les premiers qualifient, en propre, les régions d'exode alors que les seconds qualifient, en propre, les régions d'accueil.

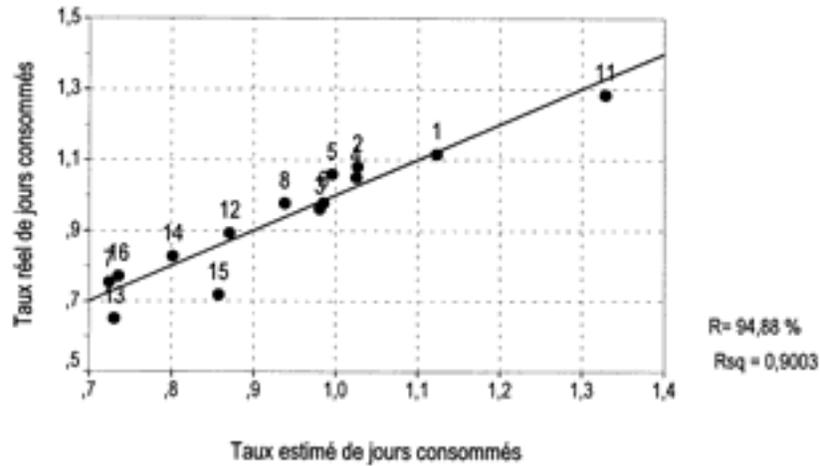
Ainsi, au Québec pour la période 1971-1991 :

- ◆ de façon systématique, les régions où le niveau de non-emploi (Tl) est supérieur à la moyenne provinciale ont un indice d'exode supérieur à zéro alors que les autres ont un indice inférieur à zéro, manifestant ainsi leur potentiel d'attraction sur les premières régions;
- ◆ dans les premières régions, la relation entre le Tl et l'indice est linéaire : plus le Tl est élevé, plus l'indice d'exode est élevé. Il s'agit ici de la manifestation de l'influence des *push factors*, pour les raisons théoriques évoquées précédemment;
- ◆ dans les régions d'accueil, au contraire, le rapport entre le Tl et l'attraction démographique est aléatoire (nul) : une manifestation de l'influence des *pull factors*. En effet, dans les régions où l'accès à l'emploi est globalement favorable a priori, le choix d'un lieu de résidence se fonde sur tout autre critère de préférence que ceux qui président à l'accès aux biens de première nécessité – « ces choses considérées nécessaires à l'existence et que l'on obtient par de l'argent ». Dans ce cas, le choix découle de critères tels l'environnement physique et « social », la proximité du lieu de travail ou même l'éloignement des zones achalandées, la proximité de proches parents, etc.

³² Voir les modèles de Beale et Lowry dans Hoover (1971, pp. 161-198).

GRAPHIQUE 13

Relation entre le taux réel de jours consommés et le taux estimé de jours par région pour 1993-1994



L'INFLUENCE DU TAUX D'INOCCUPATION DE 1987 SUR LA PROPORTION DE PERSONNES ÂGÉES PAR RÉGION EN 1993

L'ensemble de ces constats implique nécessairement que la corrélation de 39 % entre le TI et la proportion de personnes âgées (PRA) – deux déterminants du niveau de consommation de services des populations régionales que nous avons déjà associée à la mesure de leur niveau de morbidité – n'est pas fortuite.

Le « problème » de multicollinéarité statistique révélé par la corrélation entre ces deux variables indépendantes du niveau de consommation implique, en réalité, l'existence d'un lien causal entre ces mêmes variables. Un lien causal où le TI influence la vitesse relative de changement de la structure d'âge des populations et par conséquent la vitesse à laquelle s'accroît la proportion de personnes âgées de 65 ans et plus dans les régions entre deux recensements et ce, distinctement du vieillissement naturel des personnes.

En effet, dans une population où les aînés sont stables et où les jeunes émigrent, la proportion de personnes âgées s'accroîtra à l'évidence rapidement, a fortiori si une telle situation prévaut sur une longue période.

Ceci permet également de trancher le dilemme théorique évoqué précédemment concernant la contribution explicative du taux d'inoccupation (TI) et de la proportion de personnes âgées (PRA) sur la variance du taux de consommation de services. En réponse à la question restée pendante, le TI explique 82,7 % de la variance du niveau de consommation et PRA ajoute 7,3 % d'explication, pour un total de 90,03 %, en 1993 (graphique 13). Pour chaque région, le taux de jours estimé (E) se calcule comme suit :

$$E = (PRA_{1993} \times 0,02627) + (TI_{1987} \times 0,02450) - 0,44685$$

LE DÉVELOPPEMENT DU MODÈLE EXPLICATIF DES NIVEAUX RÉGIONAUX DE CONSOMMATION DE SERVICES

UN PREMIER BILAN

Le plan théorique

Le niveau de non-emploi (NE) des populations régionales détermine leur niveau de morbidité (M), par l'intermédiaire du niveau de pauvreté et de l'influence qu'exerce celle-ci sur les déclencheurs des maladies multifactorielles : situation chronique de stress, d'insuffisance alimentaire, d'exposition à des environnements défavorables. Ces déclencheurs agissent sur les personnes en fonction de leurs prédispositions biologiques. La morbidité d'une proportion plus ou moins grande de personnes affectées au sein des populations détermine leur niveau respectif de consommation.

La structure d'âge des populations (SA) influence également leur niveau de consommation; mais, à cette étape, rien ne permet d'avancer, au plan théorique, que la structure d'âge contribue à expliquer la morbidité des populations. Cette hypothèse reste encore à vérifier.

Le niveau de non-emploi (NE) détermine la vitesse du changement de la structure d'âge des populations (SA), attribuable à d'autres facteurs que le vieillissement naturel des personnes.

Les correspondances au plan méthodologique

Les relations identifiées, au plan théorique, correspondent à la formalisation suivante au plan méthodologique.

Dans chaque population régionale, le taux d'inoccupation (TI) influence le taux de jours consommés, soit le nombre de jours d'hospitalisation divisé par les effectifs de la population (NJ/P).

La proportion des personnes âgées de 65 ans et plus dans les populations (PRA) influence le nombre de jours consommés, en moyenne, par personne (NJ/P).

Le taux d'inoccupation (TI) influence la proportion de personnes âgées (PRA).

Première ébauche d'un modèle explicatif « social »

Les diagrammes 3 et 4 ébauchent un premier modèle explicatif sur les plans théorique et méthodologique. Les corollaires de ce modèle impliquent que le niveau de pauvreté, la vitesse à laquelle des populations régionales perdent la capacité de remplacement naturel et la morbidité des mêmes populations constituent trois effets interreliés et attribuables à un déterminant commun : la chronicité des niveaux de non-emploi de certaines populations régionales comparativement aux autres.

Dans ce contexte, les trois effets sont autant de manifestations d'un même besoin social : une carence chronique relative « aux choses considérées nécessaires à l'existence des *populations* et que l'on obtient aussi avec de l'argent », soit un niveau d'emploi raisonnable.

DIAGRAMME 3. Le plan théorique du modèle « social »

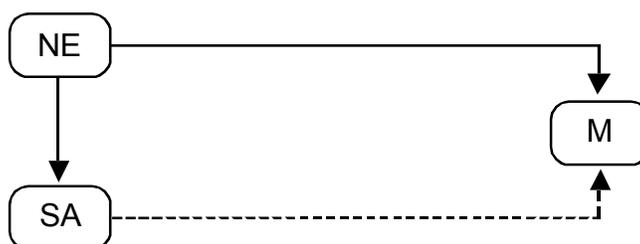
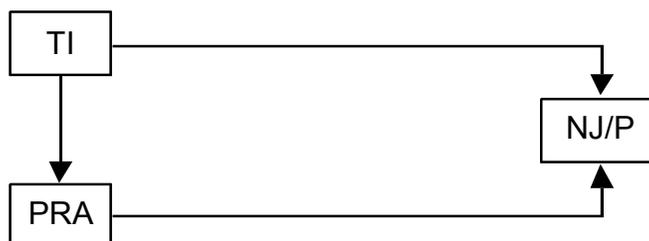


DIAGRAMME 4. Le plan méthodologique du modèle « social »



Critique du modèle théorique

Le concept de « morbidité » a déjà été défini par les « causes qui peuvent produire les maladies ». Dans ce contexte, il serait sémantiquement erroné d'établir, d'emblée, une équivalence entre la morbidité et la consommation de services, variable dépendante de deux causes présumées (NE et SA). Tel qu'évoqué dans le cas de l'espérance de vie, la morbidité ne coïncide pas avec la consommation de services hospitaliers, au plan conceptuel.

En réponse à cette objection, il est vrai que le concept de morbidité réfère aux causes de la maladie. Mais il est également vrai que le même concept se définit aussi par « le nombre relatif [la proportion] de [personnes] malades dans un groupe donné [une population] pendant un laps de temps donné » (v. Le Petit Robert). On constate donc que *le concept de morbidité s'applique à la fois aux causes et aux effets démographiques de la maladie.*

Et dans le contexte où, abstraction faite des prédispositions biologiques, le niveau de non-emploi et la structure d'âge des populations semblent constituer les deux seuls déterminants possibles du niveau de consommation de services hospitaliers attribuables aux caractéristiques des populations régionales, alors cette critique apparaît non fondée. Par ailleurs, l'influence de la structure d'âge sur la morbidité des populations reste à vérifier, et si telle hypothèse ne pouvait être soutenue, il faudrait alors en déduire et conclure que le niveau de non-emploi est le seul facteur explicatif du niveau de morbidité des populations régionales.

Critiques du modèle méthodologique

À ce qui vient d'être énoncé, on pourrait objecter que le taux de jours consommés en milieu hospitalier ne mesure pas explicitement l'incidence des personnes malades dans les populations régionales : on ne peut donc, de prime abord, établir d'équivalence entre la morbidité des populations et les taux de jours d'hospitalisation consommés par ces dernières. On peut aussi objecter que les personnes hospitalisées ne comprennent pas l'ensemble des personnes malades d'une population donnée, pour une période de temps déterminée.

En réponse à ces objections, s'il est vrai que le taux de jours consommés en milieu hospitalier n'est pas égal à la proportion de personnes hospitalisées d'une population, il est vrai aussi que les taux de jours consommés sont essentiellement un attribut des personnes hospitalisées. Ce faisant, s'il est vrai que les taux de jours sont en partie explicables par des facteurs organisationnels – tels la performance d'établisse-

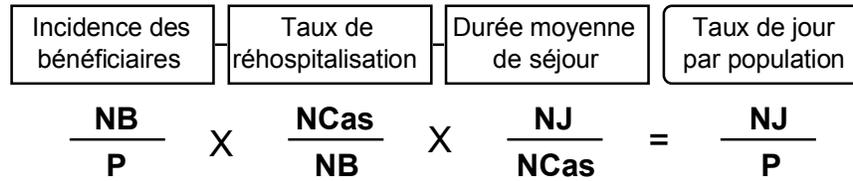
ment, etc. – il faut se rappeler que ces facteurs expliquent, tous ensemble, *moins de 10 %* de la variation interrégionale des taux de jours consommés. Ce qui sera démontré plus explicitement dans les prochains paragraphes.

Par ailleurs, en regard de l'objection sur l'inégalité entre le taux de personnes malades et le taux de personnes hospitalisées, celle-ci renvoie à un questionnement plus fondamental : celui de l'existence d'une ligne de démarcation, d'un seuil permettant d'établir la différence entre une simple indisposition organique – un rhume, une indigestion, une éraflure, une migraine passagère, un malaise au lendemain d'une soirée trop bien arrosée – et un trouble organique suffisamment sérieux, développé et observable chez une personne pour qu'un tiers, un professionnel de la santé ayant les compétences requises pour diagnostiquer l'état de santé physique des individus, prenne la décision d'admettre cette personne dans un milieu approprié, le milieu hospitalier, pour qu'elle y reçoive les soins exigés par son état.

