

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

COMPARAISON DE L'EFFICIENCE DES SYSTÈMES DE SANTÉ
PROVINCIAUX CANADIENS

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR
JONATHAN RAMACIERI LAFONTAINE

OCTOBRE 2018

REMERCIEMENTS

Je tiens en premier lieu à remercier mon directeur de mémoire, Monsieur Pierre Ouellette, pour son engagement et sa confiance à mon égard. Merci infiniment pour vos nombreux conseils, votre disponibilité et votre générosité. Votre dévouement envers les étudiants et votre grande humanité font de vous un enseignant hors pair. Vous avez grandement contribué à la réalisation de ce mémoire et je vous en suis extrêmement reconnaissant.

Je tiens également à remercier mes collègues Marie-Josée Nadeau, François Turcotte, Alexis Chénard ainsi que Guillaume Provost, qui ont fait du parcours universitaire une expérience captivante et enrichissante. Je garde un excellent souvenir de nos nombreuses discussions et des bons moments passés ensemble.

Un merci particulier à mes parents, qui m'ont toujours soutenu tout au long de mon parcours scolaire et pour m'avoir toujours appuyé peu importe mes décisions. Vous m'avez permis de réaliser mes études dans les meilleures conditions possible, notamment grâce à votre support et vos encouragements qui ont été une source constante de motivation. Merci de croire en moi et d'être les parents que vous êtes.

Enfin, je tiens à remercier ma copine pour ses nombreux encouragements, ainsi que tous ceux et celles qui ont de près ou de loin contribué à la réalisation de ce mémoire. Que ce soit par le biais d'idées, de réflexions ou autres, votre appui a été grandement apprécié.

TABLE DES MATIÈRES

LISTE DES FIGURES	v
LISTE DES TABLEAUX	vi
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES	vii
RÉSUMÉ	viii
INTRODUCTION	1
CHAPITRE I	
REVUE DE LITTÉRATURE	7
1.1 Efficience des services de santé	9
1.2 Efficience en santé	10
CHAPITRE II	
MÉTHODOLOGIE	11
2.1 Méthodologie générale	11
2.2 Concept d'efficience	12
2.3 Revue des méthodes	14
2.3.1 Méthodes paramétriques	14
2.3.2 Méthodes non-paramétriques	17
2.4 Choix de la méthode	18
2.5 Méthode de Greene	19
CHAPITRE III	
DONNÉES	21
3.1 Ressources en santé	22
3.2 Services de santé	23
3.3 Niveaux de santé	24
3.4 Données socio-démographiques	25

CHAPITRE IV	
RÉSULTATS	28
4.1 Services de santé en fonction des ressources en santé	29
4.1.1 Résultats de l'estimation	29
4.1.2 Tests de la théorie	32
4.1.3 Mesures de la technologie et de l'inefficience	33
4.2 Niveaux de santé en fonction des services de santé	40
4.2.1 Résultats de l'estimation	40
4.2.2 Tests de la théorie	43
4.2.3 Mesures de la technologie et de l'inefficience	44
4.3 Niveaux de santé en fonction des ressources en santé	48
4.3.1 Résultats de l'estimation	48
4.3.2 Tests de la théorie	50
4.3.3 Mesures de la technologie et de l'inefficience	51
4.4 Explication du rejet de l'agrégation	56
CONCLUSION	59
RÉFÉRENCES	62

LISTE DES FIGURES

Figure		Page
1.1	Schéma de la production en santé	7
2.1	Représentation de l'efficacité technique et allocative	13
2.2	Ensemble de possibilités de production	16
4.1	Évolution de l'inefficacité des services de santé	38
4.2	Évolution de l'inefficacité des niveaux de santé	47
4.3	Évolution de l'inefficacité des niveaux de santé	55

LISTE DES TABLEAUX

Tableau	Page
3.1 Statistiques descriptives	27
4.1 Coefficients de la frontière des services de santé	31
4.2 Test de Hausman sur la production de services de santé	32
4.3 Monotonicité positive des services de santé	33
4.4 Rendements d'échelle des services de santé	34
4.5 Mesure du changement technologique des services de santé	35
4.6 Mesure de l'inefficience des services de santé	36
4.7 Quantification de l'inefficience des services de santé selon des années types (en millions de dollars réels)	40
4.8 Coefficients de la frontière des niveaux de santé	42
4.9 Test de Hausman sur la production des niveaux de santé	42
4.10 Monotonicité positive des inputs en santé et négative du taux de survie .	43
4.11 Rendements d'échelle des niveaux de santé	44
4.12 Mesure de l'inefficience des niveaux de santé	45
4.13 Coefficients de la frontière des niveaux de santé	49
4.14 Test de Hausman sur la production des niveaux de santé	49
4.15 Monotonicité des variables des niveaux de santé	50
4.16 Rendements d'échelle des niveaux de santé	51
4.17 Mesure du changement technologique des niveaux de santé	53
4.18 Mesure de l'inefficience des niveaux de santé	54

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

SFA	<i>Stochastic Frontier Analysis</i>
DEA	<i>Data Envelopment Analysis</i>
FDH	<i>Free Disposal Hull</i>
ICIS	Institut canadien d'information sur la santé

RÉSUMÉ

Depuis plusieurs années, les différents gouvernements provinciaux peinent à trouver des solutions face au poids grandissant qu'accaparent les sommes investies en santé. Une pression supplémentaire émerge avec l'accélération du vieillissement de la population et assombrit du même coup les perspectives pour l'avenir. Ce mémoire tente de comparer l'efficacité des différents systèmes de santé provinciaux. Par efficacité, on entend la capacité des systèmes à produire la quantité maximale d'output permise compte tenu de la quantité d'inputs utilisés. La méthode utilisée pour y parvenir porte le nom d'*Analyse des frontières stochastiques* et permet le tracé de frontières de production efficaces. La production de services de santé et l'atteinte de niveaux de santé de la population, que nous pouvons percevoir comme étant les produits du système de santé, seront étudiées. Les banques de données de l'Institut canadien d'information sur la santé et de Statistique Canada ont permis d'obtenir des données de qualité sur les ressources utilisées, les services de santé rendus à la population et les indicateurs du niveau de santé atteints par cette même population. Des données sociodémographiques caractérisant les différentes régions sont introduites afin de tenir compte de l'hétérogénéité des provinces. Quelques résultats émergent de notre mémoire. Premièrement, l'étude de l'efficacité du système de santé doit être décomposée en deux étapes. Il faut évaluer séparément les productions des services de santé et l'atteinte des niveaux de santé. L'analyse simultanée des deux aspects est perturbée par la relation entre les termes d'erreurs et les variables causales. Deuxièmement, les rendements d'échelle des institutions de santé sont décroissants ce qui implique que la taille des unités de production est trop grande. Cela appelle une réorganisation du système de santé afin de mieux tenir compte des possibilités organisationnelles et technologiques des institutions de santé. Troisièmement, les niveaux d'efficacité productive sont comparables entre les provinces et généralement plutôt élevés. On observe que les inefficiences se retrouvent principalement du côté de la transformation des ressources de santé en services de santé plutôt que dans la transformation des services de santé en niveau de santé. D'une certaine façon, cela est heureux, car cet aspect est celui qui est le plus directement sous le contrôle des gouvernements provinciaux. Quatrièmement, on note que ce sont les provinces ayant une importante démographie qui ont les niveaux les plus faibles d'efficacité productive. Finalement, étant donné le niveau des dépenses en santé, les inefficiences productives, bien que faibles, occasionnent des pertes monétaires se chiffrant en milliards de dollars.

Mots-clés : Efficacité, système de santé, analyse par les frontières stochastiques, provinces, rendement d'échelle.

INTRODUCTION

Les dépenses de santé représentent une part importante des dépenses des États presque partout dans le monde et cette part est en croissance. Cela constitue un sujet de préoccupation pour les gouvernements. Cette hausse s'explique par les changements technologiques et par les changements démographiques (Rapport Ménard, 2005). Non seulement le vieillissement de la population sera-t-il une source de dépenses de santé supplémentaires importantes au cours des prochaines années, mais la hausse du nombre de retraités s'accompagnera d'une réduction de la capacité fiscale des États concernés. Pour assurer des soins de santé de qualité et en assez grande quantité pour ces personnes et la population, les gouvernements devront utiliser efficacement leurs ressources dans ce contexte.

La situation est particulièrement préoccupante au Québec. En 2017-2018, les dépenses au Québec pour la santé et les services sociaux s'élèveront à 36,7 milliards de dollars ce qui équivaut à 50,65 % des dépenses de programmes du budget québécois. Cette hausse des dépenses par rapport à l'année précédente est de 4,2 % comparativement à une croissance économique prévue estimée à 1,7 % en 2017 (Le plan économique du Québec, 2017). À titre de comparaison, le montant accordé à la santé pour la Colombie-Britannique s'élève à 18,8 milliards de dollars pour le budget de 2017-2018, ce qui correspond à 37,06 % du budget total de la province (Budget and fiscal plan, 2017). Cette somme sera de 5,6 milliards de dollars pour la Saskatchewan ce qui représente 38 % du budget de la province ainsi qu'une hausse de 0,7 % par rapport à l'année précédente. (Saskatchewan provincial budget 17-18, 2017).

Face à ces difficultés budgétaires, les gouvernements des différentes provinces ont à surmonter un défi supplémentaire étant donné que nos hôpitaux n'ont pas d'incitatif à performer puisqu'ils n'opèrent pas dans un marché concurrentiel comme c'est le cas aux États-Unis (Bilodeau *et al.*, 2000, 2002, 2009 ; Chalkley et Malcomson, 2000 ; Ma, 1994 ; Mougeot et Neagelen, 2005). Dans ce contexte, les gouvernements doivent mettre en place des mesures de contrôle de l'efficacité afin de s'assurer que les ressources allouées à la santé le soient efficacement.

Notre mémoire a pour but d'évaluer et de quantifier l'efficacité des systèmes de santé provinciaux au Canada. L'efficacité du système de santé se situe à deux niveaux. Bien que cela puisse sembler paradoxal, dans les faits le système de santé ne produit pas de santé. Il produit des services de santé qui sont utilisés par la population (les patients) pour améliorer leur propre santé. Le système de santé utilise des ressources comme les médecins, les infirmières, les médicaments, etc., afin de rendre des services médicaux comme des consultations médicales, des traitements hospitaliers pour les patients hospitalisés et des consultations en cliniques externes, des opérations chirurgicales, etc. Il est légitime de se demander si cette étape de production de services de santé est efficace. Toutefois, produire des services médicaux n'est pas une fin en soi. Cela ne se justifie que si ces services permettent à la population d'accroître son niveau de santé. Dans cette deuxième étape, la population utilise les services de santé afin de se guérir ou de prévenir l'apparition de maladies ou d'accidents. Naturellement, les services de santé ne sont qu'un des aspects assurant une meilleure santé de la population. D'autres déterminants auront un impact comme le bagage génétique, l'éducation, l'hygiène publique, etc. Nous appellerons ces déterminants les variables socio-démographiques. Là aussi, on peut se demander si ce processus « socio-médical » est parfaitement efficace.

La « production » de santé est primordiale pour la population de chaque province : sans santé, rien ne vaut. Étant donné l'importance des sommes dépensées en santé, la population est en droit de se demander si la totalité de l'argent prélevé pour les soins de santé remplit bien son rôle et si une gestion différente pourrait s'avérer plus efficace qu'actuellement. Avant de poursuivre, il importe de préciser le sens du mot efficience.

Premièrement, on peut se demander si la frontière de l'ensemble des possibilités de production des services de santé est atteinte pour chaque province dans le domaine de la santé. En d'autres mots, est-ce que l'argent investi et qui sert entre autres à l'achat de matériel médical ainsi qu'aux salaires des médecins et des autres personnels pourrait permettre d'obtenir davantage de services de santé? Réciproquement, serait-il possible d'obtenir le même niveau de soin de santé qu'actuellement en réduisant la quantité des inputs? Ce terme d'efficience est connu sous le nom d'efficience technique.

Deuxièmement, un autre concept d'efficience est celui des rendements d'échelle. Le coût moyen dépend de la nature des rendements d'échelle. On sait que le coût moyen est minimisé lorsque les rendements sont unitaires. Dans ce cas, on parle d'efficience d'échelle.

Troisièmement, la substituabilité entre les nombreux inputs des systèmes de santé est aussi une source d'efficience. Étant donnés les prix sur le marché, est-ce que la combinaison d'inputs observée permet la minimisation des coûts? Même si l'efficience technique est atteinte (on produit sur l'isoquante), cela ne veut pas dire que la combinaison d'inputs est optimale. Il faut respecter des conditions de tangence entre l'isoquante et l'isocoût. Si ces conditions sont respectées, on dira que l'efficience allocative est atteinte. Comme exemple, on peut se demander si, étant donné les taux de salaire relatifs des médecins et des infirmières, une réduction des coûts serait possible dans le cas où les infirmières joueraient un rôle plus important dans le système de santé.

La connaissance de l'ensemble des possibilités de production est au cœur des solutions à ces problèmes. Plus précisément, il faut connaître la frontière de cet ensemble. Nous appellerons cette frontière la fonction de transformation (dans le cas à plusieurs outputs) ou fonction de production (dans le cas à un seul output). La position de l'observation par rapport à cette frontière permet la mesure de l'efficacité et la forme de la frontière détermine la nature des rendements d'échelle et des possibilités de substitution entre les inputs.¹

À l'aide de données, plusieurs méthodes sont envisageables pour mesurer la frontière de production du système de santé des provinces canadiennes. Une des principales méthodes s'effectue à l'aide d'une régression pour récupérer la fonction de production. Cette méthode est connue sous le nom de *Stochastic Frontier Analysis* (SFA) que nous pouvons traduire sous le vocable *Analyse des frontières stochastiques*. Dans le cas à un output, la frontière de production peut s'écrire $y \leq f(\mathbf{x}, t)$ où y est l'output, \mathbf{x} est le vecteur d'inputs et t est un terme de tendance captant le changement technologique.² Comme d'habitude, on ajoute un terme d'erreur observationnel, noté μ . Il s'ensuit $y \leq f(\mathbf{x}, t) + \mu$. Ce qui est particulier aux méthodes de mesure de l'inefficacité est l'ajout d'un deuxième terme d'erreur unilatère, noté ν (≥ 0), mesurant la distance entre la frontière de production et l'observation. On obtient : $y \leq f(\mathbf{x}, t) + \mu - \nu$. La principale difficulté consiste à définir la distribution conjointe des termes d'erreur. Le choix de la forme fonctionnelle de f et celui de la distribution des termes d'erreur sont importants puisqu'ils peuvent avoir un impact direct sur nos résultats (Aigner *et al.*, 1977; Meeusen et van Den Broeck, 1977).

L'autre principale méthode est basée sur des outils de programmation linéaire issue de la recherche opérationnelle pour « envelopper » les données. Les principales variantes

¹ L'absence de prix dans les données ne nous permettra pas d'étudier l'efficacité allocative.

