

Impact des investissements dans la recherche et développement dans le
domaine de la santé sur la santé de la population

Par : Jean-François Leroux
(4083057)

Mémoire présenté au département d'économie de l'Université d'Ottawa
dans le cadre du programme de Maîtrise ès Arts en économie
Superviseur : M. David Gray

ECO 7997

Ottawa, Ontario
Août 2010

Table des matières

Sommaire exécutif	2
Introduction.....	3
Révision de la littérature	6
Caractérisation du modèle et des variables	20
Premier aperçu des données et de l'échantillon.....	23
Vérification des hypothèses statistiques	24
Description des résultats du modèle statique	31
Le modèle dynamique.....	33
Limitations et recommandations.....	35
Bibliographie.....	38
Annexe	39

Sommaire exécutif

De nos jours, les dépenses dans la recherche en santé suscitent de plus en plus d'intérêt au niveau gouvernemental ainsi qu'au niveau de la population en général. Depuis plusieurs années, des chercheurs se sont intéressés à développer des méthodes afin de pouvoir évaluer l'impact de ces investissements. En 2009, l'Académie canadienne des sciences de la santé a développé un cadre conceptuel ayant pour objectif d'évaluer cet impact. Lors de l'étude, celle-ci en est venue à la conclusion que «l'objectif ultime de tous les projets de recherche en santé consistait à l'amélioration de la santé des gens et, par le fait même, de leur longévité et de leur qualité de vie.¹» Toutefois, personne n'a à ce jour utilisé ce cadre afin de développer un modèle visant à évaluer l'impact des investissements dans la recherche en santé, sur la santé de la population. Or, ce rapport explore cet aspect en tentant d'identifier l'impact de ces investissements sur l'espérance de vie de la population. D'après les résultats, les dépenses en recherche et développement dans ce domaine augmentent significativement l'espérance de vie de la population. Un résultat similaire est obtenu pour celui des hommes. Toutefois, ce résultat ne tient pas compte du fait que cet impact peut se faire ressentir avec un certain retard. En effet, représenter ce type d'aspect lors d'une analyse comme celle décrite à l'intérieur de ce rapport, représente un très grand défi. Néanmoins, ce rapport offre quelques procédures à suivre afin de pouvoir faire une telle analyse. De plus, il identifie les limitations du rapport ainsi qu'une description des études futures qui devront être accomplies afin de pouvoir évaluer ce type d'impact.

¹ Académie canadienne des sciences de la santé (2009), «Créer un impact », http://www.caahs-acss.ca/e/pdfs/ROI_FrenchReport.pdf, p.46, consulté le 10 juillet 2010.

Introduction

De nos jours, la santé des gens est l'un des aspects qui attire le plus d'attention des différents paliers du Gouvernement au Canada, notamment au niveau des provinces. En effet, selon le rapport sur les tendances des dépenses nationales de santé publié par l'Institut canadien d'informations sur la santé, la part des dépenses gouvernementales reliées à la santé est passée de 28,2% en 1993 à 35,9% en 2008, ce qui constitue une augmentation d'environ 27% en 15 ans². De plus, le vieillissement de la population au Canada indique vraisemblablement que cette statistique risque d'augmenter durant la prochaine décennie. La raison est fort simple : des dépenses publiques plus soutenues dans le domaine de santé devraient accroître la santé de la population et ainsi améliorer son bien-être. Malheureusement, cette relation de cause à effet n'est pas toujours aussi simple contrairement à ce que les gens peuvent croire. En effet, plusieurs études ont tenté de démontrer que les dépenses dans le domaine de santé amenaient nécessairement des résultats positifs pour la santé sans toutefois être concluantes. Entre autres, John Nixon et Philippe Ulmann ont démontré qu'il y avait bel et bien une relation positive entre les deux mais qu'il était difficile de déterminer si l'augmentation des dépenses de santé avait un effet significatif direct sur la santé de la population³.

De ce fait, un aspect beaucoup moins général concernant les dépenses totales liées à la santé commence à susciter de plus en plus d'intérêt au Canada et ailleurs dans le

² Institut canadien d'Information sur la santé, «Tendances des dépenses nationales de la santé», http://secure.cihi.ca/cihiweb/products/National_health_expenditure_trends_1975_to_2009_fr.pdf, p.53, consulté le 18 mai 2010.

³ Nixon, J, Ulmann P. (1996) «The Relationship between Health Care Expenditures and Health Outcomes», *The European Journal of Health Economics*, vol.7, 1 n° 1, pp.7-18.

monde, soit les investissements dans la recherche et le développement en santé. Au Canada, plusieurs fonds publics servent notamment à financer ce genre de recherche. Au niveau fédéral, la principale source de financement provient des Instituts de Recherche en Santé du Canada, qui finançaient pour près d'un milliard de dollars dans la recherche en santé en 2009. Au niveau provincial, l'Alliance canadienne des organismes provinciaux de recherche en santé (NAPHRO) regroupe 11 institutions parmi les 10 provinces canadiennes. Celles-ci ont pratiquement le même mandat que les IRSC, c'est-à-dire de financer la recherche et le développement dans le domaine de santé.

En effet, au Canada en 2005, les Instituts de Recherche en Santé du Canada ont réuni une vingtaine d'experts dans la matière afin de créer un cadre conceptuel pour mesurer les différents impacts financés par les IRSC⁴. De ce groupe d'experts, on retrouvait notamment le professeur de l'Université de Brunel, Martin Buxton, fondateur du Health Economics Research Group (HERG), qui s'intéresse depuis plusieurs années à ce type d'analyse. Or, cette réunion a permis aux différents experts de déterminer l'information nécessaire afin d'estimer cet impact et également d'identifier certains défis afin de pouvoir faire de telles estimations. En 2009, l'Académie canadienne des sciences de la santé a soumis un rapport aux différentes institutions canadiennes finançant ce type de recherche afin d'élaborer un modèle pour que l'on puisse faire ce type d'estimation. Dans ce rapport, l'ACSS a également suggéré quelques indicateurs qui pourraient servir à évaluer l'impact de la recherche en santé. Malheureusement, aucune estimation de cet impact n'a encore été faite. Ce travail tentera donc de faire une première estimation de

⁴ Institut de Recherche en Santé du Canada, «Élaboration d'un cadre des IRSC pour mesurer l'impact de la recherche en santé», http://www.cihr-irsc.gc.ca/f/documents/meeting_synthesis_f.pdf, consulté le 21 mai 2010.

l'impact de la recherche et du développement en santé sur la santé de la population en utilisant les méthodes économétriques.

Elle débutera par une révision de ce qui a été accompli à ce sujet depuis les 15 dernières années. Cette révision de la littérature portera principalement sur les travaux de Martin Buxton et de Steve Hanney en 1996 et sur ce que les institutions canadiennes ont apporté à cette recherche depuis ce temps. Par la suite, celle-ci caractérisera les différentes variables à utiliser afin de modéliser les dépenses en recherche et développement en santé, sur la santé des gens. Elle identifiera également à ce moment, les sources de données utilisées pour ce projet. L'étude offrira ensuite une analyse descriptive de l'échantillon et des variables afin de pouvoir identifier certains problèmes qui pourraient survenir lors de l'analyse. Une fois cette première analyse terminée, elle pourra s'attarder aux analyses plus complexes en vérifiant si les nombreuses hypothèses économétriques sont satisfaites afin de pouvoir rapporter des résultats non biaisés. Nous passerons par la suite à une analyse du modèle statique, c'est-à-dire du modèle où les effets de retard ne sont pas considérés. Comme deuxième analyse, nous tenterons d'ajouter ces effets de retard au modèle afin d'identifier à quel moment les dépenses en recherche et développement en santé auront un impact sur la santé des gens. Après avoir rapporté ces résultats, l'étude se penchera sur les différentes limitations des modèles. Elle conclura en offrant certaines recommandations afin de pouvoir mieux évaluer ce type d'impact. Elle offrira également à ce moment un résumé de ce qui doit être accompli dans le futur pour bien répondre à cette question.

Révision de la littérature

Les investissements dans la recherche et le développement dans le domaine de la santé prennent de plus en plus de place pour de nombreux pays. Au Canada par exemple, ces investissements sont passés de 2 954 millions de dollars en 1998 à environ 6 387 millions de dollars en 2009, pour une augmentation de 116% en 12 ans⁵. En annexe 1, nous pouvons remarquer que depuis 2001, l'accroissement de ces dépenses a été beaucoup plus important que la hausse des dépenses dans la recherche et développement en général. Cela fait donc en sorte que la proportion des dépenses totales dans la recherche et le développement, liée au domaine de la santé, a significativement augmenté depuis les dernières années au Canada. Or, beaucoup de ces investissements au Canada sont financés à partir des fonds publics, que ce soit au niveau fédéral ou provincial. Toutefois, lorsque l'argent des contribuables sert à financer des projets de telle envergure, il est normal de vouloir déterminer si oui ou non de tels investissements sont bénéfiques.

C'est pourquoi depuis un certain temps, plusieurs chercheurs et organismes s'intéressent à ce genre de question. Prenons l'exemple des Instituts de Recherche en Santé du Canada. Ils ont créé une unité qui tente d'évaluer l'impact de la recherche en santé, financée par les IRSC. Ailleurs dans le monde, plus précisément en Grande-Bretagne, l'un des membres fondateurs du *Health Economics Research Group* à l'Université de Brunel, Martin Buxton, s'intéresse depuis plusieurs années à ce type d'évaluation, et son travail a été utilisé par plusieurs autres chercheurs et institutions partout dans le monde. C'est d'ailleurs avec l'une de ses publications les plus célèbres

⁵ Statistiques Canada, «*Science Statistics*», <http://www.statcan.gc.ca/pub/88-001-x/88-001-x2010003-eng.pdf>, consulté le 8 juillet 2010.

que nous débiterons la révision de la littérature afin de connaître ce qui a été accompli dans le domaine.