Cette ligne de démarcation peut se situer au niveau de l'hospitalisation qui correspond à une *décision* de consommation de services qui ne relève pas de la personne qui consomme, mais plutôt de la décision d'un tiers professionnel dont la compétence médicale a été préalablement sanctionnée. La personne hospitalisée est donc retirée temporairement de son milieu à partir d'une décision prise par un tiers ayant seul les compétences requises pour juger du bien-fondé d'une telle décision. À notre point de vue, ce sont là deux critères qui permettent de distinguer ce type de consommation de toutes les autres, telles la consultation en cabinet privé parce qu'on se sent mal, ou celle d'acheter un médicament, sans ordonnance, contre les symptômes d'un rhume.

Une objection plus fondamentale touchant le modèle méthodologique

Partant de la réponse qui précède, il apparaît logiquement impossible que le taux d'inoccupation (TI) et la proportion de personnes âgées (PRA) influencent *directement* le taux de jours consommés. En effet, les hospitalisations et leurs conséquences relèvent, en première instance, de décisions distinctes et individuelles arrêtées exclusivement par des professionnels de la santé. On doit donc, à l'évidence, constater que les deux modèles, théorique et méthodologique, sont clairement incomplets. En effet, comment des facteurs sociaux, mesurés par TI et PRA, peuvent-ils rendre compte des taux de jours, sans avoir à rendre compte de la propension des médecins à hospitaliser ou réhospitaliser les personnes qui consomment les jours d'hospitalisation?

DIAGRAMME 5**Démonstration algébrique de la mesure du taux de jours
par le produit de trois concepts**

Une telle critique permet même de mettre au jour l'existence d'un modèle explicatif d'un type particulier qui, sans égard à l'explication du niveau de consommation par des facteurs sociaux – des caractéristiques de populations indépendantes des pratiques professionnelles ou hospitalières – explique et mesure à 100 % ce niveau par des facteurs organisationnels. Ces facteurs particuliers sont liés essentiellement aux pratiques propres au secteur de la santé, en général, et au milieu hospitalier, en particulier.

LE DÉVELOPPEMENT DU MODÈLE EXPLICATIF « SANITAIRE »³³**Trois concepts**

Le nombre de jours d'hospitalisation consommés, en moyenne, par une population quelconque (NJ/P) s'explique et se mesure entièrement par le produit de trois facteurs (diagramme 5) :

- ◆ *l'incidence des bénéficiaires par population* (NB/P) est égal au rapport entre le nombre de personnes différentes ayant été hospitalisées au moins une fois au cours d'une période annuelle et l'effectif global de la population considérée (il s'agit dans ce cas de la mesure exacte du concept de *morbidité*);
- ◆ *le taux de réhospitalisation* (TRH) est égal au rapport entre le nombre d'admissions hospitalières – les cas d'hospitalisation – et le nombre de bénéficiaires originaires de chaque région, pour la même période (NCas/NB);
- ◆ *la durée moyenne de séjour* (DUR) est égale au nombre de jours de courte durée consommés, en moyenne, par admission en milieu hospitalier (NJ/NCas).

³³ Ce modèle est dit « sanitaire » en ce qu'il réfère aux caractéristiques mêmes du système de santé. Il est désigné ainsi pour le distinguer du modèle « social » présenté précédemment.

Les relations logiques et explicatives entre les concepts

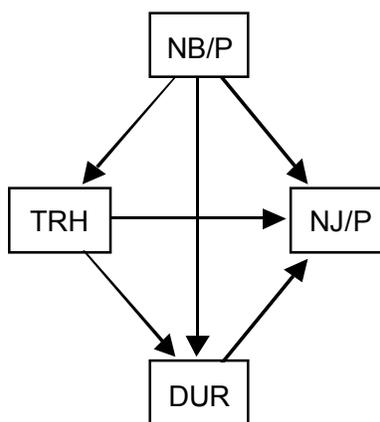
Les décisions arrêtées par des professionnels de la santé d'hospitaliser les personnes malades précèdent leur hospitalisation et leur durée de séjour.

L'hospitalisation des personnes précède leur durée de séjour en milieu hospitalier.

Le modèle « sanitaire » pour expliquer le niveau de consommation de services des populations régionales

Compte tenu des concepts et de leurs relations, on peut donc représenter comme suit ce modèle qui décrit à 100 % les niveaux régionaux de consommation (diagramme 6).

DIAGRAMME 6
Le modèle sanitaire



LE DÉVELOPPEMENT D'UN MODÈLE EXPLICATIF SOCIO-SANITAIRE

Les modèles « social » et « sanitaire » (diagrammes 4 et 5) sont explicatifs de la même variable : le nombre de jours consommés par population (NJ/P). Le modèle « social » explique cette variable par deux déterminants présumés du niveau de morbidité des populations régionales (TI et PRA). Le modèle « sanitaire » l'explique par trois concepts ou paramètres distincts dont le produit rend compte à 100 % du nombre de jours consommés par les populations. Sous un autre angle, ce modèle l'explique aussi par l'ensemble des activités sanitaires qui

sous-tendent la fonction d'hospitalisation des personnes affectées par des maladies organiques, notamment :

- ◆ *la pratique médicale*, c'est-à-dire des actes professionnels débouchant sur la décision d'hospitaliser, ou non, des personnes, une ou plusieurs fois au cours d'une période définie;
- ◆ *des facteurs organisationnels* pouvant entraîner, pour une même personne, une suite d'hospitalisations dans un laps de temps défini, par exemple, le fait pour une personne de devoir amorcer sa convalescence dans un centre hospitalier autre que celui où certains soins intensifs lui auraient été dispensés;
- ◆ dans chaque cas d'admission hospitalière, *une durée de séjour* plus ou moins longue en milieu institutionnel, celle-ci variant en fonction de la gravité des états, de la disponibilité de « ressources alternatives » ou encore d'autres facteurs.

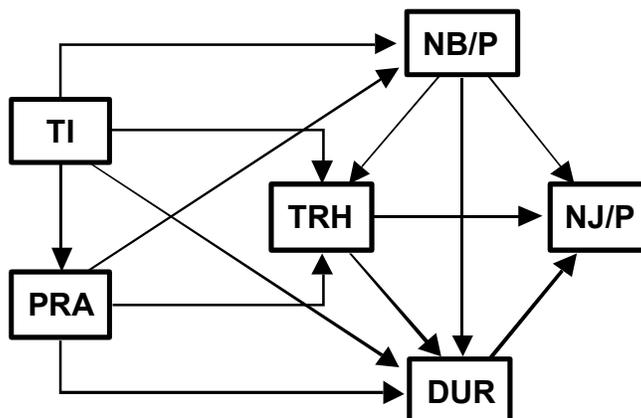
À l'égard des modèles « social » et « sanitaire », se pose l'alternative suivante :

- ◆ l'un des deux modèles est fondé théoriquement et empiriquement, alors que l'autre ne l'est pas;
- ◆ les deux modèles sont fondés théoriquement et dans la réalité, mais l'un des deux est lui-même explicatif de l'autre.

Le premier volet de l'alternative, s'il était vrai, poserait le problème suivant : comment les déterminants sociaux des taux de jours peuvent-ils déterminer ceux-ci sans être capables de rendre compte des facteurs empiriquement valables qui fondent la consommation, telles l'incidence des personnes hospitalisées, leur durée de séjours, leur propension à la réhospitalisation? Il apparaît évident, a priori, qu'il faille écarter l'idée d'une non-correspondance entre les deux modèles. Si le modèle « social » ne peut rendre compte du modèle « sanitaire », il faudrait alors comprendre que le modèle social est non valide et que, par conséquent, il n'y a probablement pas de facteurs appartenant en propre aux populations pour expliquer leur état de santé.

Le second volet de l'alternative suppose donc *nécessairement* que chacun des facteurs sociaux (TI et PRA) influence chacun des facteurs du modèle sanitaire (NB/P, TRH et DUR). L'inverse est difficilement imaginable, du moins de façon significative. Cette fusion des deux modèles impliquerait que les déterminants présumés de la morbidité influencent nécessairement la réponse institutionnelle à la maladie : les décisions d'hospitaliser ou non, les fréquences de réhospitalisation et les durées de

DIAGRAMME 7
Le modèle sociosanitaire



séjour. Ainsi, chaque variable explicative du modèle « sanitaire » serait elle-même influencée par chacune des deux variables du modèle « social » et cela nécessairement³⁴. En fait, le modèle « sanitaire » est intercalé dans le modèle social pour constituer avec celui-ci *le modèle sociosanitaire* (diagramme 7).

Compte tenu des mises au point qui précèdent et sous réserve des vérifications présentées plus loin, le modèle rend compte explicitement du niveau de morbidité des populations régionales, au sens où ce concept est défini par « le nombre relatif (la proportion) de personnes malades dans une population » (v. Le Petit Robert). Il s'agit donc d'une influence spécifique qu'exercent théoriquement les variables TI et PRA sur la variable « incidence des bénéficiaires », soit le taux de personnes différentes ayant été hospitalisées au moins une fois au cours d'une année dans une population (NB/P)³⁵.

³⁴ Le chapitre suivant qui traite de la démonstration du modèle causal, traite aussi des éléments qui attestent du bien-fondé de cette déduction logique. En particulier, concernant la mesure des influences exercées par chaque variable du modèle, sur chacune des autres variables.

³⁵ La vérification de ce lien théorique apparaît au quatrième chapitre. Il y sera notamment vérifié que seul le TI influence significativement la morbidité des populations.

Critique théorique du modèle sociosanitaire

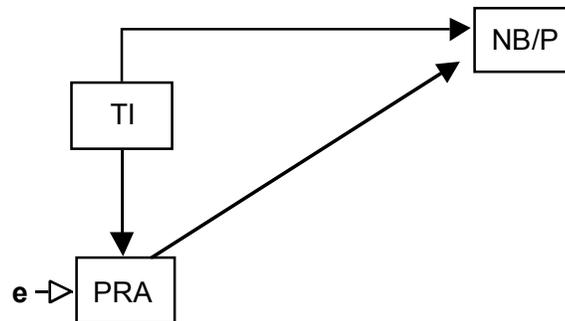
D'autres facteurs, *exogènes* au modèle ceux-là, doivent nécessairement influencer les variables du modèle.

À l'évidence, *ni le niveau de non-emploi ni le taux d'inoccupation n'ont d'existence autonome dans la réalité*. Ils sont eux-mêmes les effets d'autres facteurs qui contribuent à en fixer la valeur, pour chaque population, à un moment donné.

La structure d'âge des populations influence les autres variables du modèle, alors que le niveau de non-emploi influence spécifiquement la vitesse de changement de cette structure d'âge des populations. Alors qu'en est-il par exemple des facteurs historiques qui ont dans le temps long contribué à modeler la structure d'âge d'une population et qui contribuent, par-là, implicitement à l'explication des autres variables du modèle explicatif? En effet, dans le cas de la structure d'âge, outre les vitesses différentes de changement de cette structure d'âge imputables au niveau d'emploi, il est clair que *l'état historique du peuplement* à un moment donné explique *aussi* les différences de structure d'âge entre les populations régionales. Cet état historique s'avère donc la variable exogène « e » contribuant à fixer la valeur de PRA pour chaque région et à déterminer une influence – indépendante du TI – agissant sur les autres variables du modèle.