² Par convention, une variable en caractères gras (resp. normaux) représente un vecteur (resp. un scalaire). Ainsi, \mathbf{x} est le vecteur des inputs et y est le niveau d'output.

sont le *Free Disposal Hull* (FDH) ainsi que le *Data Envelopment Analysis* (DEA). À l'aide des données, la méthode FDH utilise l'hypothèse de libre-disposition en inputs et en outputs pour engendrer un ensemble de possibilités de production. L'ajout de l'hypothèse de convexité de l'ensemble donne la méthode DEA. La méthode DEA constitue la méthode la plus populaire pour calculer l'inefficience (Bilodeau *et al.*, 2004). L'avantage de ces méthodes est qu'elles n'exigent pas l'imposition d'une forme fonctionnelle ni d'une forme stochastique particulière. En revanche, elles sont sensibles aux données aberrantes et le calcul des intervalles de confiance repose sur des méthodes de bootstrap (Ouellette et Petit, 2018).

À cause des difficultés rattachées au calcul des intervalles de confiance avec la méthode DEA, nous opterons plutôt pour la méthode paramétrique SFA.

Nous devons mettre sur pied une banque de données comprenant les inputs du système de santé (nombre de médecins, autres personnels, équipement, bâtiments, etc.), la quantité des services de santé produits vus comme des outputs intermédiaires (nombre de consultations, de visites en cliniques externes, etc.), l'état de santé de la population (espérance de vie des hommes et des femmes, taux de mortalité à la naissance, etc.) ainsi que les déterminants socio-démographiques. La principale source de données pour effectuer cette étude est l'Institut canadien d'information sur la santé (ICIS). Cet Institut fournit des données sur l'ensemble des éléments constituant notre système de santé, dont la quantité de main-d'œuvre, le type de soins accordés, les dépenses ainsi que les services de santé rendus à la population (ICIS, 2017). Pour compléter, nous aurons recours à la base de données de Statistique Canada, CANSIM, afin de mettre la main sur les niveaux de santé ainsi que les variables socio-démographiques qui ont une influence sur le niveau de santé de la population. Le résultat de cette cueillette de données sera une banque de données en panel. Cet aspect sera incorporé dans la méthode retenue.

Pour être en mesure de bien répondre au problème, au chapitre 1, une revue de littérature permettra de cerner les différents aspects de l'analyse en santé. Au chapitre 2, les différentes méthodes permettant la résolution du problème ainsi que les raisons derrière le choix de la méthode retenue seront présentées. Nous exposerons ensuite au chapitre 3 les différentes sections de la base de données et aborderons les transformations effectuées sur les variables. Le chapitre 4 sera quant à lui dédié à la présentation des résultats de nos estimations. Une conclusion suivra.

CHAPITRE I

REVUE DE LITTÉRATURE

Plusieurs inputs sont mis à la disposition du système de santé pour maintenir et améliorer la santé de la population canadienne. Des ressources sont acheminées pour la production de services de santé et une fois cette étape atteinte, les services de santé sont utilisés par les patients ou la population pour traiter leurs maladies ou améliorer leur état de santé. Le schéma de la figure 1.1 inclut aussi les facteurs socioéconomiques qui caractérisent les sociétés et les régions, ainsi que des caractéristiques individuelles qui aident chaque individu ou qui expliquent comment chaque individu réussit à transformer les services de santé en gain de santé. Ces facteurs ou déterminants peuvent être de différentes natures. La littérature en santé publique nous informe que parmi ces déterminants, on retrouve les facteurs génétiques, culturels ou encore environnementaux.

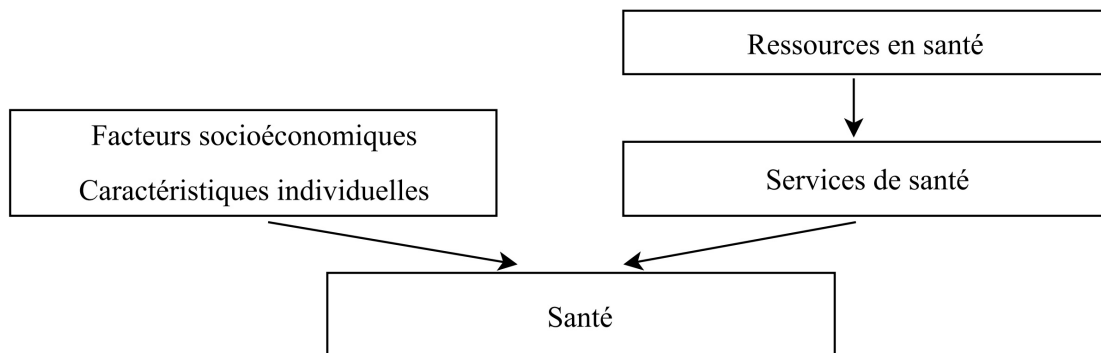


Figure 1.1 Schéma de la production en santé

L'étude de l'efficience du système de santé pris dans son ensemble peut se faire de deux façons distinctes : soit globalement sans tenir compte de l'étape intermédiaire de production de services de santé, soit en décomposant le processus en deux étapes distinctes. La première méthode repose sur la relation agrégée :

$$\text{Santé} = f(\text{Ressources en santé}, \text{Déterminants sociodémographiques}), \quad (1)$$

alors que la deuxième repose sur les deux relations :

$$\text{Services de santé} = f(\text{Ressources en santé}) \quad (2)$$

et

$$\text{Santé} = f(\text{Services de santé}, \text{Déterminants sociodémographiques}). \quad (3)$$

La première méthode est celle qui a été principalement retenue par les experts en santé publique. D'innombrables études portent sur cette relation. Le principal but est de quantifier l'impact des divers déterminants sur le niveau de santé des populations.

Une caractéristique de ces études (voir Ouellette et Petit, 2018) est qu'elles font presque systématiquement l'hypothèse que les divers systèmes de santé sont efficaces en ce sens que la frontière de production est toujours atteinte. Greene (2004 et 2005) a publié deux études en éliminant cette hypothèse et a proposé des méthodes économétriques de panel permettant de distinguer l'inefficience des effets idiosyncratiques tout en maintenant la relation agrégée. Nous retiendrons la proposition de Greene, mais pas le cadre d'analyse globale qui ne permet pas de décomposer les sources de l'inefficience. Nous opterons plutôt pour un modèle où l'efficience du système de production de services de santé sera mesurée indépendamment, puis celle de la population à transformer en santé les services de santé en tenant compte des déterminants sociodémographiques.

1.1 Efficience des services de santé

En ce qui concerne la production de services de santé, les méthodes économétriques et empiriques ont pris le dessus pour tenter de mesurer les déterminants et l'efficience du système de santé. Initialement, les études de coût en milieu hospitalier étaient de nature *ad hoc*. Evans (1971) a eu une influence déterminante dans ce genre d'études. Typiquement, on utilise des régressions avec le coût variable hospitalier comme variable dépendante et un ensemble de variables explicatives dont on pouvait penser qu'elles avaient un impact sur les coûts. Ces fonctions de coût n'étaient pas des fonctions de coût découlant du comportement optimisateur des institutions de santé. Pour cette raison, on les appelle des fonctions de coût *ad hoc* ou encore hédoniques. Il faudra attendre Cowing et Stevenson (1982) et Cowing et Holtman (1983) pour voir les premières estimations basées sur les fonctions de coût néo-classiques. Ce sera un point de non-retour et à partir de ce moment, les fonctions de coût ou de production seront toujours tirées de la théorie économique.

On observe que ces premières études ont systématiquement fait l'hypothèse d'efficience, c'est-à-dire que les chercheurs ont supposé que les hôpitaux produisaient toujours en un point sur la frontière des possibilités de production. Pourtant, dès 1951, Debreu et Koopmans avaient établi les bases du calcul de l'inefficience en termes de distance par rapport à la frontière de production. Ces travaux étant de nature exclusivement théorique, il faudra attendre Farrell (1957) pour montrer comment on pouvait mesurer l'inefficience à partir de méthodes de recherche opérationnelle. Les travaux de Farrell seront réactualisés par Charnes *et al.* (1978) et cela donnera lieu à la méthode connue sous le nom de *Data Envelopment Analysis*. En ce qui concerne les méthodes de régression, il faudra attendre Aigner *et al.* (1977) et Meeusen et van Den Broeck (1977) pour obtenir des fondements solides. Ces deux articles ont fondé la méthode appelée *Stochastic Frontier Analysis*. La méthode DEA et la méthode SFA

sont à la base de pratiquement tous les travaux empiriques en efficience. Ouellette et Petit (2018) font une synthèse théorique de ces travaux. La combinaison des travaux en efficience et la fonction de coût a été utilisée à de très nombreuses reprises pour mesurer l'efficience des services de santé à travers le monde. Ouellette et Petit (2010) recensent plusieurs centaines d'études sur l'efficience en santé.

1.2 Efficience en santé

Le processus de production de la santé a évolué de façon parallèle. Ce courant de la littérature trouve sa source dans le travail de Newhouse (1977) qui met en évidence la forte corrélation entre les dépenses de santé et le PIB par habitant ce qui suggère que la santé est un bien de luxe. Cet article a montré la voie à de très nombreux travaux. Certains auteurs notent une relation directe entre les dépenses en santé et les résultats sanitaires. En prenant en compte les facteurs socioéconomiques et les caractéristiques individuelles comme les habitudes de consommation, une baisse des dépenses en santé augmente la mortalité infantile et fait baisser l'espérance de vie au Canada (Crémieux *et al.*, 1999). Gerdtham et Jönsson (2000) subdivisent ces travaux en trois sous-groupes en fonction de la méthode économétrique utilisée. Le premier groupe recourt essentiellement aux moindres carrés ordinaires ; le second, à la suite de Hitiris et Posnett (1992), utilise plutôt des méthodes de panel afin de pallier le manque de degrés de liberté ; et le troisième, influencé par Roberts (2000), transforme les données en première différence afin de contrer la présence de racines unitaires. Toutefois, toutes ces études faisaient explicitement l'hypothèse d'efficience technique. Il faudra attendre Greene (2004, 2005) pour faire le lien entre les méthodes de panel et l'estimation en frontière stochastique.

CHAPITRE II

MÉTHODOLOGIE

2.1 Méthodologie générale

Le système de santé utilise des inputs pour produire de multiples outputs permettant d'améliorer la santé de la population. Dès lors, il est possible de concevoir l'existence d'un ensemble de possibilités de production. Cet ensemble montre les différents niveaux de production qu'il est possible d'atteindre en fonction des quantités d'inputs. Lorsque l'observation se retrouve sur la frontière de cet ensemble, l'efficacité parétienne est atteinte. Les fonctions de production sont toutefois soumises à la technologie qui est différente dans le temps et selon leur accessibilité. Les changements technologiques impliquent un déplacement de la frontière de production vers le haut.

L'importance de la production pour satisfaire le bien-être de la population a incité les économistes à vouloir évaluer les meilleures possibilités qui s'offrent dans les domaines qui régissent notre environnement. L'étendue des implications du domaine de la santé a fait l'objet d'études pour tenter de comprendre, élaborer puis quantifier les méthodes appliquées au fil du temps.

2.2 Concept d'efficience

On peut subdiviser l'efficience en deux composantes, soit l'efficience technique et l'efficience allocative. Debreu (1951) est l'un des premiers à vouloir s'attaquer à ce problème. Il démontre que le calcul de l'efficience s'effectue en calculant la distance d'un ensemble d'inputs par rapport à la frontière efficiente. Cela représente ce que cet ensemble peut nous offrir lorsqu'employé optimalement. Les pertes d'efficience sont représentées soit sous une forme de sous-emploi des ressources physiques, d'inefficience dans la production ou encore selon une imperfection dans l'organisation économique. Pour calculer cette perte, un paramètre peut venir capter ces trois sources d'inefficience lors du problème de maximisation de la production sous contrainte technologique. L'autre concept d'efficience est proposé par Koopmans (1951). L'inclusion d'un système de prix en lien avec l'efficience indique que non seulement la production doit être maximale, mais que celle-ci doit être au moindre coût. On parlera alors d'efficience allocative et cela apporte une approche monétaire à la théorie.

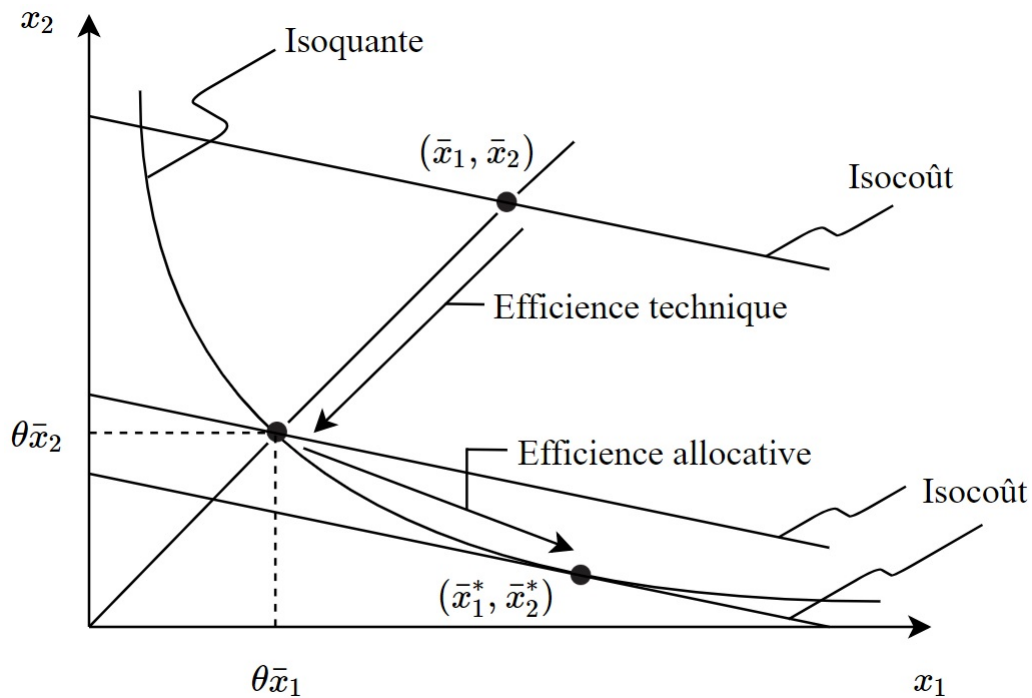


Figure 2.1 Représentation de l'efficacité technique et allocative

La figure 2.1 représente un exemple de production à l'aide de deux inputs dans lequel on peut distinguer les deux types d'efficacité. L'isoquante représentée est une courbe reliant toutes les combinaisons d'inputs permettant de produire une même quantité d'output. Le coût d'opportunité devient ainsi plus important à mesure où l'on tente de substituer un certain input. La même logique s'applique pour les trois droites d'isocoûts où l'on parlera d'un même coût le long de la droite. Lorsque l'on cherche à produire à une certaine quantité sur l'isoquante, on doit s'assurer que cela est fait avec une combinaison d'inputs qui n'excède pas la quantité minimale permise. Ce n'est pas le cas ici au départ. La combinaison d'inputs observée (\bar{x}_1, \bar{x}_2) pourrait permettre la production d'une quantité plus importante d'output puisque des isoquantes plus loin de l'origine sont accessibles. À l'inverse, pour s'assurer du meilleur rendement de nos inputs, on doit chercher à les réduire. Si on opte pour une réduction proportionnelle,

on obtient un facteur de réduction θ (avec $0 < \theta \leq 1$) qui représente la plus petite fraction des inputs initiaux permettant de produire l'output observé. On se déplace ainsi dans le plan et la distance parcourue entre l'observation et l'isoquante représente l'inefficience technique. L'efficience globale n'est toutefois pas atteinte à ce point, il est encore possible de passer à une droite d'isocoût inférieure, ce qui correspond à la minimisation du coût. Pour y parvenir, il suffit de réallouer efficacement nos ressources en substituant certains inputs entre eux. L'exemple ci-haut montre qu'en substituant du \bar{x}_2 au profit de \bar{x}_1 , on atteint un point de tangence entre la courbe d'isoquante et la droite d'isocoût. La distance parcourue lors du déplacement le long de l'isoquante représente alors l'inefficience allocative. À ce point, on ne peut atteindre une droite d'isocoût inférieur si l'on veut demeurer sur l'isoquante ce qui fait du point $(\bar{x}_1^*, \bar{x}_2^*)$ l'optimum où l'efficience globale est atteinte. L'atteinte de ce point est importante dans l'organisation de la production, sans quoi il y aura un coût excédentaire attribuable aux deux types d'efficience. L'inefficience technique entraînera un coût pour chaque unité d'input additionnel alors que l'inefficience allocative occasionnera une perte correspondant à une mauvaise combinaison d'inputs. Le domaine de la santé est soumis aux deux types d'efficience. Même si la taille du système impressionne, cela ne suppose pas qu'il ne peut être modifié. La connaissance de la production et des différentes combinaisons d'inputs qui composent le système de santé est donc nécessaire pour obtenir les plus hauts niveaux d'efficience possible dans la production de santé.