En 1996, lui et son acolyte, le Dr. Stephen Hanney, ont développé un premier cadre conceptuel visant à évaluer l'impact total de la recherche en santé⁶. À ce moment, les auteurs avaient déterminé qu'il devenait essentiel de justifier les dépenses faites dans la recherche pour les services de santé en démontrant les bénéfices qui résultent de cette recherche, puisque celles-ci avaient largement augmenté depuis un certain temps. Or, cette étude a permis de diviser les rendements de ces investissements en 5 grandes catégories, soit les connaissances, les avantages de la recherche, les avantages politiques et administratifs, les avantages dans le secteur de la santé et finalement les avantages économiques de la recherche. Durant les années suivant cette publication, la caractérisation des différentes catégories a grandement évolué, et plusieurs institutions, dont les IRSC et celles des provinces canadiennes membres du NAPPHRO, ont utilisé les résultats de cette recherche afin de développer leur propre cadre conceptuel et évaluer les retours liés aux investissements dans la recherche en santé. C'est pourquoi comme deuxième analyse littéraire, nous allons nous intéresser au processus que les institutions canadiennes, principalement les Instituts de Recherche en Santé du Canada, ont utilisé afin de développer leur propre cadre conceptuel évaluant l'impact de la recherche en santé.

Le processus débute en 2004, à la suite du premier plan stratégique développé par la nouvelle organisation fondée en 2000, où les IRSC s'engagent à « élaborer et mettre en

⁶ Buxton MJ, Hanney S (1996) «How can payback from health services research be assessed?», *Journal of Health Services Research & Policy*, vol 1, no 1, pp 35-43.

œuvre un cadre qui permet l'évaluation du rendement de l'organisation et la valeur de ses programmes de soutien à la recherche.»⁷ À ce moment, l'institution organise une réunion de deux jours, qui a eu lieu au début de l'année 2005, avec des experts dans le domaine, dont le professeur Martin Buxton, afin d'obtenir des conseils vis-à-vis leurs préoccupations.

Lors de cette réunion, les experts se sont mis d'accord sur le fait que le terme impact de la recherche en santé, au lieu de rendements sur l'investissement, était plus approprié afin de répondre aux besoins des différentes organisations finançant cette recherche. Également, les membres des IRSC ainsi que les experts en la matière ont identifié certains défis quant à l'utilisation de ce cadre, notamment l'attribution de l'amélioration de la santé de la population à ces investissements. Bien qu'une ébauche ait été bien dessinée, il restait beaucoup de travail à accomplir. La description de chacune des catégories devait être élaborée et les indicateurs servant à mesurer l'impact devaient être identifiés. Les méthodes utilisées afin de mesurer cet impact restaient toujours à être déterminées. C'est pourquoi les IRSC, en collaboration avec les institutions membres du NAPPHRO ainsi que d'autres institutions canadiennes finançant la recherche en santé au Canada, ont donné à l'Académie canadienne des sciences de la santé (ACSS) le mandat de répondre à ce type de question sous forme de rapport. Ce n'est toutefois qu'en janvier 2009 que le rapport de l'ACSS fut rédigé. Toutefois, son mandat ne constituait pas seulement à définir le cadre conceptuel des IRSC. Il se voulait également une référence pour comprendre ce qui avait été accompli dans le passé, en ce qui a trait à l'évaluation

⁷ Instituts de Recherche en Santé du Canada (2005), «Élaboration d'un cadre des IRSC pour mesurer l'impact de la Recherche en Santé», http://www.cihr-irsc.gc.ca/f/documents/meeting_synthesis_f.pdf, p.1, consulté le 8 juillet 2010.

de l'impact de la recherche en santé, et à ce qui devra être fait dans le futur avant de pouvoir mesurer et quantifier cet impact adéquatement.

Bien que le mandat principal de l'ACSS fût d'élaborer un modèle logique afin de bien expliquer l'impact global de la recherche en santé, plusieurs aspects devaient être pris en considération avant le développement de celui-ci. Par exemple, afin de bien cerner l'impact ou le rendement sur l'investissement, l'ACSS a jugé nécessaire de définir ces termes. D'après celle-ci, trois composantes majeures doivent être considérées afin de bien cerner ces rendements, soit les bienfaits pour la santé, les avantages sociaux et les avantages économiques. Les bienfaits pour la santé sont définis comme étant : «des changements apportés à la santé des personnes (physique ou psychologique) aux services de santé ou aux déterminants de la santé plus généraux.»⁸ Cette définition sera particulièrement importante lorsque nous ferons notre propre estimation des bienfaits liés aux investissements dans la recherche et le développement dans le domaine de la santé.

Cette description a permis par la suite à l'ACSS de déterminer les rendements attendus de la recherche en santé au Canada afin d'élaborer les différentes étapes du modèle logique décrivant l'impact global de cette recherche. D'après l'Académie, «l'objectif ultime de tous les projets de recherche en santé consistait à l'amélioration de la santé des gens et, par le fait même de leur longévité et de leur qualité de vie.»⁹ Cette considération devait également se refléter à l'intérieur du modèle logique. Puisque cette

⁸ Académie canadienne des sciences de la santé (2009), «Créer un impact », http://www.cahs-acss.ca/e/pdfs/ROI_FrenchReport.pdf, p.59, consulté le 10 juillet 2010.

⁹ Académie canadienne des sciences de la santé (2009), «Créer un impact », http://www.cahs-acss.ca/e/pdfs/ROI_FrenchReport.pdf, p.46, consulté le 10 juillet 2010.

amélioration était l'objectif ultime, l'ACSS devait par la suite déterminer les étapes nécessaires à franchir afin de pouvoir arriver à cet objectif.

Pour ce faire, l'ACSS s'est basée sur trois ouvrages qui ont tenté d'identifier les déterminants de la santé. D'après sa recherche, on est arrivé à la conclusion que ces différents déterminants peuvent être classés selon quatre différents niveaux :

1. L'âge et l'hérédité
2. Les soins de santé et le comportement personnel
3. Les conditions de vie et de travail
4. Le contexte économique, social, culturel et environnemental

Suite à la détermination de ces différents niveaux, l'ACSS s'est penchée sur la question du modèle logique. Les experts se sont inspirés du modèle initialement développé par Buxton et Hanney en 1996, afin de créer leur propre modèle qui se divise en six différentes étapes.

La première étape consiste en la détermination des besoins en matière de santé et de bien-être au sein de la population. Pour ce faire, l'institution doit être en mesure de déterminer les coûts de la recherche dans chacun des domaines et les différents avantages associés à cette recherche. En comparant ces deux aspects, l'institution peut donc choisir de déboursier plus d'argent dans un domaine particulier au détriment des autres afin que les investissements aient plus d'impact.

La deuxième étape quant à elle correspond à l'étape même de la recherche. C'est à cette étape que les chercheurs vont dans les laboratoires afin de faire leur recherche dans différents domaines, selon les besoins déterminés à l'étape 1. Ce n'est toutefois qu'à la troisième étape que les premiers résultats concrets de la recherche et du développement

en santé apparaissent. À cette étape, la recherche produit des résultats et des connaissances qui peuvent se traduire en publications, formations, collaborations, etc. Cette production de connaissances est essentielle afin d'augmenter le plus possible l'impact de la recherche en santé. En effet, cette production permet un échange des connaissances entre les chercheurs afin de pouvoir contribuer à l'avancement des résultats. Par exemple, les organisations canadiennes peuvent utiliser des résultats obtenus à partir de recherches faites à l'étranger afin d'améliorer la santé des Canadiens et vice-versa. Ces échanges de connaissances peuvent donc améliorer la santé et le bien-être des gens partout dans le monde. La quatrième étape du modèle peut également contribuer à ce type d'échange. À ce moment, les preneurs de décisions et les différents groupes qui peuvent potentiellement diffuser cette recherche, comme les médias, utilisent ces résultats afin de faire le lien entre les chercheurs et la population en général.

À la cinquième étape, on retrouve tous les changements qui auront potentiellement des effets sur la santé et le bien-être des gens. Le modèle divise ces changements en deux niveaux, soit le niveau des soins de santé et le niveau des déterminants de la santé. Au niveau des soins de santé, on considère notamment les changements apportés au système de santé, par exemple, les effets sur les temps d'attentes, les méthodes utilisées et la sécurité des patients. Or, cette production ou cette amélioration de nouvelles méthodes tend à améliorer la santé et le niveau de vie des habitants d'un pays. Pour ce qui est du niveau des déterminants de la santé, l'ACSS fait plutôt référence aux changements personnels, en particulier les modes de vie, ainsi qu'aux changements au niveau des sociétés, par exemple, les déterminants

environnementaux. Finalement, tous ces changements contribuent à des améliorations potentielles au niveau de la santé et du bien-être de la population ainsi que pour la prospérité économique et sociale, ce qui constitue la dernière étape du modèle logique. Toutefois, ce cycle est rétroactif puisqu'à la suite de toutes ces étapes, les besoins en termes de santé et de bien-être changent ou évoluent. En annexe 2, nous retrouvons le schéma développé par l'ACSS afin de représenter les différentes étapes du modèle.

Comme prochaine étape, nous pouvons facilement faire le lien entre ce modèle logique et les cinq catégories d'impacts développées par Buxton et Hanney en 1996. En effet, nous pouvons remarquer que l'impact concernant la production de connaissances et le renforcement des capacités se produisent à l'étape 3 du modèle tandis que l'impact concernant la prise de décision éclairée peut être facilement lié à l'étape 4. Ces deux catégories se traduisent donc par l'impact des échanges qui existent entre les différents groupes afin d'améliorer la santé et le bien-être des gens. Pour ce qui est de l'impact qui correspond aux avantages socio-économiques, il peut être associé à la composante de la dernière étape du modèle concernant la prospérité économique et sociale. En ce qui concerne la catégorie qui nous intéresse plus particulièrement, les bienfaits pour la santé, celle-ci se différencie des autres puisqu'elle peut être attribuée à plus d'une étape du modèle logique. En effet, les experts divisent cette catégorie en trois sous-catégories, soit l'état de santé, les déterminants de la santé et les systèmes de santé. Comme mentionné auparavant, l'état de santé se caractérise à la dernière étape tandis que les deux autres sous-catégories s'associent beaucoup plus aux facteurs qui sont influencés par la recherche et qui auront par la suite un impact sur la santé des gens.

En termes d'indicateurs, l'ACSS en identifie 19 différents qui peuvent être associés aux bienfaits pour la santé, dont 5 qui représentent l'état de santé. Afin de représenter la mortalité, les experts ont opté pour l'indice estimant les années de vies potentiellement perdues. Curieusement, l'espérance de vie de la population ne fait pas partie de cette liste d'indicateurs de la santé de la population. Pour ce qui est des indicateurs développés afin de représenter les comportements des gens, l'ACSS suggère d'utiliser des variables telles que l'obésité, la consommation d'alcool et le niveau de pollution. En ce qui concerne la qualité des services et des systèmes de santé, l'ACSS suggère des indicateurs comme les temps d'attente, la satisfaction des patients et le taux de réadmission. Cette liste d'indicateurs devra également être considérée lors des choix des différentes variables à utiliser lors de la création de notre modèle.