DIAGRAMME 8

L'influence d'une variable exogène



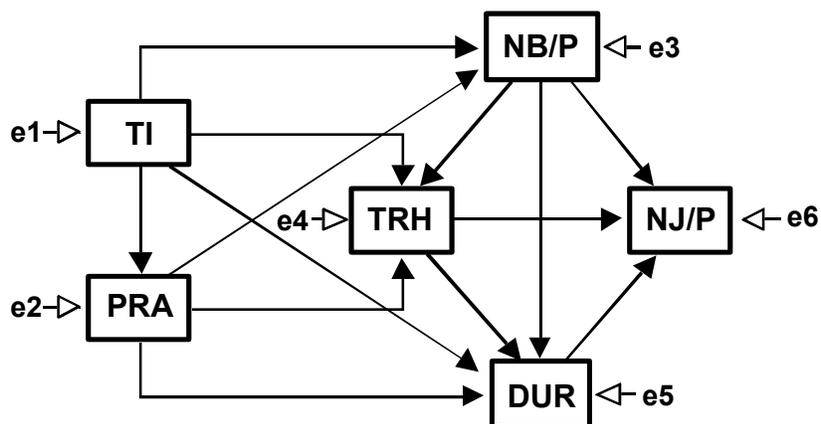
De même, en regard de chacune des autres variables du modèle – l'incidence de bénéficiaires (NB/P), les taux de réhospitalisation (TRH), les durées moyennes de séjour et même les taux de jours consommés (NJ/P) – n'existerait-il pas, dans chaque cas, au moins une variable « cachée » et exogène susceptible d'exercer, comme pour la structure d'âge, une influence implicite sur les autres variables du modèle?

LES VARIABLES EXOGÈNES QUI INFLUENCENT LE MODÈLE SOCIO-SANTITAIRE

Dans la vérification du modèle sociosanitaire, aucune variable dépendante d'une ou de plusieurs autres variables indépendantes n'est expliquée par ces dernières à 100 %. En preuve, aucune corrélation entre les variables n'est égale à 100 %. Ce résidu non expliqué par l'influence des variables indépendantes est donc *nécessairement* attribuable à l'influence de variables exogènes « e » agissant sur chacune des variables du modèle. D'ailleurs, au plan méthodologique, ce sont ces variables exogènes qui sont, en définitive, *responsables de l'accroissement de la variance totale* expliquée par l'ajout cumulatif d'une variable indépendante dans le modèle explicatif général.

Ainsi, l'ajout de PRA dans le modèle n'accroît la variance expliquée du niveau de consommation que dans la mesure où une partie de l'influence de PRA n'est pas elle-même dérivée de l'influence du TI sur PRA : en effet, la variance du niveau de consommation expliquée par le TI comprend toutes les influences directes et indirectes de TI. Donc la part d'explication additionnelle obtenue par l'inclusion de PRA dans le modèle est due uniquement à « e ». En fait, les variables exogènes sont *les véritables variables explicatives du niveau de consommation de services* alors que les variables du modèle en sont les « supports ». L'ajout de ces variables exogènes permet de parachever le modèle explicatif complet du niveau de consommation des populations régionales.

DIAGRAMME 9
Le modèle explicatif complet



Les variables exogènes sont :

e1 = L'ensemble des causes qui, de l'extérieur du modèle contribuent à déterminer le niveau de non-emploi des régions : les décisions à incidence économique *incluant les modes de répartition des fonds publics entre les régions*.³⁶

e2 = *L'historique du peuplement des régions*, soit le seul facteur possible qui, ajouté à la vitesse de changement de la structure d'âge déjà expliquée par le TI, détermine, en totalité, la structure d'âge des populations régionales à un moment donné.³⁷

e3 = Les différences de *pratique professionnelle* (médicale) en usage dans les régions de traitement. Ce facteur, conjointement avec la pauvreté des populations expliquée par le TI et présumément avec leur structure d'âge mesurée par PRA, contribue à expliquer, en totalité, l'incidence du nombre de bénéficiaires hospitalisés provenant de chaque région (NB/P).

e4 = Les différences régionales mettant en cause *l'organisation des services* des régions s'ajoutent aux trois facteurs précédents, pour expliquer, en totalité, le taux de réhospitalisation des clientèles dans les populations régionales. À titre d'exemple, les déséquilibres en disponibilité de lits courte et longue durée.

e5 = Les différences mettant en cause *la performance des établissements*, de concert avec les quatre facteurs précédents du modèle, contribuent à expliquer, en totalité, la durée moyenne de séjour des clientèles hospitalisées.

e6 = *L'erreur attribuable à la méthode*. Tel que démontré, le taux de jours est à 100 % égal au *produit* de trois variables, alors que la seule méthode utilisable pour rendre compte de l'ensemble du modèle est fondée sur *l'addition* du poids des facteurs par la régression multiple. En découle une légère distorsion entre la réalité étudiée et le résultat obtenu par régression (v. le principe d'isomorphie des méthodes). Cette erreur est techniquement corrigible en exprimant les valeurs des trois variables explicatives sous forme logarithmique, ce qui n'a pas été fait compte tenu du caractère très négligeable de cette erreur. Ceci sera démontré au chapitre suivant.

³⁶ Les chapitres 6 à 9 en traitent.

³⁷ Les chapitres 4 à 7 en traitent.

RÉSUMÉ DU MODÈLE EXPLICATIF THÉORIQUE

LES CONCEPTS IMPLIQUÉS

Le modèle théorique développé pour expliquer la nature du rapport entre le niveau d'emploi et le niveau de consommation de services est un modèle causal applicable à de grandes populations, comme celles des régions du Québec. Ce modèle rend compte des facteurs biologiques permettant d'expliquer les maladies organiques, des facteurs économiques et « socio-économiques » susceptibles d'expliquer le déclenchement de la maladie et d'engendrer des impacts démographiques agissant sur le vieillissement des populations. Il intègre aussi les divers aspects sous lesquels la consommation de services doit être étudiée : les personnes, les cas, les jours. Enfin, le modèle rend compte des principaux facteurs qui peuvent par ailleurs influencer le fonctionnement des réseaux de services : les pratiques professionnelles, l'organisation des services, la performance des établissements.

La pertinence des principaux concepts mis en relation dans ce modèle explicatif a été démontrée algébriquement. Le modèle mesure, en totalité, le taux de jours consommés. Ainsi, pour chaque région, le taux de jours consommés (NJ/P) est égal :

- ◆ *au plan social*, à la somme des produits, membre à membre, des taux de jours consommés par chaque groupe d'âge multipliés par la proportion que représentent les effectifs de ce groupe d'âge sur la population totale (pr_i) :

$$\frac{NJ}{P} = \sum_{i=1}^{16} \left(\frac{NJ_i}{P_i} \times pr_i \right)$$

- ◆ *au plan sanitaire*, au produit de l'incidence des bénéficiaires (NB/P), des taux de réhospitalisation (TRH) et des durées moyennes de séjour (DUR) :

$$\frac{NJ}{P} = \frac{NB}{P} \times TRH \times DUR$$

LES RELATIONS ENTRE LES CONCEPTS

Les relations ont été établies, d'une part, sur la base des théories reconnues et admises dans les domaines scientifiques pertinents. D'autre part, les relations principales ont été vérifiées empiriquement sur la base des données primitives, notamment le rapport de causalité entre le niveau de non-emploi et la pauvreté, et la vitesse de changement de la structure

d'âge des populations. Par ailleurs, plusieurs autres relations se déduisent logiquement et nécessairement, compte tenu de la nature même des concepts impliqués.

Chapitre 4

La validation de l'explication

La vérification du modèle explicatif :
confrontation avec la réalité

« Les mathématiques rendent l'esprit juste en mathématiques, et les lettres le rendent juste en morale. Les mathématiques apprennent à faire des ponts, et la morale apprend à vivre. »

JOSEPH JOUBERT

« La vérité est comme Dieu : elle ne se montre pas à visage découvert. »

GOETHE

LA VÉRIFICATION DU MODÈLE EXPLICATIF : CONFRONTATION AVEC LA RÉALITÉ

Ce chapitre traite de la vérification du modèle théorique. Celui-ci y est soumis à deux niveaux de vérification comprenant :

- ◆ la vérification des éléments « sémantiques » du modèle, c'est-à-dire la vérification mathématique de la *mutuelle exclusivité* de la contribution de chaque variable explicative à l'explication de la variance de *chaque* variable dépendante. Un des résultats dérivé de cet exercice est la mesure de chacune des six variables exogènes pour chacune des quatorze régions.
- ◆ la vérification des relations « syntaxiques » entre les variables, c'est-à-dire la vérification mathématique de la *causalité des relations* existant entre les variables du modèle. Un des résultats dérivé de cet exercice départage l'influence réelle et respective du niveau de non-emploi et de la structure d'âge sur le niveau de morbidité des populations régionales.

LA VÉRIFICATION DES ÉLÉMENTS « SÉMANTIQUES »

Avant de décrire la marche à suivre, il importe de rappeler ici que ce sont les variables exogènes qui constituent les véritables facteurs explicatifs du modèle, conformément à l'énoncé théorique au troisième chapitre, sous la rubrique « Les variables exogènes... ».

LES ÉTAPES

Étape 1

Le positionnement des variables dans le modèle détermine l'ordre dans lequel chacune ajoute une contribution supplémentaire à l'explication de la variance des taux de jours (NJ/P). Ainsi, T1 occupe le premier rang de la séquence, PRA le deuxième (2), puis viennent dans l'ordre NB/P (3), TRH (4), DUR (5), jusqu'à NJ/P qui occupe le sixième rang (6).

Étape 2

En utilisant la régression multiple, on mesure d'abord, pour l'ensemble des régions, la valeur estimée ou prédite des taux de jours, en ajoutant successivement à chaque fois une nouvelle variable indépendante, selon l'ordre établi. Ainsi, le taux de jours est prédit (Y') par cinq équations successives. La dernière valeur ajoutée à cette suite correspond au taux de jours réel (Y).

$$Y'_1 = B_0 + (B_1 \times TI)$$

$$Y'_2 = B_0 + (B_1 \times TI) + (B_2 \times PRA)$$

$$Y'_3 = B_0 + (B_1 \times TI) + (B_2 \times PRA) + (B_3 \times NB/P)$$

$$Y'_4 = B_0 + (B_1 \times TI) + (B_2 \times PRA) + (B_3 \times NB/P) + (B_4 \times TRH)$$

$$Y'_5 = B_0 + (B_1 \times TI) + (B_2 \times PRA) + (B_3 \times NB/P) + (B_4 \times TRH) + (B_5 \times DUR)$$

$$Y = \text{le taux de jours réel} = NJ/P$$

Dans les cinq premières équations, « B » correspond à un coefficient de régression, c'est-à-dire au poids optimal qu'il faut attribuer à chaque variable indépendante – ce poids étant constant dans chaque équation – pour estimer la variable dépendante Y' . Chaque variable dépendante ou indépendante Y' , TI, PRA, etc. varie évidemment d'une région à l'autre, par définition.

Étape 3

On mesure ensuite la valeur prédictive de chaque équation : cette mesure est obtenue en élevant à la deuxième puissance (au carré), le coefficient de corrélation multiple existant entre la variable réelle Y et chaque estimation de celle-ci, Y' . La dernière valeur de cette suite correspond à la corrélation entre la variable réelle Y et elle-même, dont le résultat est nécessairement égal à 1 :

$$R^2_{yy'_1}, R^2_{yy'_2}, \dots, R^2_{yy}$$

Étape 4

Ensuite en mesurant les différences successives entre les R^2 , on peut mesurer la part d'explication de la variance totale des taux de jours, ajoutée globalement par chaque variable du modèle. Or, cet ajout successif de prédictivité est spécifiquement et nécessairement attribuable à chacune des variables exogènes qui influence une et une seule des variables du modèle. Le « résidu » des influences subies par chaque

variable exogène est nécessairement attribuable à l'influence exercée par les autres variables du modèle.