2.3 Revue des méthodes

2.3.1 Méthodes paramétriques

Essentiellement, les méthodes paramétriques recourent à des régressions afin de calculer l'ensemble de production. La méthode est simple à décrire. À partir d'observations, on commence par spécifier une forme paramétrique pour la frontière de production (ou

de coût). Les formes les plus populaires en recherche ont été la Cobb-Douglas et la forme Translog. La Cobb-Douglas peut être vue comme une expansion de Taylor de premier ordre après transformation logarithmique des variables. Comme cette forme limite le domaine de variation de certaines mesures de la technologie, on a suggéré de prendre plutôt une expansion de deuxième ordre et on obtient ainsi la forme Translog. Cette forme est dite flexible (Diewert, 1976) puisque justement elle n'impose aucune restriction sur les mesures de la technologie. Il reste à définir la structure des termes d'erreur et un critère d'estimation, moindres carrés ou autre.

La difficulté découle du fait que l'ensemble de production (ou de coût) est défini par une inégalité, c'est-à-dire $f(\mathbf{y}, \mathbf{x}, t) \leq 0$ dans le cas multi-output ou encore $y \leq f(\mathbf{x}, t)$ dans le cas à un seul output.³ Le fait de choisir une forme fonctionnelle pour f modifie légèrement les formules précédentes. On obtient, dans le cas multi-output : $f(\mathbf{y}, \mathbf{x}, t, \boldsymbol{\beta}) \leq 0$ où $\boldsymbol{\beta}$ est le vecteur des paramètres liés à la forme fonctionnelle retenue. Si on ajoute un terme d'erreur observationnel, il s'ensuit : $f(\mathbf{y}, \mathbf{x}, t, \boldsymbol{\beta}) \leq \mu$, où μ est le terme d'erreur observationnel. La présence de l'inégalité rend les méthodes usuelles inopérantes. La parade consiste à introduire un terme de distance entre l'observation et la frontière (Aigner *et al.*, 1977; Meeusen et van Den Broeck, 1977). Cette méthode porte le nom de *Stochastic frontier analysis*. Cette application amène la forme $f(\mathbf{y}, \mathbf{x}, t, \boldsymbol{\beta}) = \mu - \nu$, où ν représente le terme d'erreur relié à l'inefficience. Le côté droit de l'équation, où les termes d'erreur sont décomposés, représente la distance entre le niveau d'output efficient y^* et le niveau d'output observé y tel que présentée à la figure 2.2. Cet écart est dû en partie à des erreurs d'observation (μ) et par de l'inefficience (ν). Comme ν mesure la distance entre la frontière et l'observation, ce terme est nécessairement positif et pour cette raison on l'appelle unilatère ($\nu \geq 0$).

³ On a ajouté un indice de temps afin de capter les déplacements dus au changement technologique.

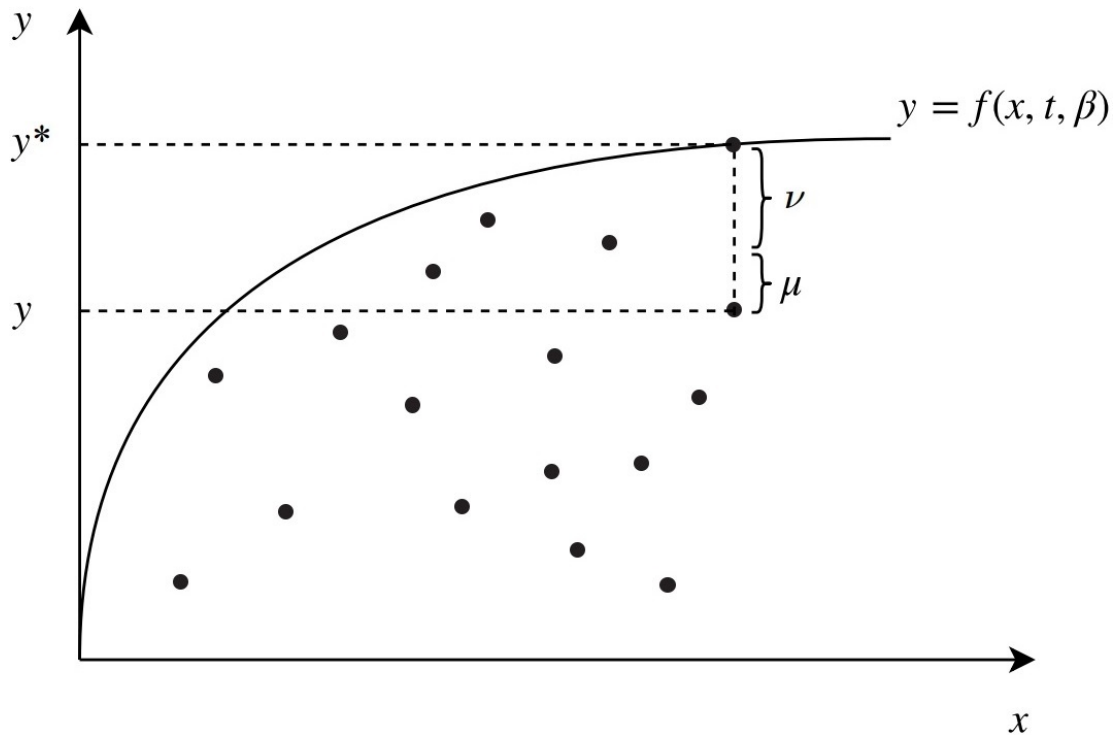


Figure 2.2 Ensemble de possibilités de production

À titre d'exemple, dans le cas à un seul output, on peut écrire la forme Translog comme suit : $\ln y = \beta_0 + \sum_i \beta_i \ln x_i + \sum_i \sum_j \beta_{ij} \ln x_i \ln x_j + \mu - \nu$ où $\nu \geq 0$.⁴ La paramétrisation de la méthode se poursuit avec la définition des distributions des termes d'erreurs. Pour le terme d'erreur observationnel, l'utilisation la plus courante est une distribution normale de moyenne nulle : $\mu \sim N(0, \sigma^2)$ où σ^2 est la variance du terme d'erreur. La complication survient lors de la définition de la distribution du terme d'erreur unilatère ν qui mesure l'inefficience technique. La distribution de ce terme d'erreur est inconnue et tout comme la forme fonctionnelle, ce choix aura des impacts significatifs sur les résultats. Il faut choisir une distribution parmi les nombreuses que nous proposent les méthodes statistiques. Parmi les plus connues, l'utilisation de la loi khi-carré, de la log-normale, de l'exponentielle et de la demi-normale est envisageable.

⁴ En posant $\beta_{ij} = 0$, pour tout (i, j) , on obtient le cas Cobb-Douglas.

2.3.2 Méthodes non-paramétriques

Le fait que les méthodes paramétriques soient soumises à de fortes hypothèses quant à la forme fonctionnelle et les distributions des termes d'erreur ont poussé les chercheurs vers d'autres méthodes dans les années 1990. Ce fut le cas pour la méthode *Data Envelopment Analysis* et une méthode apparentée, le *Free Disposal Hull*.

Le concept de base est la fonction de distance, une représentation duale de l'ensemble des possibilités de production. Bien que définie il y a plusieurs décennies, l'utilisation de cette fonction a été réintroduite par Caves *et al.* (1982) avec l'hypothèse d'efficacité. Färe *et al.* (1992) ont généralisé la méthode de Caves *et al.* en laissant tomber l'hypothèse d'efficacité. Farrell (1957) propose une mesure d'efficacité qui met de l'avant la construction des isoquantes servant au calcul de la fonction de distance. Cette approche sera reprise et généralisée par Charnes *et al.* (1978) pour donner la méthode DEA. Banker *et al.* (1984) approfondiront la méthode en incluant une technologie à rendements d'échelle variables.

Essentiellement, cette méthode ne fait aucune hypothèse sur la forme fonctionnelle ni sur la structure stochastique. Elle se contente de faire des hypothèses de libre-disposition tout en convexifiant l'ensemble résultant. Sans la convexification, on obtient la méthode *Free Disposal Hull* proposée par Deprins *et al.* (1984) qui peut être vue comme un cas particulier de la méthode DEA. Tant la méthode DEA que FDH remplacent l'ensemble des possibilités de production par un ensemble de contraintes linéaires. Alors que l'absence d'hypothèses peut être vue comme une amélioration, le problème d'optimisation résultant exige le recours à des techniques d'optimisation appropriées. Dans ce cas, ce sera la programmation linéaire. De plus, cette méthode ne fournit pas directement des mesures de confiance suite au calcul de l'inefficacité. Simar et Wilson (1998) ont proposé des méthodes de bootstrap pour y arriver.

2.4 Choix de la méthode

La méthode paramétrique est retenue pour estimer la technologie. Cette méthode, populaire dans les années 1970, s'est vue accorder moins d'importance lors des deux dernières décennies compte tenu de ses désavantages en ce qui a trait entre autres au choix d'une forme fonctionnelle et de la distribution de ses termes d'erreurs. Les avancées des dernières années en matière de programmation ont toutefois réduit l'impact de ses inconvénients. De plus, le recours à la méthode bootstrap pour calculer les intervalles de confiance avec les méthodes non-paramétriques n'a toujours pas prouvé sa validité. C'est pourquoi nous assistons à un retour des méthodes paramétriques depuis quelques années (Ouellette et Petit, 2010).

Une fois la base de données en main et la paramétrisation complétée, la frontière de production peut être estimée à l'aide de logiciel statistique. Nous estimerons la frontière efficiente à partir du logiciel *Stata* qui se démarque par la contribution de ses utilisateurs. C'est notamment le cas avec la commande *sfp* dont le fonctionnement est moins limitatif que la commande *xtfrontier*. On privilégie donc *sfp* dans cette situation d'efficacité en santé puisqu'elle offre la possibilité d'évaluer un plus large éventail de modèle d'inefficacité variant dans le temps. Au regard de cet avantage, l'aspect le plus important reste l'inclusion des modèles à effets fixes et à effets aléatoires. Ces modèles mesurent la stochasticité de la frontière à partir d'une composante d'hétérogénéité et d'inefficacité (Belotti *et al.*, 2013). L'ajout de l'hétérogénéité amène une modification notable des résultats de l'estimation et varie dans le temps (Greene, 2004 et 2005). La commande *sfp* avantage par conséquent la frontière du modèle.

La résolution du problème déterminera les raisons de l'inefficacité en soulevant la part de l'inefficacité technique. De cette façon, nous pourrions quantifier en valeur monétaire ou en perte d'année de vie l'inefficacité de nos systèmes de santé.

2.5 Méthode de Greene

Le modèle de Greene sert à estimer une frontière stochastique en panel :

$$y_{it} = \alpha_{it} + f(x_{it}) + \mu_{it} - \nu_{it}$$

où y est l'un des outputs, x représente l'ensemble des autres variables du modèle comprenant les autres outputs, α est la constante de la régression, μ est le terme d'erreur observationnel et $\nu (> 0)$ est la mesure de l'inefficience. L'indice i représente l'individu et t le temps. Ce qui caractérise cette équation, c'est d'une part la présence du terme d'erreur unilatère, ν mais aussi le fait que la constante dépende de i et de t . C'est ce dernier aspect qui a été pris en compte par Greene (2004 et 2005) dans le calcul de la frontière stochastique.

La distribution du terme d'erreur observationnel est souvent représentée par une loi normale bilatérale où $\mu_{it} \sim N(0, \sigma^2)$. La partie complexe est de déterminer la distribution de ν_{it} étant donné que la dispersion des termes d'inefficience est inconnue. Néanmoins, on peut faire l'hypothèse d'une distribution exponentielle, tronquée ou demi-normale. Pour la première, $\nu_{it} \sim exp(\lambda)$ où λ est un paramètre d'échelle. La distribution demi-normale peut être représentée comme $\nu_{it} \sim |\mu_{it}|$. Enfin, la loi normale tronquée permet la dispersion des termes d'erreurs en restreignant le domaine de la loi normale. La probabilité d'obtenir une inefficience élevée diminue donc à mesure que celle-ci croît. Pour le cas qui nous intéresse, on suppose donc qu'il est peu probable qu'un système soit excessivement inefficent. Ces trois distributions sont représentées sur un domaine ayant comme intervalle $[0, \infty[$ ce qui satisfait la positivité de la distribution.

Le choix de la forme fonctionnelle est aussi une caractéristique importante du modèle. Plusieurs choix sont possibles pour représenter la fonction de production, mais les plus populaires demeurent la forme Cobb-Douglas et Translog. La forme Cobb-Douglas est plus restrictive puisqu'elle impose certaines restrictions sur les élasticités des firmes. Cependant, on la privilégie pour sa facilité d'interprétation et le peu de paramètres nécessaires à son estimation. Une fois la linéarisation complétée par une transformation logarithmique des variables, on obtient :

$$\ln y_{it} = \beta_0 + \sum_n \beta_n \ln x_{nit} + \mu_{it} - \nu_{it}$$

La forme fonctionnelle Translog est quant à elle plus flexible du fait qu'elle incorpore des termes quadratiques en log. Cela amène une plus grande précision que la forme précédente et permet de relâcher certaines hypothèses sur les élasticités. Cela s'obtient au détriment d'une augmentation du nombre de paramètres à estimer. Une fois la transformation appliquée, on obtient :

$$\ln y_{it} = \beta_0 + \sum_i \beta_i \ln x_{it} + 0,5 \sum_i \sum_j \beta_{ij} \ln x_{it} \ln x_{it} + \mu_{it} - \nu_{it}$$

Ces deux formes fonctionnelles sont linéaires dans leurs paramètres et peuvent être estimées à l'aide de la méthode des moindres carrés ordinaires. Il est aussi possible d'imposer des conditions sur les paramètres qui les régissent.

Étant donné la trop grande simplicité de la Cobb-Douglas et la complexité de la Translog, nous opterons pour une forme à mi-chemin entre les deux. Cela s'effectuera avec la sélection de termes croisés. Nous choisirons donc des termes de deuxième ordre selon une ligne, une colonne ou une diagonale.