Une fois le modèle spécifié et les différents indicateurs identifiés, l'ACSS a également reçu le mandat d'élaborer des stratégies pour une future utilisation du cadre conceptuel en identifiant certains problèmes qui pourraient survenir lors de son application. L'ACSS a identifié trois problèmes potentiels majeurs lors de l'application de ce cadre conceptuel. Le premier concerne le fait que l'utilisateur du modèle ne prenne en considération que les impacts positifs de la recherche en santé. En effet, une augmentation des investissements dans la recherche et le développement en santé au détriment de ceux qui sont faits dans un autre domaine peut parfois engendrer des effets négatifs au niveau de la population à court terme.

Le deuxième problème potentiel que l'ACSS a identifié concerne l'attribution de l'impact dû aux investissements versus les autres facteurs. Par exemple, plusieurs

facteurs, comme les comportements personnels et le contexte économique, peuvent avoir des effets directs sur la santé de la population. Ces facteurs doivent donc être pris en considération lors d'une évaluation de l'impact. L'une des méthodes suggérées par l'ACSS afin de contrevenir à ce type de problème consiste à utiliser les méthodes de régression multiple. Finalement, le troisième problème potentiel à considérer lors d'une telle évaluation concerne le délai entre le moment où les subventions sont attribuées et le moment où les résultats des recherches surviennent. De plus, il est tout à fait normal de penser que ce délai varie selon le domaine ou l'initiative de recherche en question. Par exemple, les résultats d'un nouveau traitement contre le cancer seront sans doute visibles dans une période d'années assez importante contrairement à ceux attribuables au vaccin contre la grippe AH1N1 où les résultats furent presque instantanés. Les délais varieront selon le domaine dans lequel les investissements dans la recherche sont faits. Tous ces problèmes doivent être considérés lors d'une évaluation de l'impact de la recherche en santé, peu importe le domaine de recherche et le type d'impact que nous voulons identifier.

Comme nous avons pu le constater, plusieurs études ont été menées afin d'élaborer une façon d'évaluer l'impact de la recherche en santé au Canada et ailleurs dans le monde. Le modèle logique de l'Académie canadienne des sciences de la santé nous permet de bien comprendre la façon dont la recherche en santé se diffuse à travers plusieurs étapes afin d'affecter la population. De plus, la caractérisation de l'impact divisé en cinq catégories distinctes, nous permet de bien identifier le niveau auquel nous voulons l'évaluer. Les indicateurs suggérés à l'intérieur de chacune d'elles nous

permettent également de bien sélectionner les différentes variables dépendantes et indépendantes nécessaires à notre évaluation. Finalement, l'ACSS a cerné quelques problèmes qui doivent être pris en considération lorsque nous voulons faire une telle évaluation.

Malheureusement, aucune étude à ce jour n'a utilisé ce cadre conceptuel afin d'estimer l'impact ou les rendements liés aux investissements en recherche et développement dans le domaine de la santé. Aucun modèle n'a été développé afin de déterminer si les investissements en santé ont des effets directs sur la santé de la population. De plus, ce cadre conceptuel a été conçu afin de pouvoir répondre à des objectifs bien précis, c'est-à-dire qu'il est préférable de l'appliquer à des recherches spécifiques. Il sera donc plus difficile de l'appliquer lorsque nous voudrions déterminer les effets de la recherche et du développement en santé en général, comme c'est le cas pour notre étude. Bref, ce cadre conceptuel consiste en un bon point de départ pour notre recherche, mais une analyse supplémentaire est nécessaire afin de pouvoir créer notre propre modèle économétrique. Afin de remédier à cette situation, nous avons identifié deux recherches qui tentaient de déterminer si les dépenses globales dans le domaine de la santé avaient des effets directs sur la santé de la population. Par la suite, nous avons lié ces recherches à notre propre problème.

La première étude que nous avons vérifiée fut celle réalisée par Nixon et Ulmann en 2006. Elle tentait de prouver qu'il existait des liens de cause à effet entre les dépenses en santé et les bienfaits pour la santé de la population pour les 15 pays de l'Union Européenne durant la période 1980-1995. Afin de tester cette hypothèse, les auteurs ont

identifié des variables pouvant avoir des effets sur la santé de la population. Ils ont par la suite construit un modèle de régression linéaire multiple. En termes d'indicateurs de la santé de la population, Nixon et Ulmann ont utilisé l'espérance de vie et le taux de mortalité infantile. Ils ont également utilisé l'espérance de vie des femmes et des hommes séparément afin de déterminer si les effets étaient différents dans les deux cas. Pour ce qui est des différentes composantes pouvant avoir des effets potentiels sur ces indicateurs, les auteurs ont cru nécessaire d'inclure 12 variables indépendantes au modèle afin de bien représenter la réalité. En plus de la variable indépendante principale, soit les dépenses dans le domaine de la santé, les auteurs ont inclus les éléments suivants lors de la construction de leur modèle : le nombre de médecins par individu, le nombre de lits d'hôpitaux par habitant, le taux d'admission des patients, le temps moyen d'hospitalisation, la couverture d'assurance médicale, le taux de chômage, la consommation d'alcool, les dépenses pour le tabac, la consommation de fruits, la consommation de protéines par année par individu ainsi que les émissions d'oxyde de soufre¹⁰. Toutes les données pour cette analyse provenaient de la base de données sur la santé offerte par l'Organisation de coopération et de développement économique (OCDE).

En terme de résultats, les chercheurs ont déterminé que les dépenses dans le domaine de la santé augmentaient significativement l'espérance de vie des hommes et des femmes et que les femmes étaient celles qui bénéficiaient le plus de ces dépenses. Un résultat semblable se reflétait sur le taux de mortalité infantile qui diminuait avec une

¹⁰ Nixon, John et Ulmann Philippe. (2006) «The Relationship between health care expenditure and health outcomes», *European Journal of Health Economics*, vol 7 n°1, p.14.

augmentation des dépenses dans le domaine de la santé. De plus, la variable représentant le nombre de médecins par habitant était celle ayant le plus d'impact sur la santé de la population.

Malgré le fait que les résultats étaient intéressants, des limitations importantes ont été considérées par les chercheurs. La première mentionnée concernait le fait que l'espérance de vie et le taux de mortalité infantile ne sont que des approximations qui ne représentent que partiellement la santé de la population. Toutefois, ces indicateurs sont à ce jour les meilleurs représentant la santé de la population selon plusieurs experts. C'est pourquoi ce problème est sans doute le moins important parmi ceux énumérés par Nixon et Ulmann.

Le deuxième problème identifié concerne la spécification du modèle. En effet, les auteurs apportent le point qu'il pourrait y avoir un biais dans les estimateurs obtenus dû à l'omission de certaines variables indépendantes importantes, qui n'étaient pas disponibles à partir de la base de données utilisée. Par exemple, l'éducation et la consommation de tabac sont des indicateurs importants, selon les auteurs, qui n'ont pas été considérés. Finalement, le dernier problème que les auteurs ont cru bon de reconnaître concerne les effets de retard qui n'ont pas été examinés lors de leur étude. En effet, lors de leur estimation, les auteurs ont pris en considération que les différentes variables utilisées avaient des effets instantanés sur la santé de la population. Toutefois, il est tout à fait normal de penser que les effets de la majorité des indicateurs se feront ressentir sur la santé de la population quelques ou même plusieurs années suivant les actions apportées. Or, ces deux derniers points apportés sont très importants lors d'une analyse

économétrique de ce type. Afin de valider les résultats obtenus à partir d'une étude économétrique, plusieurs tests doivent être faits et certaines modifications au modèle peuvent être nécessaires. Par exemple, pour que les tests statistiques concernant la significativité des résultats soient valables, il faut vérifier certaines hypothèses économétriques comme l'hétéroscédasticité, l'autocorrélation et la normalité des erreurs. Si ces hypothèses ne sont pas respectées, il faudra utiliser des méthodes d'estimations comme la méthode des moindres carrés généralisés, afin que ces tests soient valides. De plus, pour s'assurer que les résultats soient non-biaisés, il faudra effectuer des tests, comme le test de Ramsey, afin de vérifier qu'aucune variable n'est omise dans le modèle.

La deuxième étude que nous avons étudiée est celle de Crémieux, Ouellette et Pilon. Ceux-ci voulaient tester la même hypothèse que Nixon et Ulmann mais pour les dix provinces canadiennes. Comme Nixon et Ulmann, les auteurs ont utilisé l'espérance de vie et le taux de mortalité infantile afin de représenter la santé de la population. La variable indépendante principale qu'ils ont utilisée était les dépenses en santé par habitant. Ils ont toutefois cru bon d'inclure 9 autres variables indépendantes afin de représenter les autres aspects qui peuvent avoir des effets sur la santé de la population : le nombre de médecins par habitant, le PIB par habitant, le nombre de gradués par habitant, le taux de pauvreté, la densité de population, le taux de fumeurs, la consommation d'alcool par habitant, les dépenses sur la viande et les dépenses sur le gras¹¹. Afin de tester leur hypothèse, ils ont utilisé les données offertes par Statistiques Canada, Santé Canada et Santé Québec pour la période 1978-1992.

¹¹ Crémieux, Ouellette et Pilon. (1999) «Health Care Spending as Determinants of Health Outcomes», *Journal of Health Economics*, vol 8 n° 1, p.630.

Tout comme Nixon et Ulmann, les auteurs ont déterminé qu'une augmentation des dépenses en santé par habitant améliorait significativement la santé de la population. Un résultat similaire fut trouvé en ce qui concerne le nombre de médecins par habitant. Pour ce qui est des dépenses sur le gras et du taux de fumeurs, Crémieux, Ouellette et Pilon ont déterminé qu'une augmentation de ces variables détériore la santé de la population.