LES RÉSULTATS MESURÉS

Les six variables exogènes expliquent 100 % de la variance du modèle :

- ◆ le TI (égal à e_1) explique 82,7 % de la variance du taux de jours (NJ/P);
- ◆ l'historique du peuplement (e_2), qui influence PRA explique 7,3 % de la variance des taux de jours;
- ◆ les pratiques professionnelles différentielles (e_3), qui influencent NB/P, expliquent 3,5 % de la variance des taux de jours;
- ◆ l'organisation des services (e_4), qui influence TRH, explique 2,7 % de la variance des taux de jours;
- ◆ la performance des établissements (e_5), qui influence DUR, explique 3,7 % de la variance des taux de jours;
- ◆ et l'erreur du modèle (e_6), qui influence NJ/P, en explique 0,1 %.

UN COROLLAIRE : LA MESURE DES VARIABLES EXOGÈNES POUR CHAQUE RÉGION

En établissant pour chaque région les différences successives entre les valeurs prédites à partir des équations de régression (les Y'), on mesure six nouvelles variables exprimées en taux de jours de courte durée, et correspondant à chacune des six variables exogènes énumérées (tableau 4A). Par exemple, pour chaque région, le taux de jours consommés en courte durée imputable à « l'historique du peuplement » (e_2) est égal à :

$$Y'_2 - Y'_1 = e_2$$

De même, le taux de jours imputable à la performance des établissements (e_5) est égal à :

$$Y'_5 - Y'_4 = e_5$$

Ces nouvelles variables pourraient ensuite permettre d'établir, pour chaque région, le coût spécifique attribuable à chaque facteur explicatif. Par exemple, e_5 multiplié par la population de chaque région est égal au nombre de jours de courte durée attribuable à la performance des établissements des régions de traitement. Ce résultat lui-même multiplié par le coût moyen d'une journée d'hospitalisation en courte durée –

TABLEAU 4A
Distribution par région des jours de courte durée imputables à chaque variable exogène pour 1993-1994

RÉ- GION	TI DE 1987 (Note)	HISTORI- QUE DU PEUPLE- MENT	PRATIQUES DIFFÉREN- TIELLES	ORGANI- SATION DES SERVICES	PERFOR- MANCE DES ÉTABLIS- SEMENTS	ERREUR DU MODÈLE	TOTAL PRÉDIT = CONSOM- MATION RÉELLE
1	222 625	358	6 568	- 11 246	3 051	385	221 741
2	309 867	- 15 984	- 10 117	8 429	15 980	1 267	309 442
3	600 179	15 316	- 26 680	- 15 230	34 604	- 4 804	603 386
4	457 990	25 890	1 558	18 680	- 8 261	- 198	495 659
5	258 742	13 824	- 165	12 632	4 128	982	290 143
6	1 571 663	186 584	8 146	- 6 948	- 2 439	- 10 354	1 746 652
7	223 508	- 10 420	2 698	9 450	- 3 393	- 140	221 704
8	153 694	- 8 905	8 860	2 534	- 3 260	- 2 007	150 915
11	137 052	- 20	- 467	- 10	- 4 877	679	132 357
12	321 279	9 008	11 461	- 9 382	5 881	755	339 001
13	241 644	- 3 217	- 10 471	- 5 082	- 13 370	2 831	212 334
14	294 530	- 21 126	636	232	8 181	- 344	282 109
15	373 408	- 10 966	- 25 211	- 9 739	- 20 523	- 3 376	303 593
16	929 540	- 16 058	43 577	- 20 006	9 232	11 351	957 636

Note : incluant l'influence des facteurs biologiques (v. graphique 10).

TABLEAU 4B
Matrice de corrélation entre les variables exogènes

	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆
e ₁	1	0	0	0	0	0
e ₂	0	1	0	0	0	0
e ₃	0	0	1	0	0	0
e ₄	0	0	0	1	0	0
e ₅	0	0	0	0	1	0
e ₆	0	0	0	0	0	1

c'est-à-dire 450 \$ environ – permet de reproduire approximativement les valeurs financières correspondantes.

LA VÉRIFICATION DES CONCEPTS

Tel que le stipule le modèle théorique complet, il ne devrait exister, en principe, aucune relation entre les variables exogènes du modèle, celles-ci devant être mutuellement exclusives. Or, cette hypothèse a été confirmée globalement, et en détail, en dressant la matrice de corrélation entre les six variables exogènes prises deux à deux. Le tableau 4B présente ce résultat.

Ce résultat confirme mathématiquement le caractère consistant et mutuellement exclusif des variables impliquées dans le modèle. En effet, il corrobore l'incongruité d'une relation pouvant impliquer, par exemple, l'historique du peuplement et la performance des établissements. Il corrobore aussi la mutuelle exclusivité des autres concepts utilisés, tels les pratiques professionnelles, l'organisation des services, etc.

Un corollaire de ce résultat : si le modèle explique 100 % de la variance des taux de jours, il doit, par conséquent, expliquer de la même manière la totalité de la variance de chacune des variables indépendantes et explicatives de cette variable dépendante. Ce corollaire a effectivement été vérifié par la technique de la régression multiple (tableau 5). Par exemple, pour chaque région il n'existe aucun écart entre l'incidence de bénéficiaires (NB/P) et l'incidence estimée à l'aide de l'équation suivante où la variable e_3 , les pratiques professionnelles différentielles, est introduite comme variable explicative de NB/P.

$$Y' = B_0 + (B_1 \times TI) + (B_2 \times PRA) + (B_3 \times e_3)$$

$$Y' = \text{NB/P}$$

En procédant de la même manière pour chacune des autres variables du modèle à l'exception de la première ($e_1 = TI$), on peut, par exemple, mesurer la contribution de la performance des établissements des régions de traitement (e_5) à la variation interrégionale des durées de séjour des clientèles originaires de chaque région, etc.

LES ÉLÉMENTS DE LA PREUVE

La matrice de corrélation entre les taux de jours de courte durée des régions attribuables à chacune des variables exogènes du modèle (e_1 à e_6) est égale à une matrice « identité ». Ce qui implique qu'il n'existe aucune multicolinéarité entre les variables explicatives – les « e » – du niveau régional de consommation.

$$R^2_{6,e1} + R^2_{6,e2} + R^2_{6,e3} + R^2_{6,e4} + R^2_{6,e5} + R^2_{6,e6} = 1$$

TABLEAU 5

Résultats relatifs à la prédiction du niveau de consommation des régions
sous les contraintes imposées par le modèle explicatif

IDENTIFICATION DES VARIABLES EXOGENES	PART CUMULATIVE DE LA VARIANCE TOTALE EXPLIQUÉE PAR LES VARIABLES INDÉPENDANTES	CONTRIBUTION DE CHAQUE VARIABLE EXOGENE À LA VARIANCE TOTALE	MESURE DES INFLUENCES DIRECTES EXERCÉES PAR LES VARIABLES EXOGENES SUR LES VARIABLES
e ₁ : les facteurs explicatifs des écarts de niveau d'emploi entre régions	R ² _{6,1} = 0,82679	e ₁ : 0,82679	R _{1e1} = β _{1e1} = 1,00000
e ₂ : l'historique du peuplement des régions	R ² _{6,12} = 0,90030	e ₂ : 0,07351	R _{2e2} = β _{2e2} = 0,921672
e ₃ : la pratique professionnelle dans les régions de traitement	R ² _{6,123} = 0,93523	e ₃ : 0,03493	R _{3e3} = β _{3e3} = 0,376040
e ₄ : l'organisation des services *	R ² _{6,1234} = 0,96256	e ₄ : 0,02733	R _{4e4} = β _{4e4} = 0,849927
e ₅ : la performance des établissements dans les régions de traitement	R ² _{6,12345} = 0,99861	e ₅ : 0,03605	R _{5e5} = β _{5e5} = 0,516467
e ₆ : erreur attribuable à la méthode utilisée	R ² _{6,6} = 1,00000	e ₆ : 0,00139	R _{6e6} = β _{6e6} = 0,037259
TOTAL :		1,00000	

* par exemple : contamination de programmes et convalescence.

LA VÉRIFICATION DES ÉLÉMENTS « SYNTAXIQUES » : LA CAUSALITÉ DES RELATIONS EXISTANT ENTRE LES VARIABLES DU MODÈLE

L'OBJET DE LA DÉMONSTRATION

La section précédente portait sur la démonstration des éléments conceptuels du modèle, les éléments « sémantiques ». Cette section traite, de façon générale, des démonstrations touchant les *relations* entre

les variables mesurant ces concepts, les aspects « syntaxiques » du modèle. Il s'agit explicitement de la mesure des rapports de causalité existant entre ces variables.

De façon particulière, il s'agit aussi de démontrer le rapport de causalité mesuré par chacune des quinze corrélations liant chacune des six variables du modèle prises deux à deux.

LA PROBLÉMATIQUE

Par définition, une corrélation mesure l'importance numérique d'une relation entre deux variables sur une échelle, un intervalle compris entre - 1 et + 1. Or, le fait que cette relation soit de nature causale, plutôt que simplement associative, ne modifie pas la pertinence de cet énoncé général. Cependant, dans ce cas, cet énoncé implique nécessairement que la corrélation corresponde à la *mesure numérique d'une relation de causalité* existant entre deux variables.

Par ailleurs, en toute logique, lorsque la relation mesurée entre deux variables est une relation causale, on doit alors s'attendre à ce qu'elle soit *égale à la somme algébrique des influences* directes et indirectes exercées par une variable indépendante et explicative sur une variable dépendante et expliquée.

Or dans le modèle théorique représenté au diagramme 7, chaque « chemin » différent qu'on peut suivre, partant d'une variable indépendante pour atteindre, en respectant le sens des flèches, une variable dépendante, identifie une influence différente : cette influence est directe lorsqu'une seule flèche réunit les deux variables; elle est indirecte s'il faut suivre plusieurs flèches pour atteindre une variable, en partant d'une autre. C'est pour cette raison que la méthode utilisée pour tester la causalité des variables les unes sur les autres a pour nom « l'analyse de sentier » (*Path Analysis*).

Dans cette analyse, les variables exogènes « e » sont exclues d'emblée, car, tel que démontré, elles ne subissent aucune influence des autres variables indépendantes du modèle; de même, elles n'exercent d'influence que sur une, et une seule, variable dépendante.

LA DÉMARCHE MÉTHODOLOGIQUE EN QUATRE ÉTAPES

Partant de ces fondements (v. Kerlinger & Pedhazur, 1973, pages 305-333), la marche à suivre comporte les quatre étapes suivantes.

Première étape : mesurer quantitativement l'influence directe correspondant à chaque flèche du modèle

Pour ce faire, il faut d'abord rendre comparables les distributions de chaque variable sur une échelle numérique standardisée et commune à toutes les variables. Cette procédure permet de neutraliser, dans les données primitives, les écarts attribuables aux moyennes (\bar{X}) et à la dispersion des observations, l'écart-type (σ), lesquels sont différents d'une variable à l'autre.

Ainsi, pour chaque région (i) et chaque variable (j) :

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{\sigma_j}$$

Pour chacune des six variables du modèle, le score standard d'une région (Z_{ij}) est égal à la valeur obtenue sur la variable (X_{ij}) moins la moyenne de la variable (\bar{X}_j), ce résultat étant lui-même divisé par l'écart-type de la variable (σ_j). On procède de la sorte pour chacune des variables du modèle.