CHAPITRE III

DONNÉES

Le traitement de la base de données en panel est une partie importante de ce projet d'étude. La réalisation de celle-ci s'effectue à l'aide de deux sources de données. L'Institut canadien d'information sur la santé fournit des données sur les ressources utilisées ainsi que les services de santé administrés à la population. La base de données de Statistique Canada complétera notre échantillon en y ajoutant les niveaux de santé de même que les données socio-démographiques qui fourniront certains indicateurs des niveaux de santé de la population.

Nous présentons les données en les regroupant selon les catégories de la figure 1.1 : Ressources en santé, Services de santé, Niveaux de santé et Données socio-démographiques. Nous utiliserons des données pour neuf provinces canadiennes. Compte tenu du manque de données disponibles concernant les trois territoires canadiens, ceux-ci ont été laissés de côté. Par ailleurs, étant donné le faible bassin de population de l'Île-du-Prince-Édouard, cette province est aussi mise de côté. La correction de l'inflation se fait à l'aide de l'indice des prix à la consommation à l'année de base 2002. Les variables monétaires sont de cette façon représentées en prix réels. En ce qui trait aux dépenses en médicaments, l'indice des prix à la consommation pour les médicaments avec et sans ordonnance a été utilisé dans le but de préserver la nature de la variable.

D'autre part, mise à part la correction de l'inflation, d'autres transformations ont été effectuées sur les variables. Ainsi, dans le but de minimiser l'impact des différentes tailles de population entre les provinces canadiennes, les variables utilisées sont exprimées sous la forme *per capita*. À noter que la transformation logarithmique ne s'applique pas aux variables représentées sous forme de taux. Cela permettra de couvrir une période de vingt ans s'étendant de 1997 à 2016. Finalement, le tableau 3.1 présentera les statistiques descriptives avec le regroupement de chaque catégorie de variables. Celui-ci comporte entre autres la moyenne, qui est calculée sur l'ensemble des années et des provinces. Pour en faciliter la compréhension, les variables présentes dans ce tableau ne sont pas représentées sous la forme *per capita*.

3.1 Ressources en santé

Les ressources utilisées dans le domaine de la santé constituent les inputs primaires. C'est grâce à leur répartition que la production de services en santé est rendue possible. Ces ressources peuvent être désagrégées en deux groupes distincts avec le personnel en santé ainsi que des dépenses relatives au système de santé. En ce qui concerne le personnel, l'hétérogénéité des traitements accordés requiert une diversité au sein du corps professionnel. La liste se résume donc aux médecins, aux infirmières autorisées, aux infirmières auxiliaires autorisées, aux physiothérapeutes et aux ergothérapeutes pour la période allant de 1997 à 2015. Étant donné une certaine équivalence entre les types d'infirmières et les types de thérapeutes, ces deux ensembles de variables seront agrégés selon leur catégorie respective. Un nombre élevé de spécialistes en santé signifie une plus grande quantité et qualité de soins et de services.

Ces spécialistes qui exercent leurs fonctions le font dans des établissements de soins à l'aide d'équipements spécialisés. Ces ressources sont représentées sous forme de

stock de capitaux en bâtiment (non résidentiel) et en équipement et machinerie. L'ICIS comptabilise d'autres types de dépenses qui sont catégorisées par affectation de fonds. Les dépenses en médicaments, en santé publique ainsi qu'en administration sont les catégories de fonds qui ont été retenues. Puisque ces séries sont colinéaires, dans la mesure où les variables sont multipliées par un terme correspondant au prix d'une année, il est possible de procéder à une normalisation. Cette façon de procéder nous permet d'obtenir des indices de quantité sans perdre de généralité. Cet ensemble de variables nous procure des données sur le sujet jusqu'en 2016.

3.2 Services de santé

Il faut ensuite considérer les données qui traitent des soins de santé accordés à la population et qui servent notamment à améliorer les indicateurs des niveaux de santé. Ces données, qui proviennent de l'ICIS, seront recensées selon trois catégories. Dans la mesure où le système de santé administre des soins selon plusieurs formes, il y a au préalable un travail qui consiste en un échange d'information entre le patient et le médecin. Cet échange s'effectue au terme de visites et de consultations pour juger du meilleur traitement à attribuer selon la situation de la personne concernée. Au nombre de celles-ci, on observe les consultations, les évaluations majeures, les autres évaluations, les jours de soins hospitaliers, les visites spéciales ainsi que la psychothérapie et le counseling. Par la suite, lorsque l'information est traitée par le système de santé, on est en mesure de procéder aux interventions, qui sont l'autre grande catégorie de services. Ceux-ci se caractérisent par les chirurgies mineures, les chirurgies majeures, l'aide chirurgicale, les services obstétricaux, les services diagnostiques et thérapeutiques, les services spéciaux ainsi que les services divers. Puisque les consultations et les interventions constituent l'essence même des services de santé, ces deux groupes de variables seront agrégés pour représenter la totalité des services de santé. Ces

données sont comptabilisées pour l'ensemble des médecins pour les exercices médicaux s'étendant de 1996-1997 à 2014-2015.

Dans le but de rendre les données comparables, il est important de préciser que ce travail nécessite le traitement des données selon le même horizon temporel. Pour faciliter la lecture des résultats, l'année civile est choisie, ce qui inclut la période s'étendant du 1^{er} janvier au 31 décembre d'une même année. Cependant, certaines variables, dont les services de santé, sont uniquement disponibles selon le format d'un exercice médical utilisant la période du 1^{er} avril au 31 mars. Une reconfiguration des exercices est donc nécessaire pour rendre compatible l'étendue des données. La conversion s'est effectuée en combinant les trois quarts d'un exercice médical au quart de l'exercice suivant pour obtenir l'année civile. Pour se faire, un premier traitement de la base de données a été réalisé avec le logiciel *Excel*. Cette même manipulation a dû aussi être appliquée à la dernière catégorie de services de santé qui est la durée moyenne du séjour hospitalier. Même si à première vue, l'hospitalisation n'est pas considérée comme un traitement médical au sens propre, celle-ci est importante puisque la réalisation des services passe aussi par l'hébergement des patients. Ainsi, plusieurs interventions ne pourraient être réalisées sans l'admission d'un patient dans un centre hospitalier dû à leur importance et à la convalescence qui en découle.

3.3 Niveaux de santé

Les indicateurs des niveaux de santé occupent une place importante pour cette étude dans la mesure où cela constitue la finalité des systèmes de santé. On y retrouve le taux de mortalité infantile (1997 à 2013) qui est calculé en termes de décès par mille naissances vivantes et l'espérance de vie à soixante-cinq ans (1997 à 2012). Ces données sont fournies par CANSIM et aideront à analyser l'impact des services de santé dans

leur ensemble. Pour l'espérance de vie à soixante-cinq ans, les estimations récentes sont fournies selon une moyenne de trois ans. La récupération des années individuelles s'est donc effectuée en désagrégant les moyennes et en isolant la dernière année lorsque les deux précédentes étaient disponibles. Puisque la mortalité touche davantage les personnes ayant un âge élevé, il est important d'avoir un indicateur relatif à cette caractéristique. Par ailleurs, nous utiliserons le taux de survie infantile, défini comme la différence entre 1 et le taux de mortalité infantile, de façon à travailler avec deux outputs positifs. La sélection de ces données s'est effectuée du fait de leur large horizon temporel, mais surtout puisqu'elles sont facilement comparables entre les provinces canadiennes. Les années plus récentes n'étaient pas disponibles au moment d'effectuer l'étude.

3.4 Données socio-démographiques

Pour conclure, le dernier groupe de données qui sera utile pour cette étude sera composé de données socio-démographiques. Celles-ci sont fournies par CANSIM et seront déterminantes pour nos résultats. Leur importance découle de la façon dont ces variables viennent modifier le procédé par lequel nous pouvons convertir les ressources et services en santé en gain de santé. Notre sélection s'est effectuée selon un facteur socioéconomique qui est le taux de chômage (1997 à 2016) de même que trois caractéristiques individuelles qui concernent les dépenses en alimentation, en alcool et en tabac (1997 à 2015). Étant donné que les variables concernant les dépenses sont en prix constants, nous pouvons les percevoir comme des indices de quantité.

Malgré la cyclicité du taux de chômage, cette variable nous renseigne sur la disponibilité des revenus des ménages. Ceux ayant des revenus plus importants ont un choix additionnel grâce au système de santé privé. Délaisser le système de santé publique

au prix d'un coût supplémentaire signifie souvent une réduction du temps d'attente. La consommation de médicaments non couverts par une assurance publique ou la possibilité de subir une chirurgie non nécessaire, mais permettant d'augmenter la qualité de vie est également à prendre en compte. Par la suite, les indices de quantité en alimentation, alcool et tabac sont à considérer puisque ces facteurs ont une incidence directe sur les niveaux de santé. Une alimentation complète et équilibrée permet d'améliorer les indicateurs de santé. À l'opposé, la consommation d'alcool et de tabac sont des habitudes nuisibles au prolongement de l'espérance de vie.

Tableau 3.1 Statistiques descriptives

Variable	Période	Moyenne	Écart-type	Minimum	Maximum
Médecins	1997 - 2015	7 228	7 781	925	30 494
Infirmières autorisées et auxiliaires autorisées	1997 - 2015	36 271	37 085	8 007	137 133
Physiothérapeutes et ergothérapeutes	1997 - 2015	3 108	3 257	259	12 864
Stock de bâtiments	1997 - 2015	3 417 280 702	3 647 661 475	356 000 000	16 298 000 000
Stock de machines	1997 - 2015	389 953 216	451 100 326	13 000 000	1 774 000 000
Indice de quantité de médicaments	1997 - 2016	2 697 901 426	3 417 233 990	186 971 946	14 629 419 399
Indice de quantité de santé publique	1997 - 2016	908 707 773	1 114 201 979	51 888 278	5 323 535 724
Indice de quantité d'administration	1997 - 2016	528 040 871	582 474 236	24 387 480	2 227 236 774
Consultations et interventions	1997 - 2014	26 984 800	32 089 351	3 299 323	109 553 400
Durée moyenne du séjour hospitalier	1997 - 2015	7,74	1,05	5,85	9,90
Espérance de vie à 65 ans	1997 - 2012	19,10	0,92	16,50	21,40
Taux de survie infantile (en %)	1997 - 2013	994,72	1,16	991,10	996,80
Taux de chômage (en %)	1997 - 2016	7,96	3,14	3,50	18,10
Indice de quantité d'alimentation	1997 - 2015	2 264	129	1 957	2 523
Indice de quantité d'alcool	1997 - 2015	496	82	340	731
Indice de quantité de tabac	1997 - 2015	500	109	277	749

CHAPITRE IV

RÉSULTATS

Ce chapitre permettra l'analyse des résultats dans le contexte actuel tout en s'assurant de leur authenticité suite à l'application des fondements théoriques. L'estimation des coefficients et le choix des termes croisés seront discutés dans un premier temps. La significativité de l'erreur unilatère, qui indique l'inefficience, de même que les tests de Hausman feront ensuite l'objet d'analyse. Les tests relatifs au respect de la théorie seront par la suite abordés.

Une fois les fondements des modèles bien établis, nous nous concentrerons sur les différentes caractéristiques de la technologie. Le changement technologique et les rendements d'échelle seront au départ analysés. Par la suite, les résultats concernant l'inefficience, qui sont au cœur de cette étude, seront présentés. Finalement, la quantification de l'inefficience permettra une comparaison entre les provinces.

Les trois modèles présentés lors de la revue de littérature seront à l'étude. Pour chacun d'eux, une distribution exponentielle a été utilisée pour le terme d'erreur unilatère. Aux fins de comparaison, les résultats du Canada concernant l'inefficience sont présentés. Une moyenne pondérée des neuf provinces canadiennes, qui a été réalisée en fonction de l'output utilisé, a permis la récupération de ce résultat. Pour terminer, des arguments sur la meilleure approche à utiliser pour faire l'analyse en santé sont présentés.

4.1 Services de santé en fonction des ressources en santé

4.1.1 Résultats de l'estimation

Le premier modèle, tiré à partir de l'équation (2), relie l'output (les services de santé) aux inputs du système de santé (médecins, infirmières, etc.). Le tableau 4.1 montre l'estimation des coefficients de la fonction de production des services de santé, de même que les valeurs p correspondantes. La durée moyenne du séjour hospitalier, qui est un indice de qualité d'output, n'est pas soumise à des conditions de monotonie. Son résultat demeure néanmoins intéressant. Par qualité, on entend que cet indice sert à qualifier (dans le sens de décrire) les services offerts par le système de santé et non de dépeindre la qualité (dans le sens de bon ou mauvais) des services. La négativité de ce coefficient nous indique qu'un long séjour hospitalier ralentit le processus de production de services de santé. À l'inverse, de plus courts séjours signifient un plus grand roulement de patients ce qui se traduit par davantage de consultations et d'interventions.

Outre cet indice de qualité, nous avons introduit des termes croisés dans les inputs. La multiplication des médecins par les autres types de personnel et stock de capitaux apparaît comme un choix logique étant donné l'importance et la complémentarité des médecins dans le système de santé. Le produit de l'indice de temps et des ressources de santé fournies par le milieu hospitalier nous informe de son côté du changement technologique dans cette sphère d'activité. Enfin, l'inclusion de la durée moyenne du séjour hospitalier au carré permet de discerner une possible non-linéarité. Un long séjour pourrait ainsi se traduire par l'octroi de multiples services de santé, ce qui est moins enclin à se produire à mesure que le séjour hospitalier est raccourci.

D'autre part, nous devons nous assurer de la significativité de l'erreur unilatère. Cela est primordial car les données de l'inefficience proviennent de ce terme d'erreur. Cela n'est cependant pas une source de problème puisque l'estimation démontre que l'erreur unilatère demeure significative selon un seuil de 5 %.

Tableau 4.1 Coefficients de la frontière des services de santé

Consultations et interventions	Coefficient	Valeur <i>p</i>
Indice de qualité d'output		
Durée moyenne du séjour hospitalier	-3,586	0,001
Inputs		
Médecins	8,303	0,079
Infirmières autorisées et auxiliaires autorisées	-1,677	0,418
Physiothérapeutes et ergothérapeutes	2,615	0,069
Stock de bâtiments	1,689	0,011
Stock de machines	-0,524	0,466
Indice de quantité de médicaments	0,618	0,000
Indice de quantité de santé publique	-0,074	0,061
Indice de quantité d'administration	-0,074	0,071
Termes croisés		
Durée moyenne du séjour hospitalier ²	1,666	0,001
Médecins ²	1,266	0,045
Médecins * Infirmières	-0,275	0,398
Médecins * Thérapeutes	0,395	0,086
Médecins * Stock de bâtiments	0,270	0,010
Médecins * Stock de machines	-0,086	0,449
Médecins * Temps	-0,004	0,777
Indice de quantité de médicaments * Temps	-0,038	0,000
Indice de quantité de santé publique * Temps	0,008	0,005
Indice de quantité d'administration * Temps	0,022	0,000
Changement technologique		
Temps	0,023	0,803
Temps ²	0,003	0,000
σ_{μ}	0,030	0,000
σ_{ν}	0,007	0,019
λ	4,083	0,000

Pour capturer les effets idiosyncratiques, nous avons estimé un modèle à effets fixes et un autre à effets aléatoires. L'inconvénient d'un modèle à effets fixes se traduit par de plus grandes valeurs p . Cela nous permet toutefois de relâcher l'hypothèse de normalité des termes d'erreur ce qui est un avantage. Nous avons eu recours au test de Hausman pour tester s'il était préférable d'utiliser un modèle à effets fixes ou à effets aléatoires pour l'estimation de la fonction de production. Le résultat du test de Hausman est indiqué au tableau 4.2. À un seuil de significativité de 5 %, on rejette l'hypothèse nulle selon laquelle la normalité des termes d'erreur n'est pas contraignante.