Contrairement à Nixon et Ulmann, les auteurs ont considéré les problèmes d'hétéroscédasticité et d'autocorrélation qui se produisent fréquemment lors de l'utilisation des données en panel. En effet, ceux-ci ont déterminé que ces deux problèmes étaient présents, et qu'il valait mieux utiliser la méthode des moindres carrés généralisés au lieu de celle des moindres carrés ordinaires, afin d'estimer les différents paramètres. De plus, les auteurs ont utilisé un modèle à effet fixe en utilisant des variables dichotomiques pour les provinces canadiennes, afin de considérer les particularités intra-temporelles spécifiques à chacune d'elles. Lors de notre étude, il faudra vérifier s'il est nécessaire de considérer ces modèles afin d'obtenir de meilleurs résultats.

Toutefois, le problème majeur de cette étude est sans doute le fait que les auteurs n'ont pas considéré les effets de retard. Encore une fois, tout comme Nixon et Ulmann, les auteurs ont testé leur hypothèse en supposant que les différentes variables avaient des effets instantanés sur la santé de la population, ce qui n'est pas nécessairement vrai. De plus, les auteurs mentionnent qu'il serait intéressant de faire ce type d'analyse en considérant des dépenses faites dans des domaines spécifiques de la santé. «Par exemple,

est-ce que l'accroissement des dépenses dans les soins de longue durée améliorent la longévité de la population d'un pays?¹²» Malheureusement, la disponibilité des données à ce moment rendait cette analyse pratiquement impossible.

Bref, les modèles créés pour ces deux études peuvent être utilisés afin de tester l'hypothèse que les dépenses dans la recherche et le développement en santé améliorent la santé de la population. En effet, ceux-ci nous permettront de choisir les variables pour caractériser les différentes composantes pouvant potentiellement avoir des effets directs sur la santé des gens. Ces études nous permettent également de connaître les différents problèmes à considérer afin de pouvoir analyser correctement les résultats obtenus. Les prochaines sections se concentreront sur ces différents aspects.

Caractérisation du modèle et des variables

Comme point de départ pour notre propre analyse, nous devons d'abord déterminer les différentes variables à utiliser pour créer notre modèle économétrique, ce qui nous aidera à tester notre hypothèse de départ. Afin de caractériser la santé de la population, nous utiliserons l'un des indicateurs de la santé de Nixon et Ulmann ainsi que de Crémieux, Ouellette et Pilon, soit l'espérance de vie. Également, nous estimerons le modèle en prenant l'espérance de vie des hommes et des femmes séparément afin de déterminer si les effets des investissements dans la recherche et le développement en santé sont différents pour les deux sexes. En résumé, trois modèles seront estimés en

¹² Crémieux, Ouellette et Pilon. (1999) «Health Care Spending as Determinants of Health Outcomes», *Journal of Health Economics*, vol 8 n° 1, p.637.

utilisant les différentes variables dépendantes décrites ci-dessus. Il reste maintenant à déterminer les différents indicateurs pouvant avoir un effet sur ces variables.

Bien entendu, la variable indépendante principale de notre modèle consistera aux investissements dans la recherche et le développement dans le domaine de la santé, puisque l'objectif de cette recherche consiste à évaluer l'impact de ce type d'investissement sur la santé de la population. Néanmoins, afin de pouvoir bien caractériser le modèle, nous devons prendre en considération d'autres aspects qui peuvent avoir un effet sur la santé de la population. D'après le modèle logique développé par l'ACSS disponible en annexe 2, la qualité des services et des systèmes de santé ainsi que les différents déterminants de la santé sont les éléments qui ont l'impact le plus direct sur la santé de la population. Afin de prendre en considération la qualité des services et des systèmes de santé, nous prendrons comme indicateur le nombre de lits d'hôpitaux par habitant. Pour ce qui est des déterminants de la santé, la consommation de fruits et de légumes, la consommation de tabac et la consommation d'alcool sont les indicateurs choisis afin de représenter les comportements des gens, tandis que le taux de chômage représentera les conditions de travail des gens. Or, toutes ces variables, à l'exception du taux de chômage, sont des indicateurs développés par l'ACSS afin de représenter les déterminants de la santé. De plus, ceux-ci ont tous été utilisés par Nixon et Ulmann lors du développement de leur modèle. Finalement, nous avons également cru important d'inclure deux autres variables afin de représenter la troisième étape du modèle logique. Le nombre de médecins praticiens ainsi que le nombre de diplômés en médecine ont donc été utilisés afin de représenter le développement de la capacité au niveau de la santé. Le

tableau 1 offre un résumé des variables dépendantes et indépendantes choisies ainsi qu'une description de chacune d'elles.

Tableau 1 : Descriptions des variables

Variables dépendantes	Description
popLE	Espérance de vie de la population à la naissance
maleLE	Espérance de vie des hommes à la naissance
femLE	Espérance de vie des femmes à la naissance
Variables indépendantes	Description
HealthRD	Dépenses totales en recherche et développement en santé (en millions de dollars US, PPP)
Physicians	Nombre de médecins praticiens par 1000 habitants
medGrads	Nombre de diplômés en médecine par 100 000 habitants
HospitalBeds	Nombre de lits d'hôpitaux par 1000 habitants
FruitCons	Consommation de fruits et de légumes (kilos par habitant)
AlcoholCons	Consommation d'alcool (litres par habitant âgé de 15 ans et plus)
TobaccoCons	Consommation de tabac (grammes par habitant âgé de 15 ans et plus)
Unemployment	Taux de chômage (% de la population active)

Il est important de noter à ce point que toutes les variables choisies pour cette étude reflètent la disponibilité des données provenant de la base de données Éco-Santé 2010 offerte par l'Organisation de Coopération et de Développement Économique (OCDE). Cette base de données contient de l'information sur plusieurs aspects liés à la santé pour 30 pays membres de l'OCDE durant la période 1960-2009. Malheureusement, plusieurs informations n'étaient pas disponibles pour certains pays durant cette période, ce qui nous a obligés à restreindre notre échantillon à sept pays : Canada, Australie, États-Unis, Finlande, France, Allemagne et République Tchèque, pour la période 1995-

2007. La prochaine section offrira un premier aperçu des données et de l'échantillon afin de déterminer certains problèmes que nous devons considérer lors de l'analyse.

Premier aperçu des données et de l'échantillon

Avant de parfaire des analyses économétriques, il est toujours important de bien comprendre l'étendue des données et de l'échantillon. C'est pourquoi il est primordial de débiter notre analyse en étudiant les statistiques descriptives de ces deux composantes pour chacune des variables dépendantes et indépendantes. Le tableau disponible en annexe 3 nous résume la moyenne, l'écart type ainsi que le nombre d'observations pour chacune des composantes. Toutefois, puisque nous travaillons avec des données en panel, il est important d'analyser cette information, non pas seulement au niveau global mais également au niveau de la partie «between», c'est-à-dire entre les différents pays, et au niveau de la partie «within», c'est-à-dire à travers la période.

Comme nous pouvons le constater d'après ce tableau, le nombre maximal d'observations pour une variable est 91, soit pour 7 pays et pour une période de temps de 13 ans. Toutefois, l'information n'est pas complète pour certaines variables. En effet, les données parmi les pays ne sont pas toujours disponibles durant la totalité de la période. Cela nous oblige donc à éliminer certaines observations qui ont des données manquantes.

Lorsque nous comparons la disponibilité des données pour chacune des variables, nous arrivons à la conclusion qu'il sera possible de faire une analyse économétrique avec 84 observations au total, où les 7 pays seront représentés et où la période de temps variera

entre 10 et 13 ans pour les pays, pour une moyenne globale de 12 ans. De plus, les statistiques que l'on retrouve dans le tableau en annexe 3 nous permettent de constater qu'il semble y avoir une certaine variation entre les pays pour la plupart des variables. En effet, lorsque nous regardons les écarts types de la partie «between», ceux-ci sont parfois assez importants. En particulier, l'écart type de la partie «between» de la variable représentant les dépenses dans la recherche et le développement en santé est assez important. Cette observation peut nous permettre de croire qu'il y aura présence d'hétéroscédasticité lorsque nous ferons nos estimations. Dans la section suivante, nous effectuerons des tests plus approfondis afin de déterminer si la variance des termes d'erreurs varie d'une observation à l'autre. Toutefois, afin d'alléger ce problème, nous utiliserons le logarithme naturel des dépenses en recherche et développement dans le domaine de la santé. Cette procédure nous aidera à stabiliser la variation de cette composante entre les pays.

Vérification des hypothèses statistiques

D'après Gujarati, les données en panel offrent «plus de données informatives, plus de variabilité, moins de colinéarité parmi les variables, plus de degrés de liberté et plus de performance»¹³, contrairement aux séries temporelles et aux coupes instantanées ordinaires. Toutefois, en exécutant des régressions avec ce type de données, des problèmes comme l'hétéroscédasticité et l'autocorrélation surviennent très souvent. La

¹³ Gujarati, Damodar N. (2004), Économétrie, 4^e édition, Bruxelles, Belgique, Édition de boeck, p.635

vérification de ce type d'hypothèse nous permet alors de choisir la méthode à utiliser afin d'estimer les paramètres du modèle.

Les données en panel font souvent en sorte que la variance des termes d'erreurs diffère d'un pays à l'autre. De plus, cette variance peut souvent être corrélée avec le temps parmi les différents pays. Malgré le fait que ces problèmes ne biaiseront pas les résultats obtenus, ceux-ci feront en sorte que les tests statistiques utilisés afin de vérifier la significativité des résultats, ne seront plus valables. C'est pourquoi il est important de corriger le problème si celui-ci est présent.

Afin de tester ces hypothèses, nous estimerons tout d'abord un modèle économétrique simple à l'aide de la méthode des moindres carrés ordinaires. Bien que cette méthode ne prenne pas totalement en considération le fait que nous travaillons avec des données en panel, celle-ci nous permettra d'utiliser le test de Breush-Pagan-Godfrey afin de tester la présence d'hétéroscédasticité. Ce test pose comme hypothèse nulle que la variance des termes d'erreurs est constante. Or, lorsque nous effectuons cette procédure, nous obtenons une statistique de 3,37 et une valeur-p de 0,0663. Ce résultat ne nous permet donc pas de rejeter l'hypothèse nulle à un niveau de confiance de 95%. Toutefois, lorsque le niveau de significativité est de 1%, nous devons rejeter l'hypothèse nulle, ce qui suggère une présence d'hétéroscédasticité. Il est donc difficile d'affirmer que la variance des résidus est bel et bien constante. Ce résultat suggère que nous devons peut-être utiliser une méthode d'estimation qui corrige les résidus. Néanmoins, nous pouvons tester si les résidus sont corrélés avec le temps afin de voir si une telle procédure est réellement nécessaire.