Il faut ensuite mesurer le poids qu'il faut donner à chaque score Z « indépendant » pour obtenir le score Z « dépendant ». Ce poids, égal à la mesure de l'influence directe recherchée, est symbolisé par la lettre grecque bêta « β ». Selon la notation conventionnelle, β_{yx} désigne la valeur de bêta lorsque la variable Y est influencée par la variable X; l'effet doit ici précéder la cause (diagramme 10).

Les variables *directement* explicatives d'une autre variable sont identifiées par les flèches du modèle théorique. Par exemple, la variable NB/P est expliquée par deux variables (TI et PRA) alors que la variable TRH est expliquée par trois variables (TI, PRA et NB/P), etc.

En utilisant, conformément à cette spécification, la technique de la régression multiple sur les distributions standardisées des variables (transformées en scores standard Z), on calcule les valeurs de chaque influence directe exercée, les unes sur les autres, par les variables du modèle, les β . Le mot « influence » est utilisé ici dans le même sens que la « pondération » qu'il faut attribuer aux scores Z de chaque variable explicative, pour qu'en mesurant la somme de ces produits, le résultat soit rigoureusement égal aux scores Z de la variable dépendante, et cela pour chacune des quatorze régions.

TABLEAU 6
Système d'équations des variables endogènes et exogènes,
pour chaque région

VARIABLE	ÉQUATION
TI	$Z_1 = e_1$
PRA	$Z_2 = \beta_{21}Z_1 + e_2$
NB/P	$Z_3^* = \beta_{31}Z_1 + e_3$ ou encore $Z_3 = \beta_{31}Z_1 + \beta_{32}Z_2 + e_3$
TRH	$Z_4 = \beta_{41}Z_1 + \beta_{42}Z_2 + \beta_{43}Z_3 + e_4$
DUR	$Z_5 = \beta_{51}Z_1 + \beta_{52}Z_2 + \beta_{53}Z_3 + \beta_{54}Z_4 + e_5$
NJ/P	$Z_6 = \beta_{63}Z_3 + \beta_{64}Z_4 + \beta_{65}Z_5 + e_6$

Note : La vérification du modèle porte sur l'interaction des variables endogènes (Z_1, Z_2, Z_3, Z_4, Z_5 et Z_6); l'influence des variables exogènes est donc considérée nulle pour cette vérification.

Or, compte tenu de la valeur absolue des β , il peut être pertinent d'effectuer certains ajustements au modèle explicatif complet (diagramme 10). Le modèle explicatif retenu (diagramme 11) correspond au résultat de ces ajustements. En effet, lorsque la valeur des β_1 est inférieure à 0,05, celle-ci peut être considérée comme négligeable, et la valeur de la pondération qu'elle implique, non significative.

Le diagramme 11 reproduit le résultat découlant de cette simplification attribuable à la valeur négligeable de l'influence de la variable PRA sur NB/P mesurée par β_{32} . Le tableau 7 reproduit les valeurs des β pour chacun des deux modèles (complet et retenu). Plus loin, sous la rubrique « Un corollaire surprenant... », seront mis en lumière les implications théoriques majeures découlant de cette observation particulière.

Seconde étape : développer le système d'équations qui permet de mesurer chaque score Z du modèle

Dans les équations suivantes, chaque score Z de chaque région est mesuré en tant que somme des influences subies par les autres variables du modèle qui l'expliquent et, par conséquent, le mesurent entièrement. Ainsi, dans l'équation suivante, Z_y – une variable dépendante et expliquée – est égale à la somme des variables indépendantes et explicatives, chacune pondérée par le poids de l'influence qu'elle exerce (β) sur la variable dépendante.

Pour toute variable dépendante du modèle :

$$Z_y = \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_m Z_m + \beta_{ey} Z_{ey}$$

Où m est égal au nombre de variables indépendantes et endogènes au modèle, qui influencent la variable dépendante (Z_y), et où Z_{ey} est égal à la variable exogène qui influence la variable Z_y .

DIAGRAMME 10
Le modèle explicatif complet

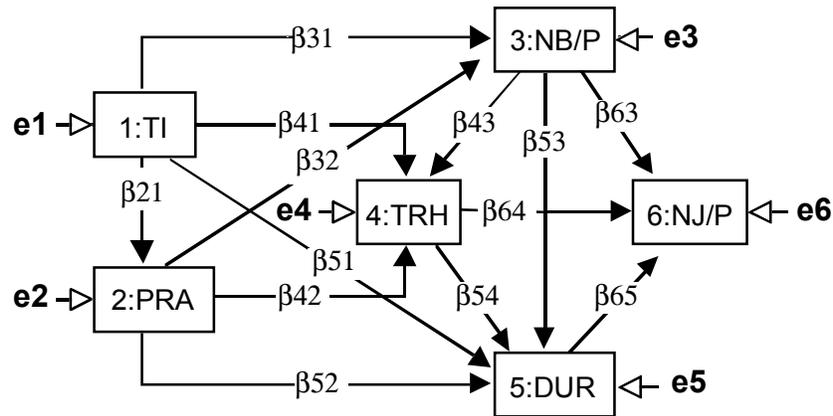
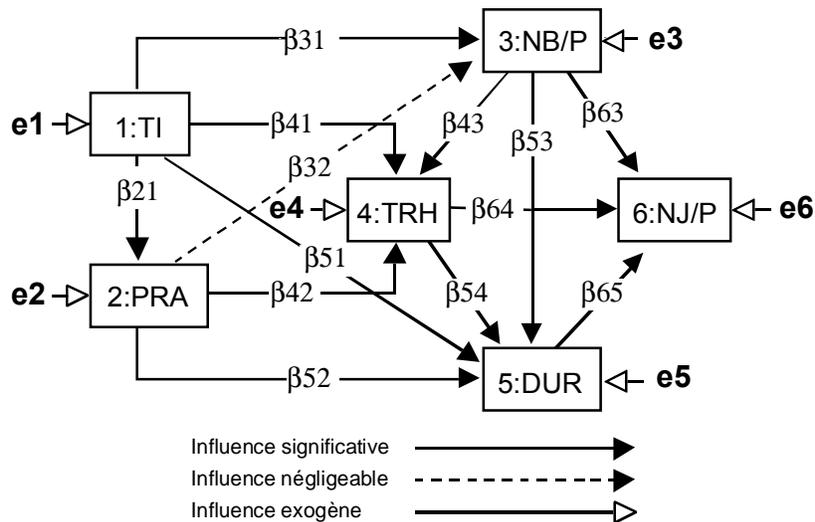


DIAGRAMME 11
Le modèle explicatif retenu



Cet énoncé découle nécessairement de la vérification précédente des concepts traitant de l'explication de chaque variable dépendante à 100 %. Les équations sont données au tableau 6.

Troisième étape : la mesure du rapport de causalité

Il s'agit ici de mesurer chaque corrélation de deux manières différentes : d'abord à partir du système d'équation obtenu, et ensuite selon les procédés usuels afin de pouvoir comparer les résultats entre eux.

La signification particulière de cette opération apparaît clairement en réalisant que la mesure des corrélations par les procédés usuels a pour objet de mesurer de simples relations associatives entre deux variables. Dans ce cas, les mêmes corrélations mesurées à l'aide du système d'équations correspondent à une situation particulière – *un rapport de causalité* – où chaque variable dépendante est égale à la somme des influences directes et indirectes exercées sur cette variable, par les autres variables du modèle. Ainsi, si la différence entre les deux résultats est nulle, on en déduira que les relations entre les variables sont des relations causales : causales, au sens où ces relations corroborent la relation d'équivalence entre ce qui existe *dans la réalité* – les données régionales entretenant entre elles des rapports associatifs – et *les énoncés du modèle théorique* qui sous-tend le modèle mathématique, impliquant l'existence d'un rapport de causalité entre les variables.

Pour les raisons invoquées précédemment, on neutralise d'abord l'influence de la variable exogène dans chaque équation du système, en postulant que, pour chaque région, $Z_{ey} = 0$. Restent alors seulement les équations susceptibles de mesurer la corrélation entre les variables prises deux à deux.

Par définition, la corrélation entre une variable indépendante et une variable dépendante quelconque « Y » est égale à l'équation :

$$R_{1y} = \frac{\sum Z_1 Z_y}{N} \text{ où } N \text{ est égal au nombre de régions (14).}$$

En utilisant cette formule de la corrélation, il faut ensuite substituer Z_y par l'équation correspondante du système d'équation décrit à l'étape précédente :

$$R_{1y} = \frac{\sum Z_1 (\beta_{y1} Z_1 + \beta_{y2} Z_2 + \dots + \beta_{ym} Z_m)}{N}$$

TABLEAU 7
Valeur des bêta du modèle explicatif selon la valeur du TI en 1987
et les données d'hospitalisation de 1993-1994

BÊTA	SELON LE MODÈLE COMPLET	SELON LE MODÈLE RETENU
β_{21}	0,387971	0,387971
β_{31}	0,916287	0,926299
β_{32}	0,025780	(retiré*)
β_{41}	0,696856	0,696856
β_{51}	0,313837	0,313837
β_{42}	- 0,276223	- 0,276223
β_{52}	0,812437	0,812437
β_{43}	- 0,144111	- 0,144111
β_{53}	- 0,234816	- 0,234816
β_{54}	0,060331	0,060331
β_{63}	0,690564	0,690564
β_{64}	0,177713	0,177713
β_{65}	0,405700	0,405700

* valeur en dessous du seuil minimal de 0,05.

où Z_1, Z_2, Z_m correspondent respectivement à chacune des m variables indépendantes influençant la variable dépendante Z_y , conformément aux spécifications du modèle.

En développant l'équation on trouve que :

$$R_{1y} = \beta_{y1} \frac{\sum Z_1^2}{N} + \beta_{y2} \frac{\sum Z_1 Z_2}{N} + \dots + \beta_{ym} \frac{\sum Z_1 Z_m}{N}$$

Sachant que :

$$\frac{\sum Z^2}{N} = 1 \text{ (par définition)}$$

on peut alors déduire que, conformément au modèle :

$$R_{1y} = \beta_{y1} + \beta_{y2} R_{12} + \dots + \beta_{ym} R_{1m}$$

TABLEAU 8A

Test de vérification du modèle explicatif complet selon la valeur du TI en 1987 et les données d'hospitalisation de 1993-1994

ÉQUATION	VALEUR CALCULÉE DE LA CORRÉLATION	ÉCART À PEARSON-BRAVAIS
$R_{12} = \beta_{21}$	0,3880	0
$R_{13} = \beta_{31} + \beta_{32}R_{12}$	0,9263	0
$R_{14} = \beta_{41} + \beta_{42}R_{12} + \beta_{43}R_{13}$	0,4562	0
$R_{15} = \beta_{51} + \beta_{52}R_{12} + \beta_{53}R_{13} + \beta_{54}R_{14}$	0,4391	0
$R_{16} = \beta_{63}R_{13} + \beta_{64}R_{14} + \beta_{65}R_{15}$	0,9093	0,0104
$R_{23} = \beta_{32} + \beta_{31}R_{12}$	0,3594	0
$R_{24} = \beta_{41}R_{12} + \beta_{42} + \beta_{43}R_{23}$	- 0,0608	0
$R_{25} = \beta_{51}R_{12} + \beta_{52} + \beta_{53}R_{23} + \beta_{54}R_{24}$	0,8410	0
$R_{26} = \beta_{63}R_{23} + \beta_{64}R_{24} + \beta_{65}R_{25}$	0,6027	0,0090
$R_{34} = \beta_{41}R_{13} + \beta_{42}R_{23} + \beta_{43}$	0,3961	0
$R_{35} = \beta_{51}R_{13} + \beta_{52}R_{23} + \beta_{53} + \beta_{54}R_{34}$	0,3895	0
$R_{36} = \beta_{63} + \beta_{64}R_{34} + \beta_{65}R_{35}$	0,9190	0
$R_{45} = \beta_{51}R_{14} + \beta_{52}R_{24} + \beta_{53}R_{34} + \beta_{54}$	0,0611	0
$R_{46} = \beta_{63}R_{34} + \beta_{64} + \beta_{65}R_{45}$	0,4760	0
$R_{56} = \beta_{63}R_{35} + \beta_{64}R_{45} + \beta_{65}$	0,6856	0

Note : le premier écart s'explique par une influence récurrente de TRH sur NB/P, attribuable au fait que, parmi les clientèles hospitalisées à plusieurs reprises au cours de la même année, certaines migrent d'une région à l'autre étant ainsi comptées deux fois dans la variable NB/P. L'autre écart, l'influence de NB/P sur DUR, s'explique indirectement par ce premier effet, car NB/P est déjà influencé par TRH. Les valeurs des bêta des équations sont données au tableau 7.