Tableau 4.2 Test de Hausman sur la production de services de santé

$\begin{aligned}\chi^2(15) &= (\beta_{TFE} - \beta_{TRE})'[\sigma^2(\beta_{TFE}) - \sigma^2(\beta_{TRE})]^{-1}(\beta_{TFE} - \beta_{TRE}) \\ &= 25,76 \\ \mathbb{P} > \chi^2(15) &= 0,0406\end{aligned}$
--

4.1.2 Tests de la théorie

Les résultats de l'estimation peuvent être utilisés seulement si les hypothèses économiques derrière la forme fonctionnelle sont respectées. Par conséquent, le modèle doit respecter des propriétés de monotonie positive en inputs. L'allocation d'un input doit donc se traduire par une production supplémentaire d'output.

Tableau 4.3 Monotonicit  positive des services de sant 

Variable	Moyenne
M�decins	0,288
Infirmi�res autoris�es et auxiliaires autoris�es	0,037
Physioth�rapeutes et ergoth�rapeutes	0,154
Stock de b�timents	0,004
Stock de machines	0,014
Indice de quantit� de m�dicaments	0,217
Indice de quantit� de sant� publique	0,010
Indice de quantit� d'administration	0,152

Les r sultats du tableau 4.3 nous indiquent que la monotonicit  du mod le est respect e. Ceux-ci ont  t  obtenus en calculant les d riv s par rapport aux inputs.  tant donn  que la mesure des m decins, des infirmi res et des th rapeutes est en nombre d'individus, on note que ce sont les m decins qui contribuent le plus   la production des services de sant , suivi des th rapeutes et des infirmi res. Les autres inputs ne peuvent  tre compar s en termes de valeurs des d riv es  tant donn  que les unit s ne sont pas d finies sur une base commune.

4.1.3 Mesures de la technologie et de l'inefficience

Les rendements d' chelle permettent d'optimiser la taille des syst mes de sant  provinciaux. Ces valeurs peuvent  tre obtenues avec la somme des d riv s des inputs. Le c t moyen  tant minimis  lorsque les rendements d' chelle sont unitaires, la gestion de l' chelle de production devient un enjeu important dans l'atteinte de l'efficience. Une province dont le rendement d' chelle de ses services de sant  est d croissant, c'est- -dire inf rieur   un, minimiserait son c t moyen suite   une r duction de la production d'output. L'inverse est aussi vrai.

Tableau 4.4 Rendements d'échelle des services de santé

Province	Rendement d'échelle
Alberta	0,78
Colombie-Britannique	0,86
Manitoba	0,98
Nouveau-Brunswick	0,72
Nouvelle-Écosse	1,01
Ontario	0,55
Québec	1,01
Saskatchewan	0,44
Terre-Neuve-et-Labrador	0,65

Dans le cas qui nous intéresse, on remarque au tableau 4.4 que seulement trois provinces, le Manitoba, la Nouvelle-Écosse et le Québec, présentent des rendements d'échelle unitaires. Ce résultat est intéressant étant donné que ces provinces proviennent toutes de régions géographiques différentes et ont des tailles différentes. La réduction du coût moyen ne semble donc pas être affectée par l'emplacement des systèmes de santé, mais plutôt par leur administration. L'autre résultat, plus prévisible, est l'absence de rendements d'échelle croissants. Certaines provinces, dont le Nouveau-Brunswick et Terre-Neuve-et-Labrador, doivent offrir une certaine quantité de services comparable à l'ensemble des provinces canadiennes en dépit d'un bassin de population moins élevé que la moyenne. La rigidité de l'offre pourrait ainsi expliquer certains de ces résultats étant donné la plus faible demande. On peut aussi penser à des indivisibilités qui expliqueraient que la province doit offrir des services complexes et des équipements de grande taille à une population trop faible ou dispersée. On doit cependant conclure que les unités de production des systèmes de santé canadiens sont actuellement trop grandes, favorisant de cette façon une perte d'efficacité.

Le changement technologique représente le déplacement de la frontière de production. Ce résultat s'obtient avec le calcul de la dérivée de la variable de temps. Le déplacement représente l'évolution de la capacité des établissements, à travers le temps, à transformer les inputs en output. De façon générale, on s'attend à ce que la frontière se déplace vers le haut, permettant ainsi une plus grande production d'output. Le tableau 4.5 dresse cependant un portrait négatif du changement technologique à travers le pays.

Tableau 4.5 Mesure du changement technologique des services de santé

Province	Changement technologique
Alberta	-0,013
Colombie-Britannique	-0,010
Manitoba	-0,010
Nouveau-Brunswick	-0,026
Nouvelle-Écosse	-0,028
Ontario	-0,023
Québec	-0,032
Saskatchewan	-0,010
Terre-Neuve-et-Labrador	-0,024

Ce résultat négatif signifie un déplacement de la frontière vers l'origine se manifestant par une baisse de l'output par année. Ce résultat semble à première vue surprenant. Néanmoins, les récentes restrictions budgétaires des gouvernements provinciaux ont incité les établissements de santé à offrir des services de santé avec des budgets moins élevés qu'historiquement. De façon certaine, cela a eu un impact sur la prestation de services en milieu hospitalier. On peut supposer que la qualité des soins n'ait pas été affectée, mais que ces restrictions aient occasionné des délais dans la prestation des services, d'où le déplacement de la frontière vers le bas. Une autre possibilité pourrait être que la main d'œuvre est un input quasi-fixe, c'est-à-dire qu'il est impossible de l'ajuster rapidement à son niveau optimal. Selon nos résultats, l'impact négatif du

changement technologique a affecté les entités administratives de différentes façons. Les provinces du centre ainsi que les provinces de l'Atlantique sont ainsi plus affectées par ce résultat négatif que les Prairies et la Colombie-Britannique.

L'inefficience est la dernière composante auquel on s'intéresse et les données s'y rattachant sont reportées dans le tableau 4.6. Ces données sont calculées à l'aide de la fonction *predict* du logiciel *Stata*.

Tableau 4.6 Mesure de l'inefficience des services de santé

Province	1997	1998	1999	2000	2001	2002
Canada	0,02	0,03	0,02	0,02	0,03	0,04
Alberta	0,05	0,01	0,01	0,01	0,03	0,06
Colombie-Britannique	0,02	0,05	0,01	0,01	0,02	0,02
Manitoba	0,03	0,02	0,01	0,03	0,02	0,04
Nouveau-Brunswick	0,08	0,03	0,02	0,01	0,04	0,03
Nouvelle-Écosse	0,01	0,07	0,05	0,04	0,04	0,05
Ontario	0,02	0,03	0,01	0,01	0,01	0,00
Québec	0,01	0,00	0,01	0,03	0,03	0,05
Saskatchewan	0,01	0,00	0,02	0,01	0,01	0,02
Terre-Neuve-et-Labrador	0,01	0,02	0,01	0,06	0,09	0,14

Province	2003	2004	2005	2006	2007	2008
Canada	0,03	0,03	0,01	0,01	0,03	0,03
Alberta	0,05	0,02	0,01	0,03	0,01	0,04
Colombie-Britannique	0,01	0,01	0,01	0,02	0,01	0,01
Manitoba	0,04	0,03	0,01	0,01	0,02	0,02
Nouveau-Brunswick	0,02	0,01	0,01	0,00	0,01	0,01
Nouvelle-Écosse	0,04	0,02	0,01	0,01	0,05	0,07
Ontario	0,01	0,05	0,02	0,01	0,03	0,01
Québec	0,01	0,02	0,01	0,01	0,04	0,05
Saskatchewan	0,01	0,01	0,01	0,00	0,01	0,02
Terre-Neuve-et-Labrador	0,11	0,10	0,05	0,00	0,04	0,04

Province	2009	2010	2011	2012	2013	2014
Canada	0,02	0,02	0,03	0,04	0,05	0,05
Alberta	0,04	0,05	0,03	0,03	0,02	0,00
Colombie-Britannique	0,01	0,01	0,01	0,03	0,05	0,09
Manitoba	0,01	0,01	0,00	0,01	0,02	0,01
Nouveau-Brunswick	0,01	0,02	0,05	0,06	0,04	0,04
Nouvelle-Écosse	0,01	0,00	0,01	0,07	0,07	0,10
Ontario	0,01	0,01	0,02	0,02	0,06	0,05
Québec	0,05	0,04	0,03	0,04	0,02	0,01
Saskatchewan	0,03	0,04	0,09	0,12	0,14	0,14
Terre-Neuve-et-Labrador	0,02	0,01	0,01	0,01	0,03	0,01

L'inefficience moyenne au pays évolue autour d'une valeur de 0,03 ce qui indique une inefficience de trois points de pourcentage. Ainsi, une situation dans laquelle un système de santé aurait planifié cent consultations, trois ne pourrait être réalisées. L'importance d'atteindre l'efficience découle des répercussions que cela peut avoir sur une population. Il ne semble toutefois pas y avoir de problème d'inefficience démesuré à travers les provinces au cours de la période couverte. Par contre, les niveaux plus élevés d'inefficience prennent du temps avant de se dissiper. Il y a donc une certaine persistance dans l'inefficience. Malgré tout, nous pouvons déceler quelques soubresauts et certaines tendances au fil du temps. On remarque ainsi des niveaux plus élevés d'inefficience lors de la première année de l'échantillon avant de connaître un certain retour à la normale. Cet apaisement est de nouveau suivi par des niveaux plus marqués d'inefficience en 2002. On retrouve une tendance similaire avec le rehaussement des niveaux d'inefficience en 2008. Ces nouveaux sommets étaient précédés d'un certain apaisement. Par contre, l'évolution subséquente demeure plus obscure. Les moins bons résultats de 2008 semblent se répercutés aux années suivantes et les derniers résultats de l'échantillon tendent à démontrent une situation qui pourrait devenir alarmante. Quoi qu'il en soit, la figure 4.1 permet de tirer d'autres conclusions grâce à la mise en contexte des différents niveaux d'inefficience dans le temps.

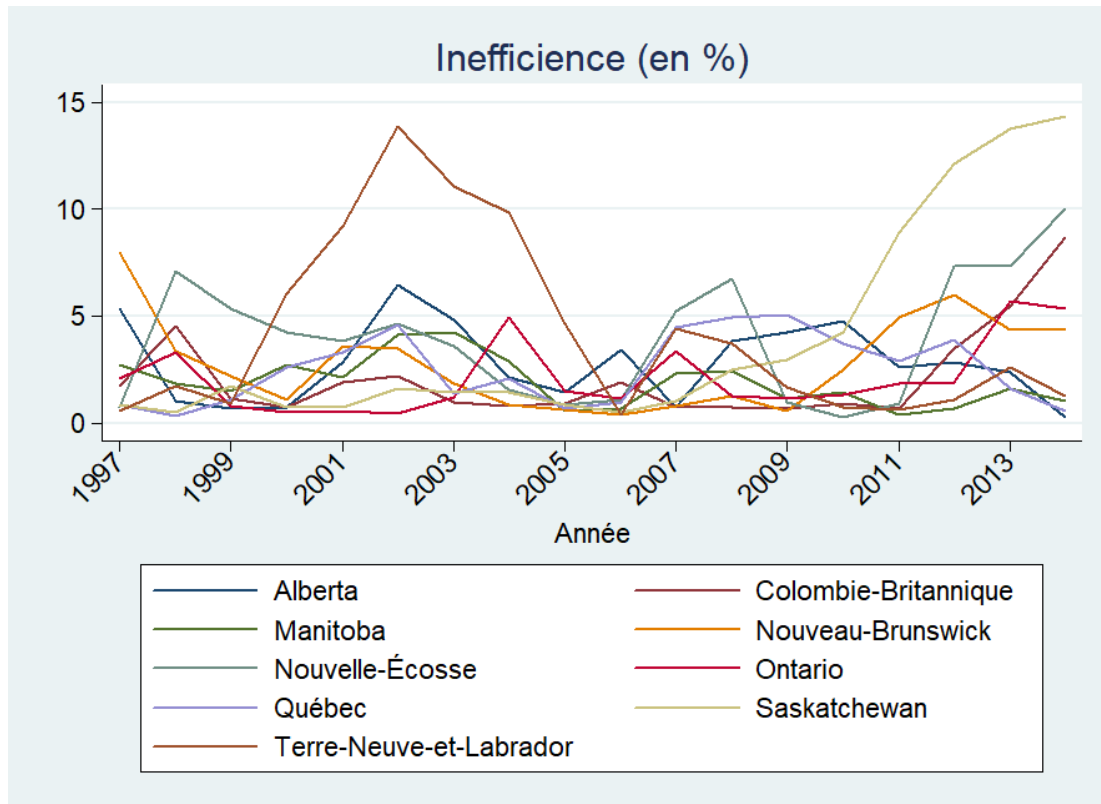


Figure 4.1 Évolution de l'inefficacité des services de santé

La comparaison de l'efficacité étant le principal objet d'étude de ce mémoire, les provinces ayant des situations d'inefficacité qui divergent de ce qui est jugé optimal attirent notre attention. Deux situations, que l'on peut qualifier d'anormales, sont observées à Terre-Neuve-et-Labrador en 2002 ainsi qu'en Saskatchewan à partir de 2011. Pour Terre-Neuve-et-Labrador, la situation de l'inefficacité s'explique par une grève de dix-sept jours des médecins en octobre 2002. Cette grève, dont l'enjeu était les hausses salariales des médecins pour qu'elles soient équivalentes à celle des autres provinces de l'Atlantique, a touché 930 médecins. Cela s'est traduit par la cessation des rendez-vous ainsi que des traitements médicaux jugés non nécessaires durant cette période. Au total, ce sont 30 000 consultations ou traitements qui ont dû être reportés. En ce qui concerne la Saskatchewan, une série de grèves concernant le personnel en santé

a touché la province à partir de 2010 avec notamment une grève de sept semaines en 2011 touchant 1000 professionnels en santé ainsi qu'une grève en 2014 touchant 3000 spécialistes d'environ trente professions. Encore une fois, ces moyens de pression ont eu un effet négatif sur la production de services de santé étant donné le report de plusieurs services. On peut donc considérer les niveaux présentés comme étant normaux, mais une attention particulière devra être portée pour les années plus récentes non couvertes par l'échantillon.