Pour ce faire, nous pouvons d'abord calculer les résidus obtenus à partir de la régression précédente et les observer en fonction du temps. D'après les graphiques en annexe 4, les termes d'erreurs semblent corrélés avec le temps pour chacun des pays pour lesquels nous avons fait ces observations. Les États-Unis montrent une tendance ascendante à travers le temps, tandis que le Canada et la France montrent une tendance cyclique des résidus durant la période 1995-2007. L'Allemagne tant qu'à elle démontre une certaine tendance cyclique mais beaucoup moins prononcée. Ces observations suggèrent donc qu'il y aurait présence d'autocorrélation. Toutefois, afin de s'assurer que nous rencontrons bel et bien ce type de problème, nous pouvons utiliser le test que Wooldridge a développé en 2002, afin de vérifier l'autocorrélation sérielle d'ordre 1. L'hypothèse nulle de ce test suggère qu'il n'y a pas d'autocorrélation sérielle. Lorsque nous effectuons ce test à l'aide du logiciel STATA, nous obtenons une statistique de 10,213. Puisque cette statistique suit une loi de Fisher avec 1 degré de liberté pour le numérateur et 6 degrés de liberté pour le dénominateur, nous pouvons rejeter l'hypothèse nulle. Ce résultat suggère donc que le problème d'autocorrélation est présent.

Or, ce problème d'autocorrélation jumelé avec la possibilité que nous fassions face à un problème d'hétéroscédasticité nous force à utiliser une méthode d'estimation qui ajuste les variances. C'est pourquoi nous utiliserons la méthode des moindres carrés généralisés. À l'aide de cette méthode, nous estimerons les mêmes paramètres que ceux obtenus en utilisant la méthode des moindres carrés ordinaires, mais la variance de ceux-ci sera corrigée afin d'éliminer le problème d'autocorrélation et la présence d'hétéroscédasticité.

Toutefois, l'un des problèmes économétriques le plus sérieux que l'on peut rencontrer lorsque nous construisons un modèle de ce type consiste à la mauvaise spécification du modèle due à l'omission de variable indépendante importante. Un problème de ce type rend les estimateurs biaisés, ce qui peut faire en sorte que les résultats obtenus ne peuvent pas être interprétés de la bonne façon. Afin de tester cette hypothèse, nous utiliserons le test développé par Ramsey avec l'aide du modèle créé pour tester l'hétéroscédasticité et l'autocorrélation. L'hypothèse nulle de ce test suggère que le modèle estimé n'a pas de variables omises. Lorsque nous exécutons ce test à l'aide du logiciel STATA, nous obtenons une statistique de 2,46. Puisque la valeur-p de cette statistique est de 0,0707, nous ne pouvons pas rejeter l'hypothèse nulle à un niveau de confiance de 95%. Ce résultat suggère donc qu'il ne semble pas y avoir de variables omises à l'intérieur du modèle.

Comme prochain aspect technique, nous regarderons quel type de modèle économétrique nous devons utiliser afin de tester notre hypothèse. En effet, l'utilisation des données en panel offre généralement deux modèles différents afin de faire les estimations voulues. Le modèle des effets fixes permet une variation de la valeur de l'ordonnée à l'origine pour les pays en ajoutant des variables dichotomiques les représentant à l'intérieur du modèle. Il est nécessaire d'estimer ce type de modèle lorsque les variables indépendantes de celui-ci varient selon le pays. Si cette méthode est ignorée lorsque celle-ci est nécessaire, un biais dans les paramètres estimés se fera ressentir à cause de l'omission de variables importantes. Le modèle des effets aléatoires quant à lui suggère que l'ordonnée à l'origine varie d'un pays à l'autre de façon aléatoire et que cette

variation est reflétée à l'intérieur du terme d'erreur. Cela fait donc en sorte que l'ordonnée à l'origine du modèle est le même pour chaque pays.

En 1978, un économiste du nom de Jerry A. Hausmann a développé un test afin de déterminer quel modèle était le plus approprié. Le test fut basé sur le fait que les termes d'erreurs individuels du modèle des effets aléatoires sont corrélés avec la variable indépendante de celui-ci. «L'hypothèse nulle du test dans ce cas est que les estimateurs du modèle à effets fixes sont différents de ceux obtenus avec le modèle à effets aléatoires¹⁴». Afin de tester cette hypothèse, nous devons estimer les deux modèles pour ensuite comparer les résultats. Toutefois, comme nous avons déterminé précédemment que la méthode des moindres carrés généralisés était plus appropriée, nous ferons nos estimations à l'aide de cette méthode. En annexe 5, nous retrouvons les résultats des deux régressions ainsi que le résultat du test de Hausmann. Notons à ce moment que pour effectuer ces tests, nous avons utilisé l'espérance de vie de la population comme variable dépendante. Comme nous pouvons le constater, la statistique du test de Hausmann obtenue en comparant les paramètres du modèle est de 12750,50. Puisque cette statistique suit une distribution χ^2 , nous pouvons rejeter l'hypothèse nulle. Ce résultat suggère donc que le modèle des effets fixes est nécessaire afin de faire nos estimations.

Le modèle que nous estimerons sera donc caractérisé par l'équation suivante :

$$(1) \quad Health_{i,t} = \beta_0 + \beta_1 \ln healthrd_{i,t} + \beta_2 physicians_{i,t} + \beta_3 hospitalbeds_{i,t} + \beta_4 fruitcons_{i,t} + \beta_5 alcoholcons_{i,t} + \beta_6 tobaccocons_{i,t} + \beta_7 unemployment_{i,t} + \beta_8 medgrads_{i,t} + \gamma_i + \epsilon_{i,t}$$

¹⁴ Gujarati, Damodar N. (2004), *Économétrie*, 4^e édition, Bruxelles, Belgique, Édition de boeck, page 647.

L'indice i à l'intérieur de l'équation est l'unité représentant les pays tandis que l'indice t est celui représentant l'unité de temps. Le paramètre γ correspond aux variables dichotomiques ajoutées au modèle afin de représenter les différents pays. Le paramètre β_1 quant à lui représente l'impact de l'investissement en recherche et développement dans le domaine de la santé, sur la santé de la population. Toutefois, deux aspects restent encore à être vérifiés afin de pouvoir analyser les résultats correctement.

L'un de ces aspects consiste en la vérification de la multicolinéarité entre les variables indépendantes. En effet, lorsque ce problème est présent, la variance des paramètres tend à s'accroître, ce qui peut faire en sorte qu'il sera très difficile d'obtenir des résultats significatifs. L'une des façons de tester si nous rencontrons ce problème, est d'observer les coefficients de corrélation entre les différentes variables. D'après Gujarati, «si le coefficient de corrélation d'ordre 0 entre deux régresseurs est élevé, par exemple supérieur à 0,8, la multicolinéarité est un sérieux problème¹⁵.» En annexe 6, nous observons ces corrélations entre les différentes variables indépendantes. D'après les résultats, seul le coefficient de corrélation entre le nombre de médecins praticiens et la consommation d'alcool est élevé. Ce résultat suggère qu'il y aurait multicolinéarité forte entre ces deux variables. Pour ce qui est des autres coefficients de corrélations, certains d'entre eux, c'est-à-dire là où la corrélation entre deux variables indépendantes est d'une valeur supérieure à 0,4 mais inférieure à 0,8, semblent indiquer qu'il y aurait un faible risque de faire face à ce problème. Par exemple, le coefficient de corrélation décrivant la

¹⁵ Gujarati, Damodar N. (2004), Économétrie, 4^e édition, Bruxelles, Belgique, Édition de boeck, page 363.

relation entre le nombre de médecins praticiens et le nombre de lits d'hôpitaux est de 0,75.

Une autre façon de vérifier cet aspect consiste à vérifier si nous obtenons un coefficient de détermination élevé mais des statistiques t peu significatives. Lorsque nous avons estimé le modèle afin de tester l'hypothèse d'homoscédasticité, nous avons obtenu un coefficient de détermination de 0,9844. De plus, les statistiques t obtenues suggèrent que seuls les paramètres estimés pour les variables représentant la consommation de tabac, la consommation d'alcool et le taux de chômage sont non-significatifs. Cela suggère donc que la multicollinéarité ne semble pas excessivement forte.

Ces résultats contradictoires ne nous permettent donc pas de dire si oui ou non, nous rencontrons ce problème lors de nos estimations. Toutefois, il est parfois essentiel d'inclure des variables importantes à l'intérieur de notre modèle malgré le fait que cette inclusion provoque de la multicollinéarité. En effet, parfois exclure de telles variables peut causer un biais dans l'estimation de nos paramètres dû à une mauvaise spécification du modèle. Or, la multicollinéarité ne cause pas de biais, ce qui est un aspect important lors de l'analyse des résultats. De plus, inclure des variables de ce type peut contribuer à augmenter le coefficient de détermination qui permet d'évaluer la qualité de l'ajustement du modèle. C'est pourquoi nous continuerons le développement du modèle en question en ne considérant pas les conséquences de ce problème.

Le dernier aspect à considérer avant de rapporter les résultats consiste à vérifier l'hypothèse de normalité des résidus. En effet, afin que les tests d'hypothèses vérifiant la significativité des paramètres soient valables, il faut que les résidus suivent une loi

normale. Si ceux-ci ne répondent pas à cette exigence, le seul moyen de contourner ce problème est d'avoir un échantillon assez grand. En effet, par la loi des grands nombres, lorsque l'échantillon est assez grand, les résidus convergeront vers une loi normale. Malheureusement, le nombre d'observations utilisées afin de faire nos régressions n'est que de 84. Ce petit échantillon nous force donc à s'assurer que cette hypothèse est respectée. En annexe 7, nous pouvons observer l'histogramme de ces résidus. D'après ce graphique, il ne semble pas y avoir de problème de normalité à première vue. Néanmoins, afin de s'assurer que la distribution des résidus suit bel et bien une loi normale, nous utiliserons le test de Skewness-Kurtosis. Ce test assume comme hypothèse de départ que la distribution des résidus suit une loi normale. En utilisant le logiciel STATA, nous obtenons une statistique de 0,281. Puisque la valeur-p de cette statistique est de 0,281, nous ne pouvons pas rejeter l'hypothèse nulle à un niveau de confiance de 95%, ce qui suggère que les résidus suivent une loi normale. Tous ces tests statistiques nous permettent d'avoir une certaine confiance dans les résultats du modèle estimé par l'équation (1). Ceux-ci démontrent également que le modèle qui a été élaboré possède une certaine puissance et une certaine fiabilité. C'est pourquoi nous pouvons maintenant passer à l'étape de la description des résultats.