Connaissant déjà les valeurs des β , identifiées à la première étape et disposant de la matrice de corrélation entre toutes les variables du modèle calculée selon les méthodes usuelles (*Pearson-Bravais*), il reste à vérifier si la valeur de corrélation mesurée par les méthodes usuelles et celle qui correspond à la résolution de la dernière équation sont égales.

En répétant le même exercice pour chacune des quinze corrélations du modèle, on peut ainsi vérifier si chacune d'elles est égale à la somme des influences directes et indirectes qu'exercent les variables indépendantes sur chaque variable dépendante. Les tableaux 8A et 8B reproduisent, à la troisième colonne, la valeur des écarts enregistrés.

TABLEAU 8B
 Test de vérification du modèle explicatif retenu selon la valeur du TI
 en 1987 et les données d'hospitalisation de 1993-1994

ÉQUATION	VALEUR CALCULÉE DE LA CORRÉLATION	ÉCART À PEARSON- BRAVAIS
$R_{12} = \beta_{21}$	0,3880	0
$R_{13} = \beta_{31} + \beta_{32}R_{12}$	0,9263	0
$R_{14} = \beta_{41} + \beta_{42}R_{12} + \beta_{43}R_{13}$	0,4562	0
$R_{15} = \beta_{51} + \beta_{52}R_{12} + \beta_{53}R_{13} + \beta_{54}R_{14}$	0,4391	0
$R_{16} = \beta_{63}R_{13} + \beta_{64}R_{14} + \beta_{65}R_{15}$	0,9093	0,0104
$R_{23} = \beta_{32} R_{12}$	0,3813	0,0219
$R_{24} = \beta_{41}R_{12} + \beta_{42} + \beta_{43}R_{23}$	- 0,0608	0
$R_{25} = \beta_{51}R_{12} + \beta_{52} + \beta_{53}R_{23} + \beta_{54}R_{24}$	0,8410	0
$R_{26} = \beta_{63}R_{23} + \beta_{64}R_{24} + \beta_{65}R_{25}$	0,6027	0,0090
$R_{34} = \beta_{41}R_{13} + \beta_{42}R_{23} + \beta_{43}$	0,3961	0
$R_{35} = \beta_{51}R_{13} + \beta_{52}R_{23} + \beta_{53} + \beta_{54}R_{34}$	0,3895	0
$R_{36} = \beta_{63} + \beta_{64}R_{34} + \beta_{65}R_{35}$	0,9190	0
$R_{45} = \beta_{51}R_{14} + \beta_{52}R_{24} + \beta_{53}R_{34} + \beta_{54}$	0,0611	0
$R_{46} = \beta_{63}R_{34} + \beta_{64} + \beta_{65}R_{45}$	0,4760	0
$R_{56} = \beta_{63}R_{35} + \beta_{64}R_{45} + \beta_{65}$	0,6856	0

Note : l'écart indiqué en **gras** provient du retrait de l'influence négligeable qu'exerce PRA sur NB/P.
 Les valeurs des bêta des équations sont données au tableau 7. Voir aussi la note du tableau 8A.

Les β sont mesurés par la technique de la régression multiple appliquée aux distributions standardisées des variables, en conformité avec les influences directes décrites dans le modèle explicatif. Les corrélations sont mesurées par les méthodes de calcul usuelles, en particulier celle de *Pearson-Bravais*. Les vérifications des corrélations correspondent aux valeurs obtenues conformément aux équations du modèle explicatif (15 en tout).

Quatrième étape : la vérification de la théorie

En vérifiant de cette manière chacune des quinze corrélations, on constatera alors que chacune peut être exprimée et mesurée par une somme de produits de β . Il faut pour ce faire substituer, dans les équations, les valeurs de corrélation (R) par les valeurs de β correspondantes, et alors chaque produit correspond au poids numérique d'un « chemin », c'est-à-dire au poids numérique d'une influence indirecte et/ou d'une ou plusieurs influences directes dont on peut suivre le tracé sur le modèle théorique et graphique.

UN EXEMPLE DE VÉRIFICATION : L'INCIDENCE DES BÉNÉFICIAIRES

Dans le modèle théorique, l'incidence de bénéficiaires (NB/P) est influencée par le TI et par PRA, cette dernière variable étant, elle-même, influencée par le TI. Or, le coefficient de corrélation entre TI et NB/P mesuré par les procédés usuels est égal à 92,63 %, alors que la valeur de chaque β mesurant les influences directes entre les trois variables est connue. Dans l'exemple illustré au diagramme 8, les valeurs des corrélations et des β utiles au calcul sont reproduites au tableau 8A.

Dans ce modèle, on peut déduire l'influence qu'exercent les variables les unes sur les autres. Ainsi, pour chaque région :

$$Z_2 = \beta_{21}Z_1 + \beta_{2e2}Z_{e2}$$

$$Z_3 = \beta_{31}Z_1 + \beta_{32}Z_2 + \beta_{3e3}Z_{e3}$$

Où Z_{e2} et Z_{e3} sont postulés égaux à zéro.

Pour convertir la valeur des corrélations (R) en mesure des influences (β), une première équivalence est ainsi démontrée :

$$R_{12} = \frac{\sum Z_1 Z_2}{N} \text{ (par définition)}$$

$$R_{12} = \frac{\sum Z_1 (\beta_{21} Z_1)}{N}$$

$$R_{12} = \beta_{21} \times \frac{\sum Z_1^2}{N}$$

$$R_{12} = \beta_{21} \times 1$$

$$R_{12} = \beta_{21}$$

Sachant, par le développement de l'équation précédente, que $R_{12} = \beta_{21}$, alors :

$$R_{13} = \frac{\sum Z_1 Z_3}{N} \text{ (par définition)}$$

$$R_{13} = \frac{\sum Z_1 (\beta_{31} Z_1 + \beta_{32} Z_2)}{N}$$

$$R_{13} = \beta_{31} \frac{\sum Z_1^2}{N} + \beta_{32} \frac{\sum Z_1 Z_2}{N}$$

$$R_{13} = \beta_{31} + (\beta_{32} \times R_{12})$$

$$R_{13} = \beta_{31} + \beta_{32} \beta_{21}$$

Enfin, en remplaçant les bêta par leur valeur numérique, on trouve que :

$$0,9263 = 0,916287 + (0,025780 \times 0,387971)$$

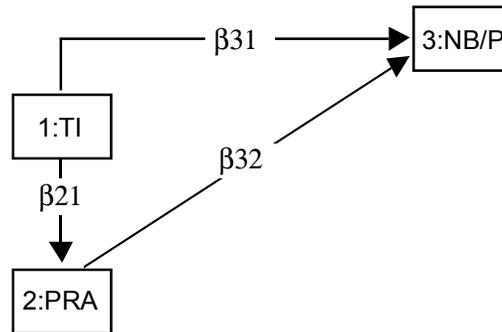
soit la somme d'une influence directe (β_{31}) et d'une influence indirecte ($\beta_{21} \times \beta_{32}$) exercée par TI sur NB/P; la valeur numérique des influences mesure le poids de chaque « chemin » exprimé par le modèle graphique de cet exemple (diagramme 12 et tableau 6).

UN COROLLAIRE SURPRENANT : L'APPARENTE INDÉPENDANCE ENTRE LA STRUCTURE D'ÂGE ET LA MORBIDITÉ

Bien que cet exemple ne représente qu'un élément de la démonstration, il amène néanmoins un des résultats les plus importants de cette étude. Élaboré à partir de données réelles, il démontre que la structure d'âge des populations mesurée par PRA exerce une influence négligeable (tableau 7, β_{32}) sur l'incidence des bénéficiaires par population (NB/P). Il faut en déduire que, des deux variables explicatives (TI et PRA), *seul le TI* exerce une influence significative sur le niveau de morbidité à l'échelle des populations régionales : en effet, la corrélation de 92 % entre TI et NB/P repose en quasi-totalité sur l'influence directe du TI sur NB/P (tableau 7, β_{31}), et non sur la contribution de la structure d'âge (PRA).

À première vue, ce résultat est surprenant. Contre toute attente, il impliquerait que le niveau de morbidité des populations régionales n'entretient aucun rapport de dépendance, ou presque, avec la structure d'âge de ces populations. Or, *c'est précisément ce qu'il faut éviter ici de déduire.*

DIAGRAMME 12
Les influences directes entre trois variables connues.



En effet, ce résultat apparemment paradoxal est tout à fait cohérent eu égard aux énoncés et démonstrations précédentes : d'abord parce que la population inoccupée, soit le numérateur du TI, inclut, par définition, l'ensemble des personnes âgées à la retraite qui forment le numérateur de la variable PRA. Ceci implique qu'indépendamment des rapports socio-économiques dont il rend compte, le TI augmente aussi en fonction de l'accroissement de la proportion de personnes âgées dans les populations.

Il faut ajouter également que le TI intègre déjà les aspects susceptibles de refléter le vieillissement relatif des populations par le fait qu'il explique la vitesse de changement de leur structure d'âge respective. Enfin, comme il a été démontré déjà, dans le contexte particulier de ce modèle causal, ce n'est pas la variable PRA qui influence le niveau de morbidité des populations, mais bien la variable exogène e_2 , soit l'état historique de leur peuplement, correspondant à l'indicateur de la structure d'âge des populations, *décapé* de l'influence déjà assumée par le TI à l'égard de la vitesse de changement de cette structure d'âge.

Par ailleurs, si ces mises au point suffisent pour expliquer méthodologiquement ce résultat de l'analyse de sentier apparemment peu conforme à la perception courante de la réalité, il n'en demeure pas moins qu'elles recèlent implicitement deux éléments fondamentaux de la problématique de la morbidité. Ces éléments ajoutent à la compréhension des rapports entre les phénomènes sociaux et les phénomènes biologiques, tels la santé des personnes; ils permettent en même temps de cerner les manifestations particulières de ces rapports en tant que rapports de causalité. Ces éléments de problématique touchent respectivement :

- ◆ l'emboîtement des phénomènes d'ordre biologique dans les caractéristiques des populations, et le fait qu'ils soient fondus à d'autres phénomènes de nature spécifiquement sociale;
- ◆ la relation d'antériorité – par opposition à simultanéité – des causes sociales sur leurs effets biologiques, conformément aux énoncés précédents, et mettant en cause la théorie des maladies multifactorielles.