En dernier lieu, la monétisation de l'inefficience nous permet d'évaluer les pertes monétaires liées à l'inefficience. Au moyen d'années types, cette présentation met en évidence l'évolution des sommes perdues au sein des différentes provinces. Le calcul pour y parvenir s'est effectué à l'aide des dépenses nationales en santé, qui recensent chaque année les dépenses de santé de chaque province. Le produit de l'inefficience et de ces données permet de dresser le portrait précis des pertes monétaires propre à chaque province au fil du temps. Lorsque l'on examine les différents montants (Tableau 4.7), on remarque que certaines provinces s'en tirent moins bien que d'autres. Ainsi, la Colombie-Britannique, l'Ontario et la Saskatchewan sont les provinces qui ont subi les pertes les plus importantes dans leur système de santé respectif à la fin de l'échantillon en 2014. Fait intéressant, ces provinces ont toutes réussi à contrôler l'évolution des sommes perdues en milieu d'échantillon avant que la situation ne se détériore à nouveau. Le Nouveau-Brunswick ainsi que la Nouvelle-Écosse ont quant à eux connu une hausse des sommes perdues mais à un rythme moins alarmant. Certaines provinces ont toutefois réussi à contrôler les pertes liées à l'inefficience au cours de la dernière année. C'est le cas notamment avec l'Alberta, le Manitoba, le Québec ainsi que Terre-Neuve-et-Labrador. Cela est même un retour intéressant pour le Québec qui a essuyé des pertes de près de deux milliards de dollars en 2008. Néanmoins, la tendance de l'évolution des sommes perdues en santé au pays est claire, il y a une nette augmentation depuis les deux dernières décennies. Passant d'un peu plus de deux

milliards de dollars au début de l'échantillon en 1997 à un peu plus de neuf milliards de pertes en 2014. Au cours de ces dix-huit années, les sommes perdues totalisent 77 milliards de dollars avec des sommets en Ontario et au Québec. Logiquement, étant donné les bassins de population respective de ces provinces, celles du centre sont les plus touchées par ces pertes monétaires. Le classement se poursuit avec les Prairies et l'Ouest canadien, avant d'être complété par les provinces ayant de plus faibles populations, soit les provinces maritimes.

Tableau 4.7 Quantification de l'inefficience des services de santé selon des années types (en millions de dollars réels)

Province	1997	2003	2008	2014	Total
Canada	2 351,7	3 041,5	4 876,4	9 326,8	77 701,6
Alberta	582,0	786,4	863,9	75,8	9 987,7
Colombie-Britannique	242,0	192,0	171,0	2 464,1	8 734,8
Manitoba	118,1	264,4	186,7	88,5	2 220,2
Nouveau-Brunswick	211,2	63,5	57,1	211,5	1 969,4
Nouvelle-Écosse	24,4	164,3	381,1	619,9	3 680,0
Ontario	915,3	763,7	991,6	4 483,1	27 494,3
Québec	214,5	439,7	1 952,0	263,2	16 743,3
Saskatchewan	33,3	73,5	155,8	1 073,3	4 841,3
Terre-Neuve-et-Labrador	10,5	293,8	116,8	46,9	2 030,3

4.2 Niveaux de santé en fonction des services de santé

4.2.1 Résultats de l'estimation

Le deuxième modèle fait référence à l'équation (3). Il conserve l'approche désagrégée en estimant les niveaux de santé par rapport aux services de santé précédemment estimés. Ce modèle intègre deux outputs, l'espérance de vie à 65 ans et le taux de survie infantile, et les inputs provenant du système de santé sont ceux du modèle

précédent, le nombre de consultations et d'interventions. On retrouve aussi l'indice de qualité de l'output, la durée moyenne du séjour hospitalier. L'ajout de variables socio-démographiques permet de caractériser l'impact des habitudes de vie des citoyens sur les niveaux de santé.

Concernant les termes croisés, un premier objectif est de caractériser l'implication des différents déterminants socio-démographiques sur la durée moyenne du séjour hospitalier. Une saine alimentation et une consommation réduite en alcool et en tabac devraient en théorie faciliter un rétablissement et ainsi réduire la durée moyenne d'un séjour hospitalier. Par la suite, étant donné l'importance de nos déterminants socio-démographiques, ceux-ci ont été élevés au carré pour inclure certaines non-linéarités. L'élévation du taux de survie infantile au carré combiné à son implication sur la durée moyenne du séjour hospitalier vient finalement caractériser la qualité des soins obstétricaux et pédiatriques des différentes provinces.

Avant de poursuivre, nous devons une fois de plus nous assurer de la significativité de l'erreur unilatère. À nouveau, il n'y a pas de problème à ce niveau. Tel qu'en témoigne la valeur p du coefficient de l'erreur unilatère (Tableau 4.8), celui-ci est significatif à un seuil de 1 %. Au tableau 4.9, on constate, à l'aide du test de Hausman, qu'on ne peut rejeter le modèle à effets aléatoires pour faire l'estimation des niveaux de santé. Toutefois, dans un objectif de cohérence interne, nous continuerons d'utiliser un modèle à effets fixes pour nos estimations.

Tableau 4.8 Coefficients de la frontière des niveaux de santé

Espérance de vie à 65 ans	Coefficient	Valeur <i>p</i>
Output		
Taux de survie infantile	-0,062	0,043
Inputs		
Consultations et interventions	0,078	0,059
Durée moyenne du séjour hospitalier	-7,755	0,124
Déterminants socio-démographiques		
Taux de chômage	0,005	0,799
Indice de quantité d'alimentation	29,011	0,016
Indice de quantité d'alcool	-1,188	0,235
Indice de quantité de tabac	-1,506	0,028
Termes croisés		
Taux de survie infantile ²	-0,003	0,058
Durée moyenne du séjour * Taux de survie infantile	0,023	0,082
Durée moyenne du séjour * Taux de chômage	0,001	0,882
Durée moyenne du séjour * Indice en alimentation	1,280	0,082
Durée moyenne du séjour * Indice en alcool	-0,373	0,198
Durée moyenne du séjour * Indice en tabac	0,064	0,642
Taux de chômage ²	-0,001	0,099
Indice de quantité d'alimentation ²	-4,090	0,015
Indice de quantité d'alcool ²	0,338	0,121
Indice de quantité de tabac ²	0,208	0,015
σ_{μ}	0,014	0,000
σ_{ν}	0,009	0,000
λ	1,637	0,000

Tableau 4.9 Test de Hausman sur la production des niveaux de santé

$$\begin{aligned}
 \chi^2(11) &= (\beta_{TFE} - \beta_{TRE})' [\sigma^2(\beta_{TFE}) - \sigma^2(\beta_{TRE})]^{-1} (\beta_{TFE} - \beta_{TRE}) \\
 &= 4,09 \\
 \mathbb{P} > \chi^2(11) &= 0,9673
 \end{aligned}$$

4.2.2 Tests de la théorie

Tel que mentionné auparavant, l'utilisation des résultats exige le respect des hypothèses justifiant l'utilisation la théorie économique. Dans ce modèle, non seulement faut-il tenir compte du respect des propriétés de monotonicité positive en inputs, mais nous devons aussi prendre en considération le respect de la monotonicité négative dans le deuxième output. En d'autres mots, l'allocation d'inputs pour la production d'un deuxième output, on parle ici du taux de survie infantile, se fera au détriment de notre output principal qui est l'espérance de vie à 65 ans. Cela explique la relation négative qui doit exister entre ces deux outputs. En dernier lieu, le respect de la théorie se conclut avec les monotonicités des déterminants socio-démographiques. La nature du signe doit correspondre à la nature des déterminants.

Tableau 4.10 Monotonicité positive des inputs en santé et négative du taux de survie

Variable	Moyenne
Output	
Taux de survie infantile	-0,002
Inputs	
Consultations et interventions	0,078
Durée moyenne du séjour hospitalier	0,131
Déterminants socio-démographiques	
Taux de chômage	0,002
Indice de quantité d'alimentation	0,030
Indice de quantité d'alcool	0,149
Indice de quantité de tabac	-0,091

On constate au tableau 4.10 que les monotonicités par rapport à l'output et aux inputs sont respectées même si deux déterminants socio-démographiques n'obtiennent pas le signe espéré. On s'attend à ce que le taux de chômage ainsi que les consommations en alcool et en tabac obtiennent des signes négatifs tandis qu'on entrevoit un signe

positif pour l'alimentation. Le taux de chômage et la consommation d'alcool ont cependant le mauvais signe, mais puisque la valeur p de ces variables dépasse un seuil de significativité de 10 %, on peut rejeter la significativité des signes positifs des monotonicités du taux de chômage et de l'indice de quantité d'alcool. On peut donc conclure que la théorie n'est pas rejetée et continuer l'analyse des résultats suivants.

4.2.3 Mesures de la technologie et de l'inefficience

Une fois de plus, les rendements d'échelle nous informent sur la taille optimale des systèmes de santé, mais cette fois en ce qui concerne les niveaux de santé. À l'aide de ce concept, les économies d'échelle sont maximisées lorsque les rendements d'échelle sont unitaires. Lorsque ce n'est pas le cas, il suffit de réduire la production d'output lorsque les rendements d'échelle sont inférieurs à un et d'augmenter cette production lorsque les rendements d'échelle sont supérieurs à un pour parvenir à l'efficience.

Tableau 4.11 Rendements d'échelle des niveaux de santé

Province	Rendement d'échelle
Alberta	0,27
Colombie-Britannique	0,26
Manitoba	0,18
Nouveau-Brunswick	0,22
Nouvelle-Écosse	0,20
Ontario	0,13
Québec	0,27
Saskatchewan	0,20
Terre-Neuve-et-Labrador	0,14

Le portrait que dresse le tableau 4.11 sur les rendements d'échelle des niveaux de santé est plutôt sombre. Dans la mesure où les neuf provinces obtiennent des rendements

d'échelle fortement décroissants, c'est donc dire que la production d'une année de vie supplémentaire à l'âge de 65 ans se fait à un coût très élevé. Toutefois, étant donné que la durée de vie humaine est contrainte par une limite biologique, il peut être normal de constater ces faibles rendements d'échelle du fait qu'une année de vie supplémentaire est difficile à atteindre. On se retrouve donc vraisemblablement dans la région où la fonction de production est quasi-plate.

L'absence du changement technologique s'explique par le retrait des variables associées au temps afin d'assurer la convergence du modèle. Nous nous tournerons donc vers l'analyse de l'efficacité pour nous informer au sujet des pertes résultant de la transformation des inputs en outputs. Le tableau 4.12 montre l'inefficacité par province et par année pour l'ensemble de l'échantillon.

Tableau 4.12 Mesure de l'inefficacité des niveaux de santé

Province	1997	1998	1999	2000	2001	2002	2003	2004
Canada	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
Alberta	0,01	0,03	0,03	0,01	0,01	0,01	0,01	0,01
Colombie-Britannique	0,01	0,00	0,01	0,00	0,01	0,01	0,02	0,01
Manitoba	0,01	0,03	0,02	0,02	0,01	0,01	0,01	0,02
Nouveau-Brunswick	0,01	0,03	0,01	0,01	0,01	0,01	0,02	0,01
Nouvelle-Écosse	0,01	0,01	0,00	0,00	0,01	0,01	0,01	0,01
Ontario	0,01	0,01	0,01	0,01	0,01	0,02	0,02	0,01
Québec	0,03	0,02	0,01	0,00	0,01	0,01	0,01	0,01
Saskatchewan	0,00	0,01	0,01	0,01	0,01	0,01	0,02	0,01
Terre-Neuve-et-Labrador	0,01	0,01	0,01	0,03	0,01	0,00	0,01	0,01

Province	2005	2006	2007	2008	2009	2010	2011	2012
Canada	0,01	0,01	0,06	0,01	0,01	0,04	0,01	0,01
Alberta	0,00	0,00	0,04	0,01	0,02	0,02	0,00	0,01
Colombie-Britannique	0,01	0,01	0,07	0,03	0,01	0,06	0,01	0,01
Manitoba	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00
Nouveau-Brunswick	0,02	0,01	0,03	0,01	0,01	0,02	0,00	0,00
Nouvelle-Écosse	0,01	0,01	0,03	0,01	0,02	0,03	0,01	0,01
Ontario	0,01	0,01	0,05	0,01	0,02	0,04	0,00	0,00
Québec	0,01	0,01	0,08	0,01	0,01	0,05	0,01	0,00
Saskatchewan	0,01	0,01	0,04	0,01	0,02	0,05	0,00	0,01
Terre-Neuve-et-Labrador	0,01	0,01	0,04	0,02	0,00	0,02	0,00	0,01

À nouveau, l'évolution de l'inefficience et la comparaison entre les provinces demeurent nos intérêts principaux. Une analyse par année démontre que le Québec, l'Alberta, le Manitoba et le Nouveau-Brunswick connaissent des niveaux d'inefficience plus élevés que la moyenne nationale dans les premières années de l'échantillon. Cela n'est toutefois pas alarmant puisque les niveaux d'inefficience sont relativement faibles. Concernant les niveaux d'inefficience plus élevés de 2007 et 2010, on se doit de demander d'où provient la source de ces bonds. À l'exception du Manitoba, on remarque une baisse de l'espérance de vie à 65 ans dans l'ensemble des provinces pour les deux années qui nous intéressent. La différence est notable et peut même atteindre plus d'une année de vie pour le Québec et la Colombie-Britannique. On note aussi une baisse considérable pour l'Ontario. Face à des inputs dont les niveaux sont légèrement croissants, voire constants, l'accroissement de l'inefficience s'explique donc par un recul dans la production de l'output. La figure 4.2 permet de suivre l'évolution de l'inefficience avec une représentation en année de vie perdue. À partir de nos données, on ne peut déterminer ce qui explique cette baisse d'une année d'espérance de vie.

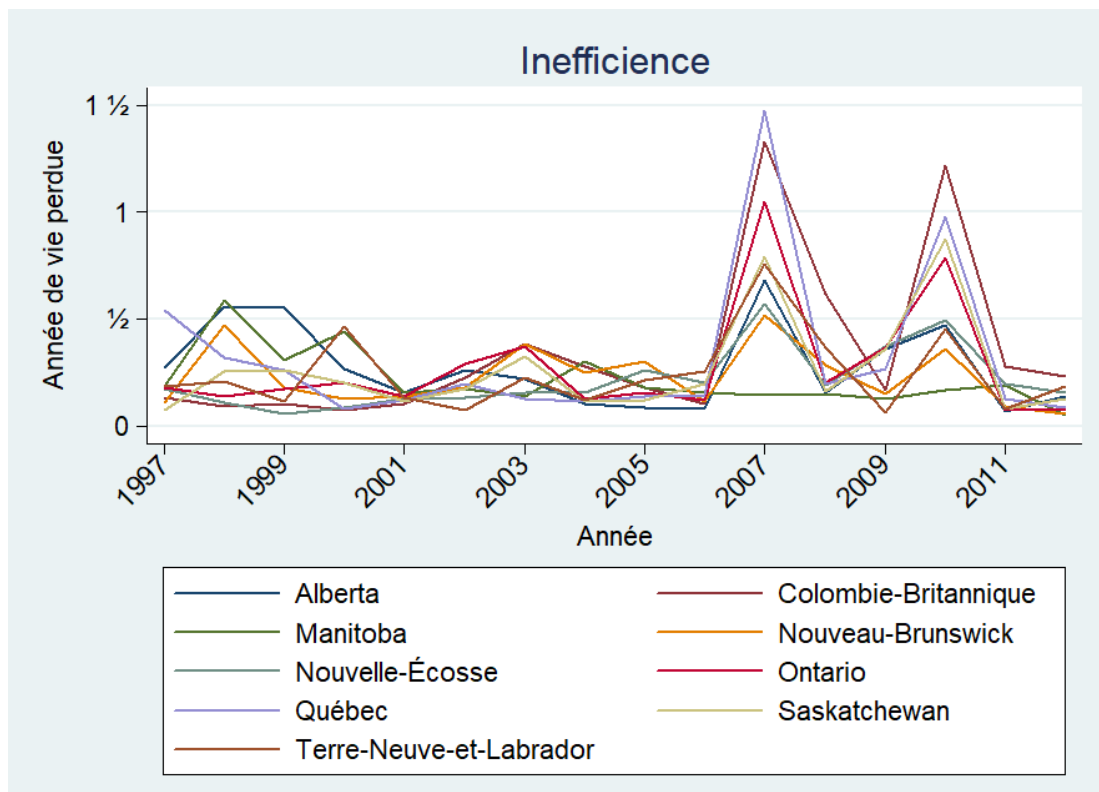


Figure 4.2 Évolution de l'inefficience des niveaux de santé

La multiplication de l'inefficience et de l'espérance de vie à 65 ans permet de bien cerner les enjeux liés à l'importance des niveaux de santé. Les niveaux plus élevés d'inefficience au début de l'échantillon se sont traduits par une plus grande perte d'année de vie. Ces résultats affectent directement la population. De plus, la hausse des niveaux d'inefficience en 2007 et 2010 s'accompagne d'une perte d'espérance de vie de plus d'une année. Une comparaison entre les provinces montre qu'encore une fois, ce sont celles dont le bassin de population est le plus élevé qui sont surtout désavantagées. L'Alberta et la Saskatchewan, ainsi que les provinces de l'Atlantique, ont des résultats semblables et près de la moyenne nationale. Le Manitoba semble être la province qui s'en tire le mieux. Le résultat de cette province, au même titre que celui de la production de services de santé au premier modèle, est remarquable. Le Manitoba apparaît donc comme un exemple à suivre.