Description des résultats du modèle statique

D'après les résultats de l'annexe 8, le paramètre décrivant l'impact des investissements au niveau de la santé de la population en général, est significatif et positif au niveau de confiance de 95%. Cela fait donc en sorte que ces investissements augmentent significativement l'espérance de vie de la population. Afin de représenter

l'ampleur de ce résultat, il faut considérer le fait que puisque les dépenses dans la recherche en santé est sous forme logarithmique, et que l'espérance de vie est sous forme normale, la relation entre les deux variables peut être représentée par la fonction suivante :

$$(2) \quad \Delta PopLE = \frac{\beta_1}{100} * \Delta healthrd \%$$

D'après cette relation, une augmentation de 1% des dépenses dans la recherche et le développement en santé augmentera l'espérance de vie de la population de $\frac{\beta_1}{100}$ année. D'après les résultats de l'annexe 8, si un pays double ses investissements, l'espérance de vie augmentera de 0,3547 année. De l'autre côté, si nous réorganisons l'équation (2), nous pouvons facilement constater qu'il faut une augmentation de 282% des dépenses dans la recherche et le développement en santé afin qu'un pays augmente l'espérance de vie de sa population d'une année. Un résultat similaire est également obtenu lorsque la variable dépendante correspond à l'espérance de vie des hommes. D'après les résultats, afin que cet indicateur augmente d'une année, un pays devra augmenter ses dépenses dans la recherche et le développement en santé d'environ 245%. Ce résultat n'est toutefois pas significatif pour l'espérance de vie des femmes. Notons à ce moment que ces augmentations supposent que toutes les autres variables indépendantes sont constantes.

En ce qui concerne les autres variables, l'indicateur qui semble avoir le plus d'impact sur la santé, tant au niveau de la population qu'au niveau des hommes et des femmes, est sans aucun doute l'augmentation dans le nombre de médecins praticiens.

Toutefois, cette augmentation sera plus importante au niveau des hommes qu'au niveau des femmes. En ce qui concerne le nombre de diplômés en médecine et la consommation de fruits et de légumes, les résultats suggèrent qu'une augmentation de ces deux aspects améliorera la santé des gens, mais de façon modérée puisque les coefficients sont inférieurs à 0,1 dans les deux cas. Le résultat le plus surprenant de notre étude est sans doute le fait qu'une augmentation dans le nombre de lits d'hôpitaux diminuera significativement l'espérance de vie des hommes, des femmes et de la population en général. Ce résultat est peut-être une indication qu'une variable importante a été omise lors de la création du modèle. En ce qui concerne la consommation de tabac et d'alcool ainsi que le taux de chômage, les résultats suggèrent que ces indicateurs n'ont pas d'effet significatif sur l'espérance de vie. En ce qui a trait aux paramètres des variables représentant chacun des pays, les résultats suggèrent que les États-Unis sont ceux ayant l'espérance de vie la plus faible. La Finlande quant à elle aurait, en moyenne, l'espérance de vie la plus élevée. Comme nous pouvons le constater, les résultats obtenus à partir du modèle statique soutiennent la plupart de nos hypothèses, à l'exception de la variable représentant le nombre de lits d'hôpitaux.

Le modèle dynamique

Le problème majeur avec le modèle statique concerne le fait que les effets de retard ne sont pas pris en considération. Afin de remédier à cette situation, nous pouvons estimer le modèle de l'équation (1) mais en considérant un retard en ce qui a trait à la variable représentant les dépenses dans la recherche et le développement en santé. Le

problème est qu'à ce jour, personne n'est capable de déterminer à quel moment ces investissements auront un impact sur la santé de la population. Toutefois, si l'on émet l'hypothèse que ces investissements peuvent avoir des effets instantanés, l'une des méthodes utilisées afin de déterminer le moment où les effets de ces investissements se font ressentir est la méthode d'estimation ad hoc du modèle à retards échelonnés. Cette méthode consiste à estimer le modèle de l'équation (1) en y ajoutant des variables explicatives retardées, jusqu'à ce que les coefficients estimés deviennent non significatifs. Dans ce cas, le modèle estimé est de type :

(3)

$$\begin{aligned} Health_{i,t} = & \beta_0 + \beta_1 lnhealthrd_{i,t} + \sum_{j=1}^{\infty} \beta_{j+1} lnhealthrd_{t-j} + \delta_1 physicians_{i,t} + \\ & + \delta_2 hospitalbeds_{i,t} + \delta_3 fruitcons_{i,t} + \delta_4 alcoholcons_{i,t} + \\ & + \delta_5 tobaccocons_{i,t} + \delta_6 unemployment_{i,t} + \delta_7 medgrads_{i,t} \gamma_i + \epsilon_{i,t} \end{aligned}$$

La séquence sommative des β_j illustrée à l'intérieur de cette sommation s'arrête dès qu'un estimateur devient non significatif. Lorsque nous estimons ce modèle pour $j=1$ à l'aide du logiciel STATA, nous nous rendons compte que les coefficients β_1 et β_2 de celui-ci ne sont plus significatifs. D'après Gujarati, «les retards ont tendance à être fortement corrélés.¹⁶» Or, cette multicollinéarité peut également faire augmenter la variance des paramètres, ce qui rend difficile d'obtenir des résultats significatifs. De plus, lorsqu'on estime ce type de modèle, le nombre de degrés de liberté diminue. Cette diminution jumelée à un petit échantillon rend les analyses statistiques beaucoup plus difficiles. Bref, afin de faire ce type de modèle, il faut augmenter le nombre d'observations de notre échantillon. Malheureusement, les données offertes par l'OCDE ne nous permettent pas

¹⁶ Gujarati, Damodar N. (2004), *Économétrie*, 4^e édition, Bruxelles, Belgique, Édition de boeck, page 662.

de faire cela. Cet aspect du rapport représente l'une des limitations de celui-ci qui sont mentionnées dans la section suivante.

Limitations et recommandations

D'après nos résultats, les variables représentant le comportement des individus n'ont pas d'effets significatifs sur l'espérance de vie des gens. Ce résultat découle sans doute du fait que l'impact de ces comportements risque de se faire ressentir avec un effet de retard, tout comme celui observé avec les dépenses dans la recherche et le développement en santé. Malheureusement, le manque de disponibilité des données nous empêche de faire ce type d'analyse dynamique. De plus, cet effet de retard risque fortement de varier selon le domaine dans lequel ces investissements sont faits. Cet aspect nous mène à la première recommandation de ce rapport.

En effet, la recherche et le développement en santé se divisent en plusieurs domaines. Or, l'impact de ces investissements diffère grandement selon les domaines. Certains d'entre eux ont un impact à court terme, tandis que d'autres ont des effets à long terme. Par exemple, les recherches financées lors la crise AH1N1 ont permis de développer un vaccin pour prévenir l'épidémie. L'impact de ces recherches s'est donc fait ressentir presque immédiatement. De l'autre côté, les recherches financées pour la maladie du VIH et du SIDA ne cessent d'augmenter. Malheureusement, personne n'a à ce jour développé un traitement pour une guérison complète du virus. Toutefois, ces recherches ont permis d'améliorer la qualité de vie des personnes qui en souffrent. L'effet de retard de l'impact de ces recherches est donc beaucoup plus important dans ce cas. C'est pourquoi les chercheurs devraient commencer à s'intéresser à l'impact des

investissements en recherche et développement pour des domaines spécifiques. Malheureusement, très peu de données sont disponibles afin de pouvoir faire ce type d'analyse. Par exemple, beaucoup d'informations sont disponibles concernant le taux de mortalité associé à certaines maladies. Par contre, l'information concernant les investissements dans la recherche et le développement par domaine est très difficile à obtenir. Différentes institutions canadiennes comme les IRSC comptabilisent ce type d'information, mais les agences statistiques mondiales ne comptabilisent que les investissements en général. Il faudrait donc développer une méthode pour rendre l'information disponible à l'avenir afin que l'on puisse faire ce type d'analyse.

Comme deuxième considération, il est clair que l'impact au niveau de la santé de la population n'est qu'une petite partie de l'impact global lié à la recherche et au développement en santé. Comme le modèle de l'ACSS l'illustre bien, plusieurs aspects sont à considérer afin de pouvoir mesurer l'impact global des investissements dans ce domaine. Toutefois, ce rapport illustre bien les défis liés à l'étude de ce problème pour une catégorie particulière, c'est-à-dire l'impact au niveau de la santé de la population. C'est pourquoi, évaluer l'impact global de ces investissements est sans doute une tâche ardue qui nécessite une recherche encore plus extensive du sujet. Il faudra notamment déterminer les besoins en termes de données et d'expertises afin de pouvoir accomplir cette tâche.