Touchant ce dernier élément, si le vieillissement des populations contribuait véritablement à accroître leur niveau de morbidité, on devrait, compte tenu des mises au point énoncées, observer une croissance simultanée de la corrélation entre le TI et le niveau de consommation de services, à mesure que l'année de prise du TI s'approche de l'année de consommation de services mesurée. Par exemple, le TI de 1993 serait supérieur à celui de 1987 pour rendre compte de la consommation régionale de l'année 1993, car le TI de 1993 *comprend* les personnes âgées de cette année de consommation. Or, tel que démontré au chapitre 1, c'est précisément l'inverse qui se vérifie dans les faits : plus l'année de prise du TI est antérieure à l'année de consommation étudiée, plus la corrélation entre les deux variables est élevée. Ce constat impliquerait, à titre d'hypothèse, contrairement à une opinion généralement admise, que le vieillissement des populations *ne constitue pas le facteur principal* de l'accroissement de leur morbidité. À ce sujet, la rubrique ci-après, intitulée « Une clarification sur le rapport... », traite des faits qui confirment cette hypothèse. Ces faits démontrent que ce résultat est attribuable *au degré d'exposition* de toutes les strates d'âge de chaque population aux déclencheurs des maladies multifactorielles, incluant les autres influences associées au vieillissement.

LES RÉSULTATS DE L'ANALYSE DE SENTIER

Les critères de vérification du modèle

L'analyse de sentier comme méthode repose sur un postulat voulant que les relations entre les variables expriment des rapports de causalité. Or, ces rapports ne peuvent être *inférés* par une méthode mathématique : cette méthode est donc utilisée *pour vérifier l'application d'une théorie explicative dans la réalité*. Dans ce cadre, la théorie est elle-même posée comme hypothèse à vérifier à l'aide de cette méthode.

L'approche suivie dans cet ouvrage respecte cette exigence, la théorie explicative du niveau de consommation de services ayant été développée au chapitre précédent. Dans ce contexte, les inégalités entre

les résultats des deux mesures de corrélation entre chaque variable (tableaux 8A et 8B) permettent de détecter :

- ◆ une erreur d'interprétation théorique, par exemple, l'omission d'une relation, d'une influence entre deux variables qu'il serait nécessaire de considérer; ou encore une influence du modèle théorique orientée en sens inverse de ce qui se vérifie dans la réalité;
- ◆ une incohérence entre les faits de la réalité et les autres postulats de l'analyse de sentier. Par exemple, concernant le postulat « d'additivité », le fait que certaines variables dépendantes soient égales au produit de variables indépendantes plutôt qu'à la somme de leurs influences. Concernant le postulat de « non-récurtivité », le fait que deux variables s'influencent mutuellement et réciproquement, étant à la fois indépendante et dépendante l'une de l'autre.
- ◆ enfin, l'effet d'une décision arrêtée délibérément dans le but de *simplifier* le modèle : lorsque la valeur numérique d'une influence directe (un bêta) est jugée négligeable ($< 0,05$). L'élimination de la relation entre PRA et NB/P en est un exemple approprié.

Le résultat général

Sur les quinze corrélations à vérifier à l'aide du modèle, treize coïncident à 100 % avec les résultats attendus. Deux corrélations ne coïncident pas : l'écart entre le résultat attendu et le résultat réel y est dans les deux cas inférieur à 1 %, donc négligeable.

Ces écarts sont essentiellement attribuables à un effet de récursivité d'une variable sur une autre. Il s'agit, en l'occurrence, d'une influence récursive de TRH sur NB/P attribuable au fait suivant : parmi les personnes hospitalisées à plusieurs reprises au cours d'une même année, certaines partent d'une région vers une nouvelle région de résidence. Elles sont ainsi comptées deux fois dans la variable NB/P. L'autre écart – soit l'influence de NB/P sur DUR – s'explique indirectement par ce premier effet car NB/P est déjà influencé par TRH.

Compte tenu du résultat obtenu, le modèle explicatif du niveau de consommation régional s'avère triplement fondé :

- ◆ *dans les faits de la réalité* manifestés dans des relations généralisées à l'échelle des populations régionales : cumul par les mêmes populations des caractéristiques systématiques de

morbidity vérifiables pour tous les groupes d'âge, la concomitance de ces caractéristiques avec le TI croissant dans le temps, etc.;

- ◆ *dans l'univers des conceptions et des théories explicatives reconnues* du domaine scientifique touchant les déterminants de la consommation et les influences qu'ils entretiennent entre eux, tel que présenté au chapitre 3;
- ◆ enfin, *dans l'univers des modèles formels* pertinents à la vérification de l'équivalence entre les conceptions théoriques et les faits de la réalité, tel que traité dans ce chapitre.

Objectivement, la pauvreté influence la morbidité, et celle-ci influence à son tour l'expression même de la maladie que mesure le niveau de consommation des populations considérées. Toutefois, il était nécessaire que la consommation soit produite et mesurée dans un régime de soins « universels ». Il fallait que soit neutralisé le biais attribuable à l'influence de la pauvreté sur la consommation. En effet, indépendamment de la morbidité, comme cela a été constaté au Québec à d'autres époques et dans d'autres régimes de santé, la pauvreté peut engendrer une sous-consommation de services de santé.

Les principaux résultats spécifiques

Mise à part l'influence mineure exercée par les pratiques professionnelles différentielles (e_3), le niveau de non-emploi (TI) explique à lui seul la quasi-totalité de la variable « incidence de bénéficiaires » (NB/P). Dans ce contexte, ce déterminant est l'unique responsable du niveau de morbidité des populations³⁸, car il mesure la propension différente des populations à générer des personnes nécessitant une hospitalisation.

La structure d'âge des populations est le principal déterminant qui fait varier le volume de services à dispenser aux personnes hospitalisées entre les régions, par le truchement des durées de séjour (DUR) et à un degré moindre, des taux de réhospitalisation (TRH).

Comparée aux taux de réhospitalisation (TRH) et aux durées de séjour (DUR), l'incidence des bénéficiaires (NB/P) est, de loin, le

³⁸ Cet énoncé inclut et ne nie pas la propension plus grande des personnes âgées à consommer des services pour des raisons physiologiques : en effet, cette caractéristique est prise en compte dans le profil de consommation, par groupes d'âge, des populations. Or, il a été vu que le TI explique la hauteur de ce profil pour tous les groupes d'âge incluant les personnes âgées. C'est pourquoi, paradoxalement, la proportion de personnes âgées (PRA) exerce une influence négligeable sur NB/P.

déterminant qui exerce *la plus forte influence* sur le niveau global de consommation de services des populations régionales (NJP).

La somme des influences directes et indirectes exercées par le Tl sur les taux de jours totalise 89,89 % alors que la corrélation entre les deux variables égale 90,93 %. Il s'agit ici d'une des deux corrélations sur quinze qui ne se vérifie pas à 100 %; ces écarts mineurs ont été expliqués à la rubrique intitulée : « Le résultat général ».

Enfin, les facteurs spécifiquement sociaux – Tl et PRA – expliquent 90 % de la variance du niveau de consommation des populations régionales. Tous les autres déterminants réunis (les facteurs organisationnels, les erreurs de modèle, etc.) comptent pour 10 % seulement.

UNE CLARIFICATION SUR LE RAPPORT DE CAUSALITÉ ENTRE LA STRUCTURE D'ÂGE ET LA MORBIDITÉ DES POPULATIONS RÉGIONALES

Les mises au point présentées précédemment sur le sujet sont-elles suffisantes? Comment concevoir que la structure d'âge, en particulier si elle est mesurée par la proportion de personnes âgées, n'entretiennent aucun rapport explicatif avec le nombre relatif de bénéficiaires produit par chaque population régionale, l'expression même du concept de « morbidité »?

En effet, ne constate-t-on pas en courte durée physique, *dans toutes les régions*, que le nombre de bénéficiaires parmi les 65 ans et plus est, en moyenne, environ $3\frac{1}{4}$ fois plus élevé que chez les moins de 65 ans, pour une même quantité de personnes appartenant à chacun de ces groupes d'âge? Ceci impliquerait, de façon presque obligée, que plus une population régionale comporte une forte proportion de personnes âgées, plus cette population sera forte consommatrice de services de santé. Plus encore, cette réalité renvoie aux constatations biologiques les plus évidentes relatives au vieillissement organique selon lesquelles les représentants de l'espèce humaine deviennent progressivement, avec l'âge, de plus en plus vulnérables et affectés par la maladie!

Une clarification au plan théorique

Avant de procéder aux clarifications d'ordre méthodologique, certaines clarifications théoriques s'imposent d'emblée, en raison de la nature même des phénomènes sociaux, par opposition à ceux qui sont de nature biologique et individuelle.

À cet égard, il a été clairement spécifié, à propos des profils de consommation par groupes d'âge des populations régionales, que ceux-ci sont *quasi identiques* dans leur forme d'une région à l'autre; par contre,

leur « hauteur » varie selon la région. Cette distinction entre la *forme* du profil et sa *hauteur* permet de départager, dans l'expression de la consommation, deux types de déterminants :

- ◆ la *physiologie des êtres humains*, exprimée par la forme du profil, celle-ci demeurant invariante quelle que soit la population étudiée;
- ◆ les *facteurs sociaux*, exprimés par la hauteur du profil (intensité), variant selon les populations.

La simple observation des profils de consommation par groupe d'âge permet donc de constater, à l'évidence, que si la morbidité des populations est envisagée uniquement et même exclusivement comme une caractéristique physiologique des sujets qui en font partie, la structure d'âge sera vue comme le facteur essentiel pour expliquer pourquoi certaines populations consomment plus que d'autres. En revanche, on ne parviendra jamais à expliquer, sous cet angle d'approche, pourquoi les représentants de chacun des groupes d'âge – y compris les personnes âgées – de certaines populations sont plus malades que les représentants de chacun des groupes d'âge d'autres populations. En effet, les facteurs propres à en rendre compte appartiennent à l'ordre des phénomènes sociaux, variables d'une population à l'autre, cette variabilité affectant d'un seul tenant tous les groupes d'âge des différentes populations impliquées.

Dans cette perspective, se demander si la structure d'âge influence le niveau de consommation des populations régionales peut aboutir à deux réponses apparemment contradictoires :

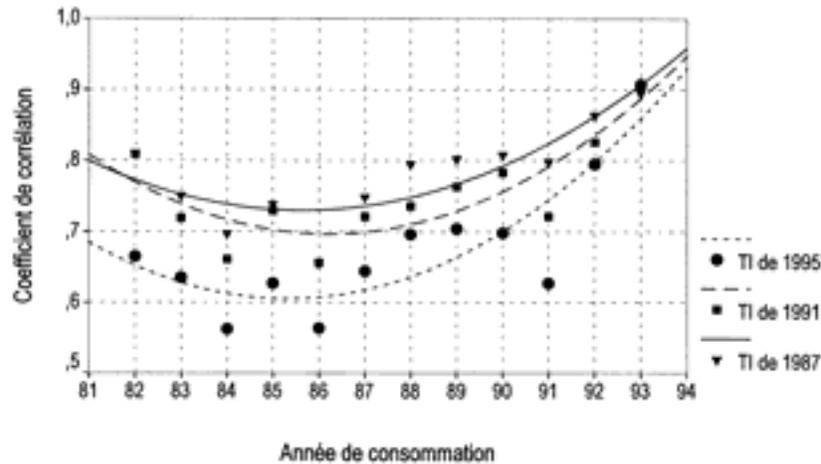
- ◆ *oui*, si la morbidité des populations est considérée seulement sous l'aspect physiologique;
- ◆ *non*, si la morbidité des populations est considérée seulement sous l'aspect social.