4.3 Niveaux de santé en fonction des ressources en santé

4.3.1 Résultats de l'estimation

Notre dernier modèle estime les niveaux de santé en une seule étape telle que présentée à l'équation (1). Par conséquent, nous excluons l'étape intermédiaire qui est la production de services de santé et utilisons directement les ressources investies en santé comme inputs. Cette façon de faire nous permettra d'examiner si l'agrégation vient changer les résultats précédemment obtenus de façon désagrégée. Pour permettre une comparabilité des résultats, l'espérance de vie à 65 ans est de nouveau utilisée et le taux de survie infantile, qui est notre deuxième output, est conservé. De plus, les inputs utilisés lors du premier modèle sont désormais combinés aux déterminants socio-démographiques. Ceci permet de caractériser les résultats en santé des différentes provinces au même titre que précédemment.

Le premier constat que l'on peut faire (Tableau 4.13) est qu'il y a plus de coefficients non-significatifs avec cette approche. On peut en effet rejeter que la majorité des coefficients estimés sont significatifs à un seuil de 10 %. Au même titre que précédemment, la valeur p du coefficient de l'erreur unilatère atteste de sa significativité selon un seuil de 5 %.

Tableau 4.13 Coefficients de la frontière des niveaux de santé

Espérance de vie à 65 ans	Coefficient	Valeur p
Output		
Taux de survie infantile	-0,002	0,095
Inputs		
Médecins	-0,472	0,533
Infirmières autorisées et auxiliaires autorisées	-0,003	0,942
Physiothérapeutes et ergothérapeutes	-0,606	0,347
Stock de bâtiments	0,006	0,533
Stock de machines	0,004	0,445
Indice de quantité de médicaments	0,054	0,083
Indice de quantité de santé publique	-0,002	0,897
Indice de quantité d'administration	-0,001	0,939
Déterminants socio-démographiques		
Taux de chômage	-0,001	0,495
Indice de quantité d'alimentation	-0,141	0,138
Indice de quantité d'alcool	0,008	0,767
Indice de quantité de tabac	-0,074	0,000
Terme croisé		
Médecins * Thérapeutes	-0,080	0,449
Changement technologique		
Temps	0,003	0,531
Temps ²	0,000	0,598
σ_{μ}	0,016	0,000
σ_{ν}	0,004	0,011
λ	4,623	0,000

Tableau 4.14 Test de Hausman sur la production des niveaux de santé

$$\begin{aligned}
 \chi^2(15) &= (\beta_{TFE} - \beta_{TRE})' [\sigma^2(\beta_{TFE}) - \sigma^2(\beta_{TRE})]^{-1} (\beta_{TFE} - \beta_{TRE}) \\
 &= 9,45 \\
 \mathbb{P} > \chi^2(15) &= 0,8531
 \end{aligned}$$

À partir du test de Hausman (Tableau 4.14), on ne peut rejeter le modèle à effets aléatoires, mais tout comme au modèle précédent, l'emploi d'un modèle à effets fixes demeure notre choix de prédilection. Les raisons demeurent les mêmes qu'auparavant, soit pour conserver une certaine cohérence entre les modèles. Les trois modèles reposent donc sur des effets fixes.

4.3.2 Tests de la théorie

Les prochains tests concernant le respect de la théorie économique vont nous renseigner sur la crédibilité d'un modèle agrégé. On doit donc s'assurer du respect des propriétés de monotonie positive en inputs et négative en output.

Tableau 4.15 Monotonie des variables des niveaux de santé

Variable	Moyenne
Output	
Taux de survie infantile	-0,002
Inputs	
Médecins	0,090
Infirmières autorisées et auxiliaires autorisées	-0,003
Physiothérapeutes et ergothérapeutes	-0,110
Stock de bâtiments	0,006
Stock de machines	0,004
Indice de quantité de médicaments	0,054
Indice de quantité de santé publique	-0,002
Indice de quantité d'administration	-0,001
Déterminants socio-démographiques	
Taux de chômage	-0,001
Indice de quantité d'alimentation	-0,141
Indice de quantité d'alcool	0,008
Indice de quantité de tabac	-0,074

Le compte rendu du tableau 4.15 témoigne du non-respect des propriétés de certaines monotonicités. Les signes contraires à ceux attendus proviennent des infirmières et des thérapeutes ainsi que l'indice de quantité de santé publique et d'administration. Il est inconcevable de reconnaître que ces ressources puissent avoir un impact négatif sur les niveaux de santé. Nous devons conclure que la théorie n'est pas respectée. Par ailleurs, deux déterminants socio-démographiques, la quantité d'alimentation et la quantité d'alcool, ne possèdent pas non plus les bons signes. La quantité d'alcool est significativement du mauvais signe. Les raisons expliquant ces résultats, et pour lesquelles on peut démontrer qu'effectuer l'analyse en santé de façon agrégée n'est pas une bonne façon de procéder, seront couvertes au cours de la prochaine section.

4.3.3 Mesures de la technologie et de l'inefficience

Malgré l'absence du respect de la théorie pour le cas qui nous intéresse, nous poursuivons avec l'analyse des rendements d'échelle. Les résultats permettront tout de même une certaine comparaison avec l'estimation des niveaux de santé de façon désagrégée.

Tableau 4.16 Rendements d'échelle des niveaux de santé

Province	Rendement d'échelle
Alberta	0,03
Colombie-Britannique	0,03
Manitoba	0,04
Nouveau-Brunswick	0,04
Nouvelle-Écosse	0,03
Ontario	0,05
Québec	0,02
Saskatchewan	0,06
Terre-Neuve-et-Labrador	0,05

On constate cette fois-ci au tableau 4.16 que les rendements d'échelle sont encore plus petits que précédemment. Nous ne sommes cependant pas étrangers à ces résultats puisque l'on se retrouve encore une fois à l'endroit où la fonction de production est quasi-plate. En revanche, la différence concernant les rendements d'échelle entre les deux modèles qui estiment les niveaux de santé est notable. Cela implique que la relation entre un service en santé et une ressource en santé est inégale. On peut expliquer cet écart par la quantité d'inputs utilisées pour permettre le prolongement d'une année de vie supplémentaire. Nous pouvons dans un premier temps supposer qu'une intervention importante puisse prolonger la durée de vie d'un patient. Par la suite, lorsqu'on analyse les implications de cette intervention, on remarque que celle-ci nécessite beaucoup de ressources. L'occupation d'un bloc opératoire, d'un médecin ainsi que des infirmières assurant le bon fonctionnement, sans compter une possible prise de médicaments pour assurer la guérison, est à prendre en compte. Il faudra donc obligatoirement plus de ressources pour produire un service, d'où le coût plus important et le rendement d'échelle plus bas.

Étant donné l'absence de termes croisés avec le temps dans ce modèle, le changement technologique est identique pour chaque province. Néanmoins, ce qui importe est que le résultat est cette fois positif, bien que très faible, ce qui implique que le déplacement de la frontière de production s'effectue vers le haut.

Tableau 4.17 Mesure du changement technologique des niveaux de santé

Province	Changement technologique
Alberta	0,005
Colombie-Britannique	0,005
Manitoba	0,005
Nouveau-Brunswick	0,005
Nouvelle-Écosse	0,005
Ontario	0,005
Québec	0,005
Saskatchewan	0,005
Terre-Neuve-et-Labrador	0,005

La valeur moyenne du tableau 4.17 nous indique une augmentation de 0,5 point de pourcentage de l'output par année, ce qui est crédible. Lorsque l'on convertit ce résultat en output, cela se traduit par une augmentation de l'espérance de vie à 65 ans d'un mois de vie supplémentaire. Nous pouvons constater ces progrès au fil du temps. On s'aperçoit ainsi que l'espérance de vie a crû considérablement au fil des années et cette progression pourrait se poursuivre. Le changement technologique permet donc de prolonger les niveaux de vie. Il ne faut toutefois pas cependant oublier les limites biologiques qui nous sont imposées.

Finalement, l'inefficience est une fois de plus à l'étude pour ce dernier modèle. Le tableau 4.18 nous montre les différents niveaux d'inefficience pour l'ensemble des provinces à l'étude.

Tableau 4.18 Mesure de l'inefficience des niveaux de santé

Province	1997	1998	1999	2000	2001	2002	2003	2004
Canada	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,01
Alberta	0,01	0,01	0,02	0,01	0,00	0,02	0,02	0,01
Colombie-Britannique	0,01	0,01	0,01	0,00	0,00	0,01	0,02	0,01
Manitoba	0,00	0,02	0,02	0,01	0,00	0,00	0,01	0,01
Nouveau-Brunswick	0,01	0,03	0,01	0,00	0,00	0,01	0,01	0,01
Nouvelle-Écosse	0,02	0,02	0,00	0,00	0,00	0,01	0,01	0,01
Ontario	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,00
Québec	0,04	0,03	0,03	0,01	0,01	0,02	0,00	0,01
Saskatchewan	0,00	0,01	0,01	0,01	0,00	0,01	0,01	0,01
Terre-Neuve-et-Labrador	0,05	0,04	0,00	0,03	0,01	0,00	0,01	0,02

Province	2005	2006	2007	2008	2009	2010	2011	2012
Canada	0,01	0,00	0,06	0,01	0,01	0,05	0,01	0,01
Alberta	0,01	0,00	0,05	0,02	0,03	0,05	0,02	0,04
Colombie-Britannique	0,01	0,01	0,08	0,03	0,00	0,07	0,01	0,01
Manitoba	0,01	0,01	0,01	0,01	0,00	0,02	0,04	0,00
Nouveau-Brunswick	0,02	0,01	0,04	0,02	0,01	0,02	0,01	0,00
Nouvelle-Écosse	0,02	0,01	0,03	0,00	0,01	0,02	0,01	0,02
Ontario	0,01	0,00	0,05	0,00	0,01	0,05	0,00	0,00
Québec	0,01	0,00	0,08	0,00	0,00	0,07	0,01	0,01
Saskatchewan	0,01	0,01	0,05	0,01	0,02	0,05	0,00	0,02
Terre-Neuve-et-Labrador	0,03	0,03	0,07	0,03	0,00	0,04	0,00	0,01

Les résultats de l'inefficience de ce dernier modèle se comparent au modèle des niveaux de santé de façon désagrégée. On y retrouve passablement les mêmes tendances, mais avec plus de volatilité. De par leurs niveaux plus importants d'inefficience comparativement à la moyenne canadienne, deux provinces se démarquent en début d'échantillon. C'est le cas du Québec ainsi que de Terre-Neuve-et-Labrador. Si le résultat de Terre-Neuve-et-Labrador contraste avec le modèle précédent, on retrouve

sensiblement les mêmes niveaux concernant le Québec. Le retour à des niveaux d'inefficience plus bas est la norme au début des années 2000 avant une hausse subite en 2007. Les mêmes provinces que précédemment, la Colombie-Britannique, l'Ontario et le Québec, soit celles ayant une importante démographie, sont les plus touchés par cette montée d'inefficience. On note aussi des niveaux plus importants qu'à l'habitude pour Terre-Neuve-et-Labrador au milieu des années 2000. Tel que présenté pour le deuxième modèle, la figure 4.3 permet de constater l'inefficience en termes d'année de vie perdue.

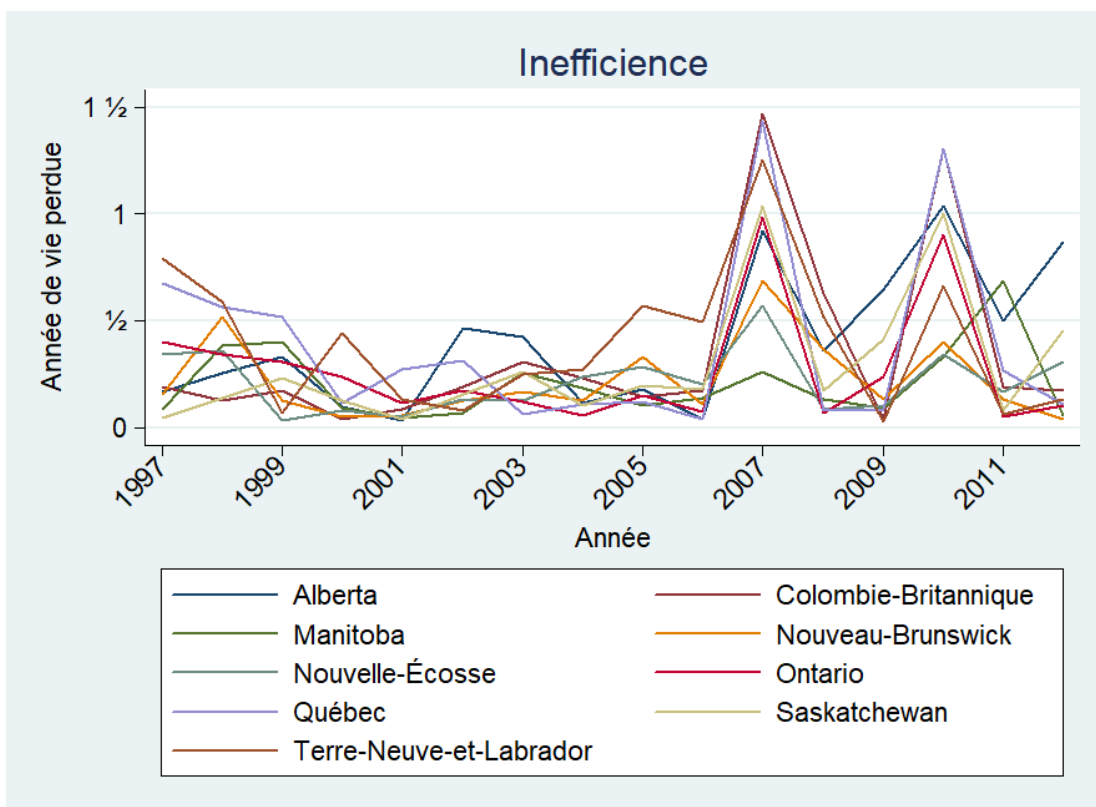


Figure 4.3 Évolution de l'inefficience des niveaux de santé

Les résultats en année de vie perdue se rapprochent grandement de ceux du deuxième modèle. On peut en tirer les mêmes conclusions. Les années 2007 et 2010 ont été la cible d'inefficience étant donné la baisse généralisée de l'espérance de vie à 65

ans des provinces canadiennes. Par ailleurs, la quantité d'années de vie perdues en Alberta pour les années 2002 et 2003 est plus importante ici que dans le modèle désagrégé. Finalement, l'inefficience semble rebondir pour certaines provinces en 2012. C'est le cas notamment de l'Alberta qui souffre davantage d'années de vie perdues comparativement aux autres provinces depuis 2010. Ce semble être aussi le cas de la Saskatchewan ainsi que de la Nouvelle-Écosse. Nous pouvons conclure cette section en mentionnant que le Manitoba fait encore une fois bonne figure lors de cette quantification de l'inefficience.