Bref, il est clair qu'il reste beaucoup de travail à faire avant de pouvoir bien évaluer l'impact de la recherche et du développement en santé. Malgré le fait que certains résultats intéressants sont identifiés dans ce rapport, plusieurs aspects sont encore à être

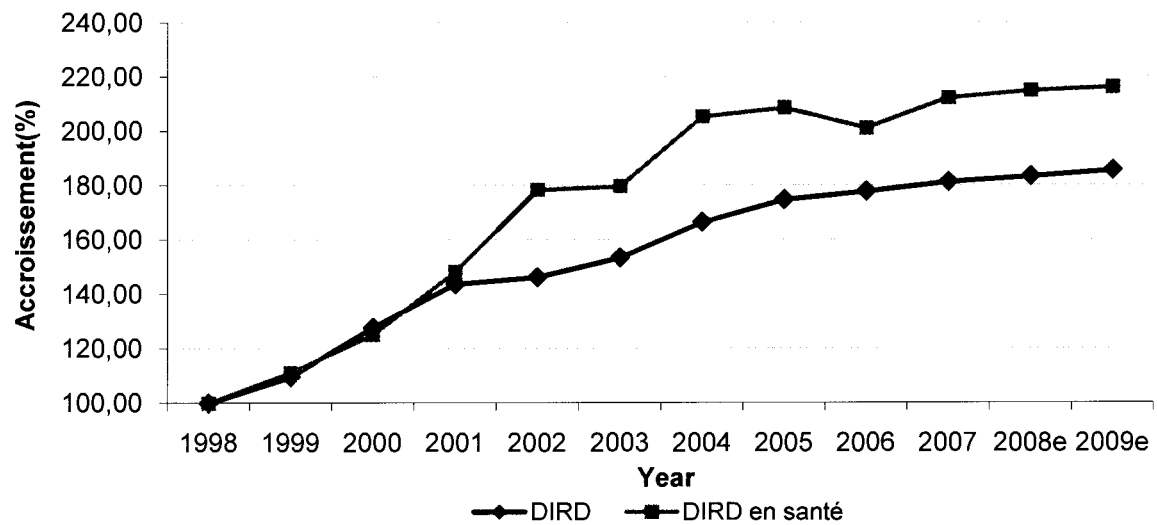
considérés afin de pouvoir bien identifier cet impact. La complexité de cette étude démontre le défi auquel les différentes institutions finançant ce type de recherche sont confrontées chaque jour. De plus, plusieurs indicateurs tels que le vieillissement de la population et les changements climatiques poussent les experts à croire que ce financement continuera à augmenter durant les prochaines années. Cela démontre également pourquoi ce type de recherche est important pour la société et continuera à l'être durant plusieurs années encore. C'est pourquoi il faut continuer à développer l'expertise dans ce domaine afin de pouvoir un jour évaluer les bénéfices ou l'impact global de la recherche et du développement dans le domaine de la santé.

Bibliographie

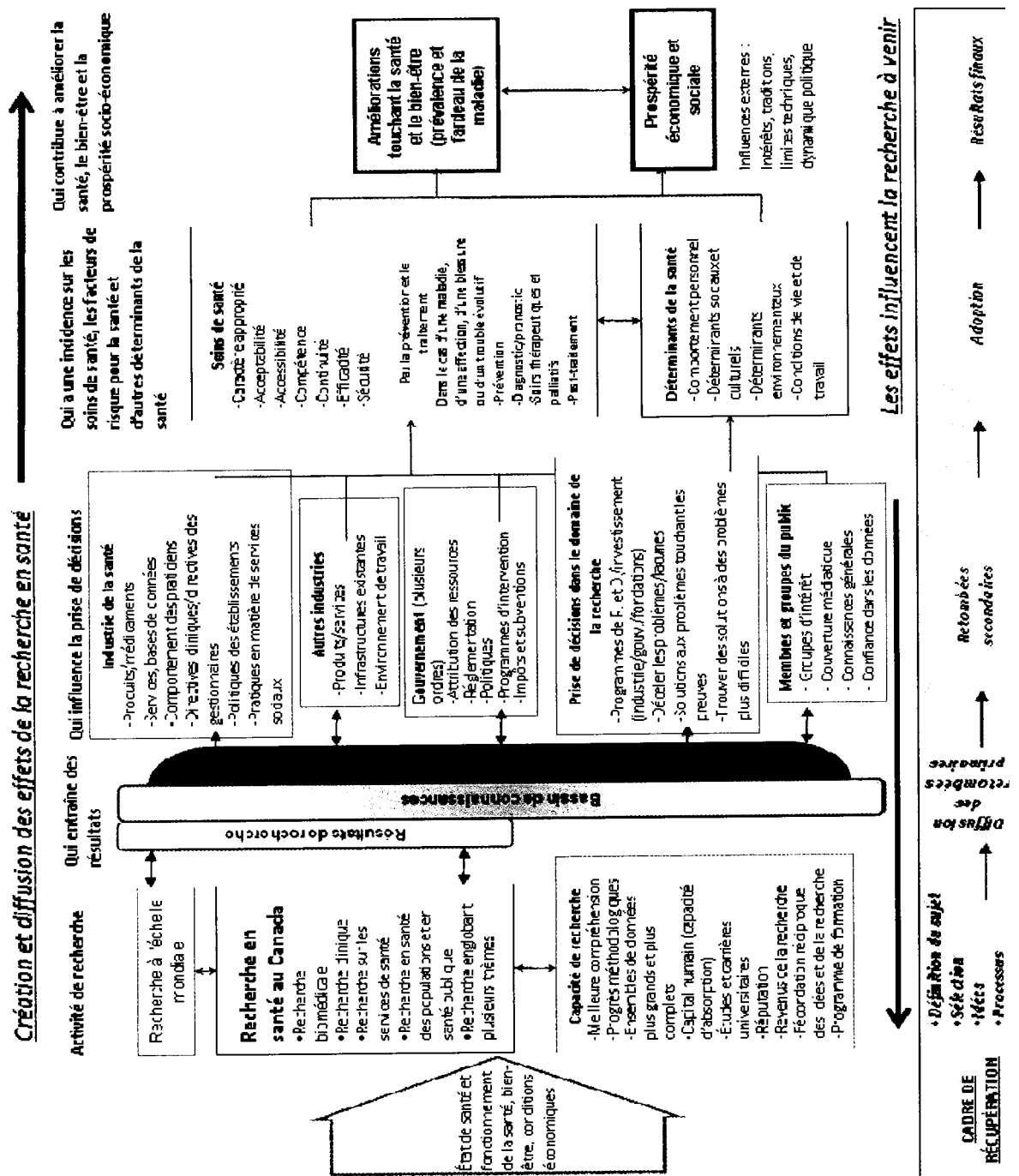
- Académie canadienne des sciences de la santé (2009), «*Créer un impact* », http://www.caahs-acss.ca/e/pdfs/ROI_FrenchReport.pdf
- Buxton MJ, Hanney S (1996) «*How can payback from health services research be assessed?*», *Journal of Health Services Research & Policy*, vol 1, no 1
- Cameron, Trivedi. (2005), *Microeconometrics*, 1^{ère} édition, États-Unis, Édition Cambridge
- Crémieux, Ouellette et Pilon. (1999) «Health Care Spending as Determinants of Health Outcomes», *Journal of Health Economics*, vol 8 n° 1, pp. 627-239
- Greene, William H. (2008), *Econometric Analysis*, 6^e édition, New-Jersey, États-Unis, Édition Pearson Prentice Hall
- Gujarati, Damodar N. (2004), *Économétrie*, 4^e édition, Bruxelles, Belgique, Édition de boeck
- Institut canadien d'Information sur la santé, «Tendances des dépenses nationales de la santé», http://secure.cihi.ca/cihiweb/products/National_health_expenditure_trends_1975_to_2009_fr.pdf
- Institut de Recherche en Santé du Canada, «Élaboration d'un cadre des IRSC pour mesurer l'impact de la recherche en santé», http://www.cihr-irsc.gc.ca/f/documents/meeting_synthesis_f.pdf
- Nixon, J, Ulmann P. (1996) «The Relationship between Health Care Expenditures and Health Outcomes», *The European Journal of Health Economics*, vol.7, 1 n° 1.
- Statistiques Canada, «*Science Statistics*», <http://www.statcan.gc.ca/pub/88-001-x/88-001-x2010003-eng.pdf>

Annexe

Annexe 1: Croissance des dépenses brutes dans la recherche et le développement en santé versus les dépenses totales en recherche et développement depuis 1998



Annexe 2 : Modèle logique du Cadre de l'ACSS17

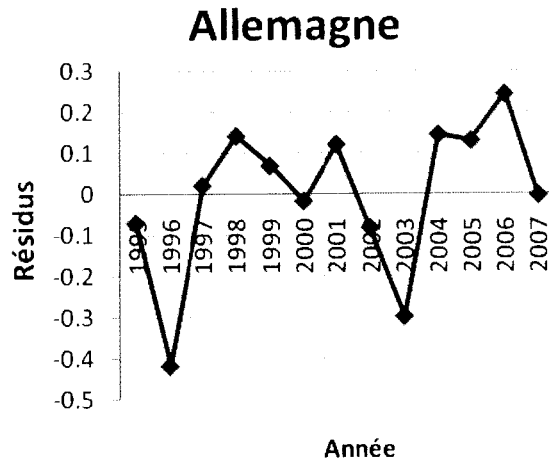
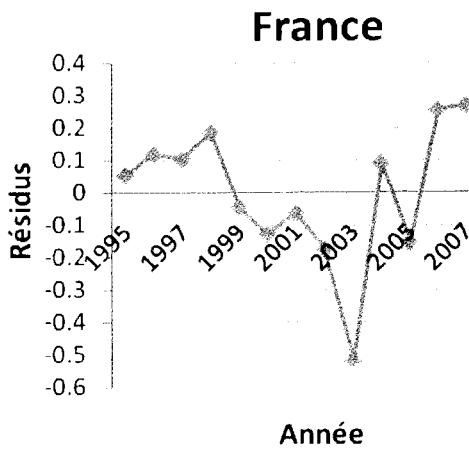
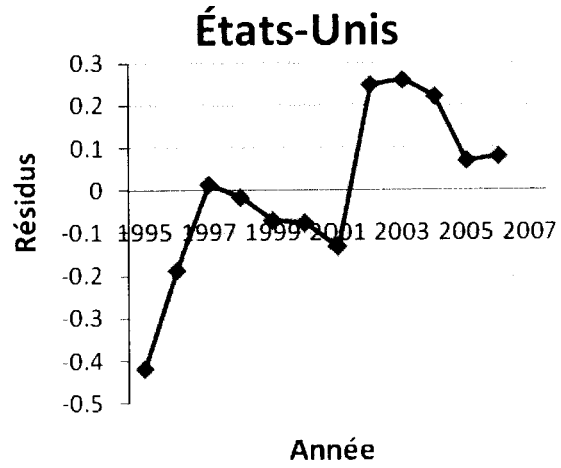
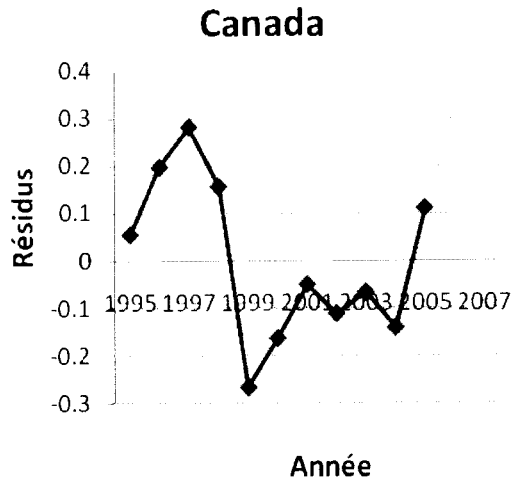


¹⁷ Académie Canadienne des sciences de la santé (2009), «Créer un impact », http://www.caahs-acss.ca/e/pdfs/ROI_FrenchReport.pdf, p. 19, Consulté le 24 juillet 2010.