Ces réponses, apparemment opposées, découlent du fait que les phénomènes qui caractérisent en propre les populations rendent compte – par définition et à l'évidence – *de façon indissociée* des phénomènes spécifiquement sociaux *et* de l'ensemble des phénomènes d'ordre biologique et physiologique qui caractérisent les individus appartenant à ces populations. L'inverse n'est pas vrai³⁹. La sommation des phénomènes

³⁹ Une cellule ne se réduit pas à la juxtaposition informe de molécules; un être humain ne se réduit pas non plus à la somme de ses cellules; de même, à une échelle supérieure, une collectivité n'est pas que la somme des individus qui la composent.

GRAPHIQUE 14

Corrélations de trois TI avec les taux annuels de jours régionaux des 0-14 ans pour dix groupes diagnostiques non spécifiques à l'enfance



* Groupes diagnostiques retenus : 4, 6, 7, 8, 9, 10, 12, 13, 16 et 17.

individuels ne rend pas compte d'un certain nombre de faits qui caractérisent en propre les populations. Ces faits transcendent les individus qui composent les populations et affectent ces mêmes individus de diverses façons et à divers degrés.

Ainsi, le résultat relatif aux facteurs explicatifs de la morbidité doit être énoncé comme suit :

Compte tenu des facteurs physiologiques propres à l'espèce humaine et qui, a priori, rendent les personnes plus âgées davantage vulnérables à la maladie que les représentants des autres groupes d'âge, le niveau de non-emploi mesuré par le taux d'inoccupation apparaît constituer le seul facteur social susceptible d'interférer sur le niveau de morbidité des populations régionales, incluant les facteurs d'ordre physiologique.

Cette proposition est corroborée empiriquement par le fait que le TI corrèle avec le niveau de consommation régional de chacun des seize groupes d'âge, subdivisant chaque population, et ce, considérés indépendamment les uns des autres. Pour tout regroupement d'individus en

Néanmoins, chaque niveau supérieur englobe, a priori et de façon indissociée, tous les phénomènes qui caractérisent, en propre, les niveaux inférieurs.

populations, le fait est que les phénomènes spécifiquement attribuables à la physiologie de l'espèce humaine apparaîtront, a priori et nécessairement, entremêlés et fondus aux phénomènes de nature spécifiquement sociale.

Le décalage chronologique entre les causes sociales et leurs effets biologiques : un rapport de non-simultanéité

Pour les raisons évoquées précédemment, l'évolution de la structure d'âge des populations ne peut être mise en cause pour expliquer les résultats de l'analyse de sentier. Pourquoi alors la corrélation entre le TI et le niveau de consommation des populations augmente-t-elle à mesure qu'on recule l'année de mesure du TI? À cette interrogation, il existe un nombre limité de réponses à la fois possibles et plausibles. On pourrait d'une part supposer que la réduction progressive des durées de séjour en courte durée physique pour les personnes âgées de 65 ans et plus masque la contribution du vieillissement de la structure d'âge à la morbidité. On pourrait, d'autre part, à une échelle plus fondamentale, y voir la principale manifestation de la *surexposition continue* de certaines populations aux déclencheurs des maladies multifactorielles dont le manque d'emplois est le premier déterminant.

Pour le vérifier, nous avons reproduit, pour de grands groupes d'âge non influencés par la présence des personnes âgées de 65 ans et plus, l'analyse réalisée au deuxième chapitre sur l'accroissement des taux de corrélation entre le TI et le niveau de consommation. Cet accroissement attestait de la capacité de l'indicateur de prédire la *consommation ultérieure* des populations régionales. Pour deux grands groupes d'âge, les 0-14 ans et les 15-39 ans, nous avons ici analysé la progression des coefficients de corrélation entre le TI de plusieurs années (comprises entre 1987 et 1995) avec le niveau de consommation des douze années (de 1982 à 1993). Ceci a porté sur un ensemble de dix catégories de diagnostics dont le niveau de corrélation avec le TI de 1987 était, dans chaque cas, supérieur à 70 %. L'analyse a été reproduite sur la base des *cas et des jours* de courte durée, de façon à isoler les facteurs attribuables aux durées de séjour.

Le résultat de cette analyse montre que la prédictivité du TI pour chacun de ces deux grands groupes d'âge de chaque population se vérifie, comme cela était le cas pour les niveaux globaux de consommation des populations, tel qu'énoncé au deuxième chapitre. Le graphique 14 illustre le résultat pour les 0-14 ans.

Ce résultat implique qu'au plan social, l'évolution de la structure d'âge des populations *ne peut être tenue* comme un facteur déterminant

agissant de façon indépendante sur l'évolution du niveau de morbidité des populations régionales. En effet, la corrélation entre le TI et la consommation augmente en fonction de l'ancienneté de l'année de mesure du TI pour chacun des grands groupes d'âge comme pour les populations considérées globalement⁴⁰, mais non en fonction de l'accroissement du nombre des personnes âgées parmi la population inoccupée, imputable au vieillissement.

Compte tenu de ce résultat, comment s'expliquer alors les résultats de l'analyse de sentier montrant l'existence d'un lien explicatif beaucoup plus évident entre PRA et deux autres variables, les taux de réhospitalisation (TRH) et les durées de séjour (DUR)? Comment ces variables peuvent-elles réagir à l'historique du peuplement (e_2)? Nous n'avons pas investigué, de façon approfondie, cette piste d'analyse. Néanmoins, plusieurs hypothèses plausibles se dégagent des résultats de l'analyse de sentier. Ces hypothèses tiennent à certaines caractéristiques de la consommation de soins de courte durée par les personnes âgées en particulier, des caractéristiques peu susceptibles de varier dans le temps compte tenu de leur caractère permanent. L'effet sur les durées de séjour pourrait notamment s'expliquer par des séjours naturellement plus longs chez les personnes âgées que chez les personnes d'autres groupes d'âge. Une exposition plus grande à la mortalité et/ou la prise en charge par un autre réseau, celui de la longue durée, pourrait aussi expliquer l'influence négative de PRA sur les taux de réhospitalisation.

En conclusion, le taux d'inoccupation semble donc constituer le seul facteur explicatif du niveau de morbidité des populations régionales. Compte tenu des postulats de l'analyse de sentier, il est pertinent d'affirmer que le niveau de non-emploi mesuré par le TI explique causalement et en première instance environ 90 % du niveau de consommation des populations régionales – mesuré par la corrélation entre cette variable et les taux de jours – et 92 % de leur niveau de

⁴⁰ Corroborant autrement la pertinence de ce résultat au plan théorique, l'analyse de la progression de la corrélation entre le taux d'inoccupation et le niveau de consommation pour des *troubles transitoires en santé mentale* démontre, elle, un rapport de simultanéité. Or, il s'agit là d'une cause d'hospitalisation où on peut difficilement imaginer l'existence d'un décalage entre la cause extérieure et l'effet induit. Comment pourrait-on, par exemple, concevoir qu'une dépression nerveuse se manifeste chez une personne plusieurs années après un important événement traumatique déclencheur de détresse psychologique? Néanmoins, selon les psychiatres mis à contribution dans nos travaux, il ne fait pas de doute que la vulnérabilité biologique joue un rôle important dans le déclenchement de ce type d'affections dont l'incidence reste fortement corrélée avec le TI.

morbidité – mesuré par la corrélation entre le TI et le nombre de bénéficiaires par population.

LES PREMIÈRES CONCLUSIONS

Le niveau de consommation de services constitue une des multiples manifestations du niveau de besoin des populations régionales, reflété de la même manière par leur niveau de pauvreté, la vitesse de leur désintégration démographique, le tout expliqué par les écarts de disponibilité d'emploi.

Les populations qui se distinguent par leur haut niveau de consommation de services hospitaliers sont des populations plus malades que les autres, car elles réunissent une plus grande proportion d'individus de tous âges encourant un risque de mort ou d'atteinte grave si elles ne pouvaient recourir aux services, conformément au concept même de « besoin ».

Dans ce contexte, les hypothèses explicatives associant l'essentiel des différences de niveaux de consommation entre les régions à des dysfonctions organisationnelles du réseau de services et aux écarts de ressources disponibles entre les régions apparaissent non fondées dans la réalité : l'ensemble des « dysfonctions » éventuelles ne peut contribuer qu'à 10 % à peine de l'explication du niveau régional de consommation de services⁴¹ : les dysfonctions pouvant toucher les pratiques professionnelles médicales, l'organisation des services et la performance des établissements.

Par contre, l'hypothèse évoquée par certains, voulant que les écarts de consommation soient l'effet des ressources disponibles, reste à traiter. Or, l'importance du sujet exige qu'il soit examiné de façon particulière. Les prochains chapitres en traiteront.

⁴¹ Il faut noter qu'il s'agit ici d'une hypothèse selon laquelle les écarts identifiés pourraient être véritablement interprétés comme des dysfonctions réelles et non comme des « comportements » justifiables.

D'AUTRES RÉSULTATS D'ÉTUDE NON EXPOSÉS

La première moitié de l'ouvrage traite des éléments principaux qui vérifient le lien entre le phénomène de la consommation de services hospitaliers en santé physique, les indicateurs qui le prédisent et les facteurs qui le déterminent ou l'influencent. D'autres matières qui n'apparaissent pas ici ont aussi été examinées. Parmi les aspects investigués mais non présentés, il faut mentionner principalement :

- ◆ la mesure des besoins à l'échelle intrarégionale;
- ◆ la problématique de la santé mentale;
- ◆ les autres champs d'intervention relevant du régime de santé.

Les lignes directrices qui fondent, au plan théorique, l'analyse de la consommation et la mesure du besoin – à l'échelle des MRC, districts de CLSC ou des municipalités – sont identifiées. Il en est de même des critères méthodologiques qui assurent la validité des analyses infra-régionales. Ces critères permettent notamment de postuler qu'il ne faut pas s'attendre à l'échelle d'un quartier, d'une ville ou d'un district de CLSC de voir se reproduire, sans particularités, tous les constats qui valent à l'échelle de plus grandes populations.

L'analyse des problématiques dites de santé mentale a été réalisée tant pour les troubles transitoires que pour les sous-catégories⁴² qui composent les troubles sévères et persistants. Les résultats de nos analyses ont été présentés à des experts de la santé mentale – psychologues, psychiatres, médecins et autres professionnels du domaine – d'établissements québécois et régionaux. Les constats faits en santé physique sur le lien entre le niveau de non-emploi et l'incidence des maladies multifactorielles *valent également* en santé mentale, à une nuance près : dans le cas des troubles transitoires, le TI du moment s'avère meilleur prédicteur que celui des années antérieures, tel qu'il convient compte tenu des particularités propres à cet élément de la problématique.

Les domaines d'activités administrativement dénommés – adaptation sociale, intégration sociale, alcoolisme et toxicomanie, etc. – répondent à des phénomènes qui, à l'échelle des populations, sont fortement et directement influencés par les facteurs sociaux : la pauvreté, le

⁴² Des travaux, à l'échelle des régions, par sexe et groupes d'âge, sur les problèmes associés aux schizophrénies, aux psychoses maniaco-dépressives et à l'Alzheimer ont permis de constater la pertinence du niveau de non-emploi comme prédicteur.

vieillesse accélérée, l'exode des jeunes et, de façon plus générale, la désintégration des populations régionales. À cet égard, la problématique d'accroissement des taux de suicide, spécialement chez les jeunes, est l'une des plus reliées aux effets sociaux que déterminent les niveaux de non-emploi que mesure le TI.