4.4 Explication du rejet de l'agrégation

Comme nous l'avons vu, les résultats du modèle en une étape sont mauvais ou au mieux comparables à ceux du modèle en deux étapes. La raison pour laquelle l'agrégation des modèles n'est pas une bonne façon de faire l'analyse en santé est en partie liée aux hypothèses stochastiques. Dans le cas qui nous intéresse, nos trois modèles peuvent être représentés sous la forme :

$$\text{Modèle 1 : } y = f_1(x) + \mu_1 - \nu_1$$

$$\text{Modèle 2 : } z = f_2(y, w) + \mu_2 - \nu_2$$

$$\text{Modèle 3 : } z = f_3(x, w) + \mu_3 - \nu_3$$

où x est le vecteur des ressources en santé, y , le vecteur des services de santé, z , le vecteur des niveaux de santé et w , le vecteur des déterminants socio-démographiques. La substitution du modèle 1 dans le modèle 2 donne :

$$\text{Modèle 3 : } z = f_2(f_1(x) + \mu_1 - \nu_1, w) + \mu_2 - \nu_2$$

Pour simplifier l'argument, nous supposons que les modèles sont linéaires. Nous obtenons :

$$\text{Modèle 1 : } y = a'x + \mu_1 - \nu_1$$

$$\text{Modèle 2 : } z = by + c'w + \mu_2 - \nu_2$$

L'agrégation des modèles implique donc :

$$\begin{aligned} \text{Modèle 3 : } z &= b(a'x + \mu_1 - \nu_1) + c'w + \mu_2 - \nu_2 \\ &= b * a'x + b * \mu_1 - b * \nu_1 + c'w + \mu_2 - \nu_2 \\ &= b * a'x + c'w + (b * \mu_1 + \mu_2) - (b * \nu_1 + \nu_2) \end{aligned}$$

Le dernier modèle est cependant estimé comme étant :

$$\text{Modèle 3 : } z = d'x + g'w + \mu_3 - \nu_3$$

Il faut donc en toute logique que :

$$b * a = d$$

$$c = g$$

$$b * \mu_1 + \mu_2 = \mu_3$$

$$b * \nu_1 + \nu_2 = \nu_3$$

Les deux premières conditions sont des conditions de cohérence entre les deux premiers modèles et le troisième modèle. Il est facile de vérifier en comparant les tableaux de coefficients que ces conditions ne sont pas respectées. Non seulement la valeur, mais aussi le signe des coefficients et leurs valeurs p sont différents.

Ce problème trouve sa source dans la nature des termes d'erreur découlant des troisième et quatrième conditions. L'estimation du troisième modèle suppose que ν_3 est distribué normalement et indépendamment des coefficients du modèle, dont le vecteur b , ce qui n'est pas le cas. L'argument qui précède vaut aussi pour μ_3 ce qui suppose un troisième modèle donnant de mauvais résultats. On comprend ainsi les raisons pour lesquelles l'analyse des résultats du modèle en une étape n'était pas concluante. La conclusion qui s'impose est que l'agrégation des modèles n'est pas une bonne façon d'effectuer l'analyse en santé. Pour obtenir des résultats cohérents, seule la méthode désagrégée prévaut.

CONCLUSION

Les dépenses en santé ont depuis longtemps commencé à croître et cette croissance n'est pas prête de s'arrêter. Dans un contexte où l'âge moyen de la population augmente, il nous faut utiliser efficacement nos ressources pour nous assurer de répondre à la demande grandissante de soins de santé. Cela s'intègre dans un cadre où la capacité fiscale des gouvernements est déjà limitée. L'allocation des ressources est donc au cœur du problème et l'on est en droit de se demander si l'allocation actuelle est efficiente.

Bien comprendre les causes de l'inefficience en santé est pertinent étant donné les répercussions directes sur les niveaux de vie de la population. Si les travaux antérieurs ne s'attardent pas spécifiquement à la problématique de l'inefficience entre les provinces canadiennes, la connaissance des causes expliquant les différents niveaux d'efficience au pays demeure nécessaire. Il importe d'inclure les disparités qui existent entre les provinces dans le vaste territoire qu'est le Canada, d'autant plus que la santé est de juridiction provinciale. Ce mémoire s'insère dans cette réflexion et vise à mesurer puis à quantifier l'inefficience dans une perspective comparative.

Plusieurs méthodes sont envisageables pour étudier les mesures d'efficience. Nous pouvons répartir ces méthodes en deux catégories avec les méthodes paramétriques et non-paramétriques. Compte tenu des avancées récentes en matière de programmation, la méthode paramétrique a été retenue. Celle-ci porte le nom *Stochastic Frontier Analysis* et a été combinée au modèle des effets fixes de Greene (2004 et 2005). L'élaboration de frontières stochastiques nous a permis de quantifier l'inefficience et les rendements d'échelle des différents systèmes de santé provinciaux.

D'une part, nous avons pu confirmer la validité des modèles désagrégés avec le respect des propriétés théoriques. Ce ne fut pas le cas avec le modèle agrégé. Nous devons donc conclure que l'approche agrégée ne peut être utilisée pour effectuer l'analyse en santé. Les raisons derrière ce rejet proviennent de la nature des termes d'erreur du modèle. L'analyse en santé doit donc passer par une approche en deux étapes. Pour cette raison, dans ce qui suit, nous ne retiendrons que les résultats du modèle en deux étapes.

D'autre part, les mesures de la technologie nous informent sur la variation de la production des centres hospitaliers. Les rendements d'échelle nous renseignent dans un premier temps de la taille optimale des institutions quant à la production des services de santé. Si trois provinces obtiennent de bons résultats pour cette mesure avec des rendements d'échelle unitaires, cela n'est pas un cas généralisé. Ainsi, à l'exception du Manitoba, de la Nouvelle-Écosse et du Québec pour notre premier modèle, les autres provinces obtiennent des rendements d'échelle décroissants, augmentant par conséquent les coûts totaux. Le changement technologique nous renseigne dans un deuxième temps sur le déplacement de la frontière de production. On s'attend normalement à un déplacement vers le haut, mais cela ne fut pas le cas pour la production de services de santé. Les récentes coupures dans les différents budgets liés à la santé et les rigidités intrinsèques au système de santé en sont probablement la cause.

En dernier lieu, les niveaux d'inefficience sont pour la plupart acceptables. On observe des niveaux d'inefficience relativement bas et semblables entre les provinces. Deux situations plus alarmantes se sont produites à Terre-Neuve-et-Labrador en 2002 et en Saskatchewan à partir de 2011. Un examen plus approfondi de ces situations démontre que les plus hauts taux d'inefficience sont attribuables à des interruptions de travail liées à des moyens de pression. Les situations extrêmes peuvent ainsi être expliquées.

Lorsqu'on s'attarde à la comparaison entre les provinces, on note que la production de services de santé est l'étape souffrant le plus d'inefficience et que le Manitoba est la province qui s'en tire le mieux à tous les niveaux. Lorsque l'on quantifie l'inefficience en termes monétaires, des divergences existent entre les provinces. Celles ayant un fort bassin de population sont celles qui souffrent le plus des sommes perdues en santé. Les provinces du centre se retrouvent donc en haut du classement, suivi des Prairies et de l'Ouest canadien avant d'être clos par les provinces maritimes. Cependant, peu importe la position géographique des entités administratives, l'évolution des sommes perdues dans le domaine de la santé au pays est sans équivoque. Les pertes en santé ont déjà commencé à croître et cette croissance s'accroîtra.

Les résultats présentés tout au long de ce mémoire permettent de mieux caractériser les situations des différentes provinces canadiennes en matière de santé. On note ainsi une trop grande utilisation des ressources compte tenu du nombre de services et de niveaux de santé produits. Une décision devra alors être prise, soit en produisant davantage de services ou bien en réduisant le nombre de ressources étant donné l'entendue des économies pouvant être réalisées. En raison de l'importance du domaine de la santé, ces questions se doivent d'être posées.

Puisque le vieillissement de la population s'accroîtra au cours des prochaines années, les décisions qui seront prises par les gouvernements auront d'importantes retombées. Dans la mesure où le Canada n'est pas le seul aux prises avec ce problème, il serait intéressant d'appliquer les mêmes procédés utilisés ici avec les banques de données des autres pays concernés. La comparaison des résultats pourrait nous renseigner sur de possibles améliorations à appliquer chez nous, mais il faut toutefois garder en tête l'aspect humain et délicat qu'ont les impacts des décisions en santé.

RÉFÉRENCES

- Aigner, D., Lovell, C. A. K. et Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1), 21–37.
- Banker, R. D., Charnes, A. et Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management science*, 30(9), 1078–1092.
- Belotti, F., Daidone, S., Ilardi, G. et Atella, V. (2013). Stochastic Frontier Analysis using Stata. *Stata Journal*, 13(4), 719–758.
- Bilodeau, D., Crémieux, P.-Y., Jaumard, B., Ouellette, P. et Vovor, T. (2004). Measuring hospital performance in the presence of quasi-fixed inputs : An analysis of Québec hospitals. *Journal of Productivity Analysis*, 21(2), 183–199.
- Bilodeau, D., Crémieux, P.-Y. et Ouellette, P. (2000). Hospital cost function in a non-market health care system. *Review of Economics and Statistics*, 82(3), 489–498.
- Bilodeau, D., Crémieux, P.-Y. et Ouellette, P. (2002). Hospital technology in a nonmarket health care system. *Southern Economic Journal*, 68(3), 511–529.
- Bilodeau, D., Crémieux, P.-Y. et Ouellette, P. (2009). Hospital performance in a noncompetitive environment. *Applied Economics*, 41(4), 459–468.
- Caves, D. W., Christensen, L. R. et Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6), 1393–1414.
- Chalkley, M. et Malcomson, J. M. (2000). Government purchasing of health services. In A. J. Culyer et J. P. Newhouse (dir.), *Handbook of Health Economics*, volume 1A, 847–890. Elsevier : New York.
- Charnes, A., Cooper, W. W. et Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of Operational Research*, 2(6), 429–444.
- Colombie-Britannique. Ministry of Finance. (2017). *Budget and fiscal plan*. Colombie-Britannique. [Document PDF]. Rapport déposé le 21 février 2017 à l'Assemblée législative de la Colombie-Britannique. Victoria : Gouvernement de la Colombie-Britannique. Récupéré le 5 mai 2017 de http://bcbudget.gov.bc.ca/2017/bfp/2017_Budget_and_Fiscal_Plan.pdf.

- Cowing, T. G. et Holtmann, A. G. (1983). Multiproduct short-run hospital cost functions : Empirical evidence and policy implications from cross-section data. *Southern Economic Journal*, 49(3), 637–653.
- Cowing, T. G. et Stevenson, R. E. (1982). Automatic adjustment clauses and allocative efficiency in public utilities. *Journal of Economics and Business*, 34(4), 317–329.
- Crémieux, P.-Y., Ouellette, P. et Pilon, C. (1999). Health care spending as determinants of health outcomes. *Health Economics*, 8(7), 627–639.
- Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, 19(3), 273–292.
- Deprins, D., Simar, L. et Tulkens, H. (1984). Measuring labour efficiency in post offices. In M. Marchand, P. Pestieau, et H. Tulkens (dir.), *The performance of public enterprises : Concepts and measurement*, 243–268. North-Holland : Amsterdam.
- Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of Econometrics*, 4(2), 115–145.
- Evans, R. G. (1971). Behavioural cost functions for hospitals. *The Canadian Journal of Economics*, 4(2), 198–215.
- Färe, R., Grosskopf, S., Lindgren, B. et Roos, P. (1992). Productivity changes in swedish pharmacies 1980–1989 : A non-parametric malmquist approach. *Journal of Productivity Analysis*, 3(1), 85–101.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253–290.
- Gerdtham, U.-G. et Jönsson, B. (2000). International comparisons of health expenditure : Theory, data and econometric analysis. In A. J. Culyer et J. P. Newhouse (dir.), *Handbook of Health Economics*, volume 1A, 11–53. Elsevier : New-York.
- Greene, W. (2004). Distinguishing between heterogeneity and inefficiency : Stochastic Frontier Analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, 13(10), 959–980.
- Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1), 7–32.
- Hitiris, T. et Posnett, J. (1992). The determinants and effects of health expenditure in developed countries. *Journal of Health Economics*, 11(2), 173–181.

- ICIS. (2017). *Disponibilité des données des banques de données de l'ICIS selon la province, le territoire et l'année*. [Base de données]. Récupéré de https://www.cihi.ca/sites/default/files/document/data_holdings_availability_fr.xlsx.
- Koopmans, T. C. (1951). Efficient allocation of resources. *Econometrica*, 19(4), 455–465.
- Ma, C. A. (1994). Health care payment systems : Cost and quality incentives. *Journal of Economics & Management Strategy*, 3(1), 93–112.
- Meeusen, W. et van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2), 435–444.
- Mougeot, M. et Naegelen, F. (2005). Hospital price regulation and expenditure cap policy. *Journal of Health Economics*, 24(1), 55–72.
- Newhouse, J. P. (1977). Medical-care expenditure : A cross-national survey. *The Journal of Human Resources*, 12(1), 115–125.
- Ouellette, P. et Petit, P. (2010). Efficience budgétaire des institutions de santé : Une bibliographie commentée. *Centre sur la productivité et le bien-être*, HEC Montréal.
- Ouellette, P. et Petit, P. (2018). Efficiency measurement : A methodological review and synthesis. *Data Envelopment Analysis Journal*, 4(1), 67–107.
- Québec. Ministère des Finances. (2017). *Le plan économique du Québec*. Québec. Québec. [Document PDF]. Rapport déposé le 28 mars 2017 à l'Assemblée nationale du Québec. Québec : Gouvernement du Québec. Récupéré le 5 mai 2017 de http://www.budget.finances.gouv.qc.ca/budget/2017-2018/fr/documents/PlanEconomique_Mars2017.pdf.
- Rapport Ménard. (2005). *Pour sortir de l'impasse : La solidarité entre nos générations*. Québec. 155 p.
- Roberts, J. (2000). Spurious regression problems in the determinants of health care expenditure : A comment on Hitiris. *Applied Economics Letters*, 7(5), 279–283.
- Saskatchewan. Ministry of Finance. (2017). *Saskatchewan provincial budget 17-18*. Saskatchewan. [Document PDF]. Rapport déposé le 22 mars 2017 à l'Assemblée législative de la Saskatchewan. Regina : Gouvernement de la Saskatchewan. Récupéré le 5 mai 2017 de <http://www.finance.gov.sk.ca/budget17-18/2017-18Budget.pdf>.
- Simar, L. et Wilson, P. W. (1998). Sensitivity analysis of efficiency scores : How to bootstrap in nonparametric frontier models. *Management science*, 44(1), 49–61.