Annexe 3 : Les statistiques descriptives

Variables dépendantes		Moyenne	Écart Type	Observations
popLE	total	78,11	1,76	N = 91
	between		1,59	n = 7
	within		0,95	T = 13
maleLE	total	75,01	1,96	N = 91
	between		1,73	n = 7
	within		1,12	T = 13
femLE	total	81,18	1,65	N = 91
	between		1,56	n = 7
	within		0,78	T = 13
Variables indépendantes		Moyenne	Écart Type	Observations
HealthRD	total	6031,22	10589,99	N = 90
	between		10698,66	n = 7
	within		3496,97	T = 12,86
lnHealthRD	total	7,24	1,93	N = 90
	between		2,06	n = 7
	within		0,35	T = 12,86
physicians	total	2,77	0,51	N = 91
	between		0,53	n = 7
	within		0,14	T = 13
medGrads	total	8,01	2,29	N = 91
	between		2,11	n = 7
	within		1,17	T = 13
HospitalBeds	total	6,26	2,18	N = 90
	between		2,29	n = 7
	within		0,42	T = 12,86
fruitcons	total	196,30	35,69	N = 91
	between		36,80	n = 7
	within		9,99	T = 13
alcoolcons	total	10,16	2,06	N = 91
	between		2,15	n = 7
	within		0,47	T = 13
tobaccoCons	total	1654,52	527,84	N = 89
	between		520,08	n = 7
	within		199,82	T = 12,71
Unemployment	total	7,89	2,19	N = 89
	between		1,81	n = 7
	within		1,46	T = 12,57

Annexe 4 : Détection de la présence d'autocorrélation pour quatre pays



Annexe 5 : Test de Hausmann pour le modèle à effets fixes

VARIABLES (pople)	Effet fixes	Effets aléatoires
Inhealthrd	0,3547** [0,1412]	-4308*** [0,9620]
canada	4,3964*** [0,3857]	
australia	4,1206*** [0,4900]	
finlande	7,3867*** [0,6452]	
france	5,0014*** [0,6381]	
allemagne	5,4788*** [0,6779]	
reptcheque	2,7996*** [0,7237]	
physicians	3,1041*** [0,2641]	3,3574*** [0,4321]
medgrads	0,0802** [0,0329]	0,4225*** [0,0864]
hospitalbeds	-1,1048*** [0,1372]	-0,3826*** [0,1032]
fruit_vegetables	0,0094*** [0,0036]	0,0536*** [0,0071]
alcoholcons	-0,0162 [0,0642]	-0,0210 [0,0954]
tobaccocons	-0,0001 [0,0002]	-0,0036*** [0,0003]
unemployment	0,006 [0,0260]	0,1338** [0,0635]
constant	67,5083*** [1,9426]	65,1632*** [1,6760]
Wald χ^2	5317,7	285,77

Observations		84
Nombre de groupes		7
Observations par groupe	min	10
	moyenne	12
	max	13

Écart type entre parenthèses

*** p<0.01, ** p<0.05, * p<0.1

Test: H0: différence significative des coefficients

χ^2 = 1270,50
p-value = 0,0000

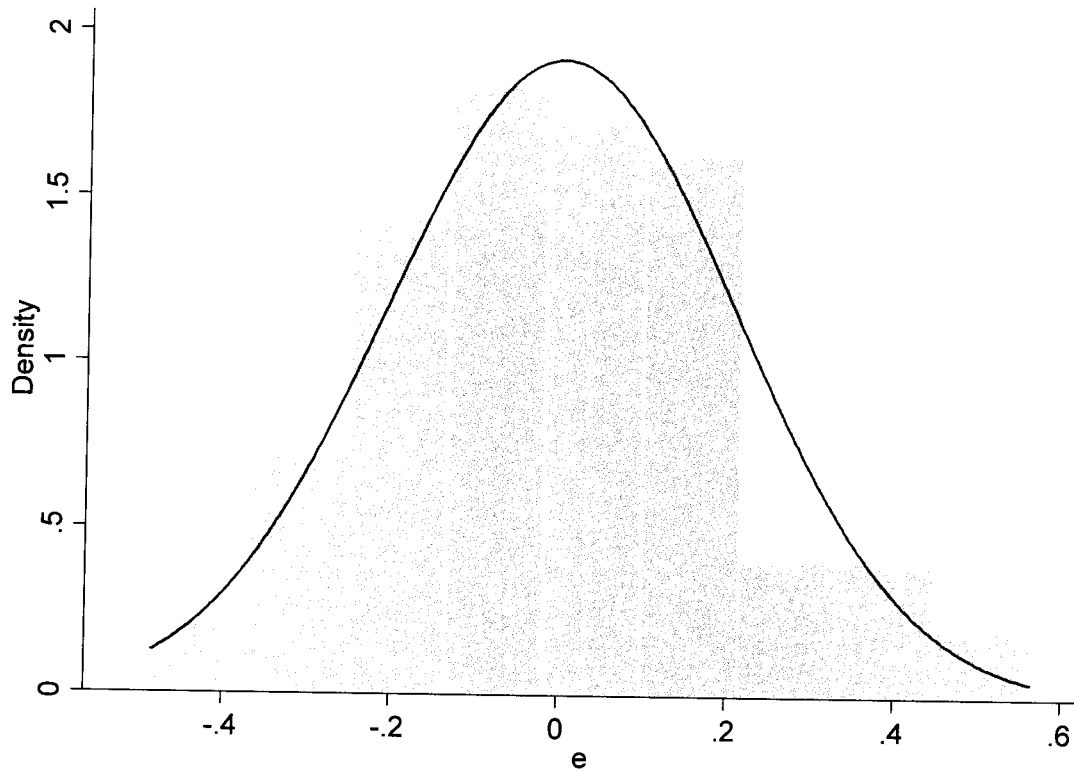
Annexe 6 : Détection de la multicollinéarité

Coefficient de corrélation

Variables	Inhealthrd	physicians	medgrads	hospitalbeds	fruitcons	alcoholcons	tobaccocons	unemployment
Inhealthrd	1							
physicians	-0,1325	1						
medgrads	-0,376	0,4486	1					
hospitalbeds	-0,3278	0,7516	0,6435	1				
fruitcons	0,7796	-0,4227	-0,6519	-0,6176	1			
alcoholcons	-0,0619	0,8393	0,1655	0,6311	-0,3416	1		
tobaccocons	-0,1848	0,5624	0,5344	0,4168	-0,1975	0,3391	1	
unemployment	0,2548	0,1704	0,2881	0,5792	-0,411	0,1884	-0,1664	1

Annexe 7 : Test de la normalité des résidus

Histogramme des résidus



Test de normalité Skewness-Kurtosis

H_0 : Les résidus sont de distribution normale

Variable	pr(Skewness)	Pr(Kurtosis)	adj χ^2	Prob > χ^2
Résidus	0,281	0,663	1,39	0,4993

Annexe 8 : Résultats des régressions

VARIABLES	pople	malele	femle
Inhealthrd	0,3547** [0,1412]	0,6041*** [0,1457]	0,0915 [0,1514]
canada	4,3964*** [0,3857]	5,2742*** [0,3979]	3,5083*** [0,4133]
Australia	4,1206*** [0,4900]	4,8810*** [0,5055]	3,2537*** [0,5252]
finlande	7,3867*** [0,6452]	7,9013*** [0,6656]	6,7394*** [0,6914]
france	5,0014*** [0,6381]	4,2019*** [0,6583]	5,6324*** [0,6838]
allemagne	5,4788*** [0,6779]	5,7996*** [0,6993]	5,0702*** [0,7265]
reptcheque	2,7996*** [0,7237]	3,5448*** [0,7465]	1,9417** [0,7755]
physicians	3,1041*** [0,2641]	3,7647*** [0,2724]	2,4523*** [0,2830]
medgrads	0,0802** [0,0329]	0,0798** [0,0339]	0,0932*** [0,0352]
hospitalbeds	-1,1048*** [0,1372]	-1,1980*** [0,1415]	-1,0008*** [0,1470]
fruit_vegetables	0,0094*** [0,0036]	0,0084** [0,0037]	0,0101*** [0,0039]
alcoholcons	-0,0162 [0,0642]	-0,0525 [0,0662]	0,0319 [0,0688]
tobaccocons	-0,0001 [0,0002]	-0,0003 [0,0002]	-0,0001 [0,0002]
unemployment	0,006 [0,0260]	0,0153 [0,0268]	-0,0109 [0,0279]
constant	67,5083*** [1,9426]	61,7284*** [2,0040]	73,4035*** [2,0818]
Wald χ^2	5317.7	6114.2	4220.71

Observations		84
Nombre de groupes		7
Observations par groupe	min	10
	moyenne	12
	max	13

Écart type entre parenthèses

*** p<0.01, ** p<0.05, * p<0.1

Annexe 9 : Estimations du modèle à retards échelonnés

VARIABLES	pople
Inhealthrd	0,2211 [0,2728]
Inhealthrd L1	0,1202 [0,2378]
Canada	4,0691*** [0,5292]
Australia	4,0691*** [0,5292]
finlande	7,1734*** [0,7445]
france	4,9449*** [0,7152]
allemagne	5,3528*** [0,7518]
reptcheque	2,5420*** [0,8277]
physicians	3,1365*** [0,2830]
medgrads	0,0737* [0,0384]
hospitalbeds	-1,0896*** [0,1525]
fruit_vegetables	0,0074* [0,0038]
alcoholcons	-0,0310 [0,0689]
tobaccocons	-0,0001 [0,0002]
unemployment	0,0039 [0,0308]
constant	68,1848*** [2,0903]
Wald χ^2	4562

Observations		77
Nombre de groupes		7
Observations par groupe	min	9
	moyenne	11
	max	12

Écart type entre parenthèses

*** p<0.01, ** p<0.05, * p<0.1