

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

L'OBÉSITÉ ET LES SALAIRES AU CANADA

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAITRISE EN ECONOMIQUE

PAR

SYLVIE THEVENIN

JANVIER 2010

# UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

## Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement n°8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Je tiens à remercier en premier lieu Madame Arianna Degan d'avoir accepté de diriger ce mémoire et pour sa patience, sa disponibilité, ses précieux conseils, ainsi que la confiance qu'elle m'a accordée.

Je remercie également les professeurs que j'ai eu le privilège de côtoyer au fil des années pour tous leurs enseignements et leurs encouragements. Je voudrais aussi souligner le professionnalisme et la prévenance de Madame Martine Boisselle-Lessard.

Pour terminer, je voudrais exprimer toute ma gratitude à ma famille qui m'a soutenue tout au long de cette démarche.

## TABLE DES MATIÈRES

RÉSUMÉ.....	iv
INTRODUCTION.....	1
CHAPITRE I.....	4
REVUE DE LITTÉRATURE.....	4
1.1. DÉFINITION DE L'OBÉSITÉ.....	4
1.2. LES DÉTERMINANTS DE L'OBÉSITÉ.....	5
1.3. LES EFFETS DE L'OBÉSITÉ SUR LA SANTÉ.....	6
1.4. LES EFFETS ÉCONOMIQUES DE L'OBÉSITÉ.....	8
CHAPITRE II.....	16
BASE DE DONNÉE.....	16
2.1. SOURCE DE DONNÉE.....	16
2.2. LES VARIABLES.....	17
2.3. CORRECTION DE L'IMC.....	20
2.4. LES VARIABLES INSTRUMENTALES DE L'IMC.....	21
CHAPITRE III.....	24
ANALYSE EMPIRIQUE.....	24
3.1. L'ÉQUATION DE SALAIRE.....	24
3.2. RÉSULTATS.....	25
CONCLUSION.....	30
BIBLIOGRAPHIE.....	32
ANNEXE.....	37
Tableau 1 : classification du risque pour la santé en fonction de l'indice de masse corporelle (IMC)...	37
Tableau 2 : revenu personnel généré par interpolation pour individu de 18 à 64 ans.....	37
Tableau 3 : statistiques descriptives des personnes âgées de 18 à 64 ans.....	38
Tableau 4 : personnes obèses auto-déclarées versus personnes obèses mesurées.....	38
Tableau 5 : coefficients de l'équation de correction de la taille (en mètres) auto-déclarée (aut).....	39
Tableau 6 : coefficients de l'équation de correction du poids (en kilogrammes) auto-déclaré (aut).....	39
Tableau 7 : corrélation entre l'IMC et instruments de l'équation de salaire.....	40
Tableau 8 : régression variable salaire avec IMC, première étape - échantillon hommes.....	40
Tableau 9 : régression variable salaire avec IMC, première étape - échantillon femmes.....	40
Tableau 10 : régression variable salaire avec IMC - échantillon hommes.....	41
Tableau 11 : régression variable salaire avec IMC - échantillon femmes.....	42

## RÉSUMÉ

L'augmentation soutenue de l'obésité au cours des dernières décennies est un phénomène partagé, à différents degrés, par tous les pays développés et en voie de développement. L'obésité constitue un problème majeur de santé qui a un impact notable sur l'économie.

D'une part, au niveau macroéconomique, les coûts d'un système de santé public sont affectés par l'obésité. D'autre part, au niveau microéconomique, le fait d'être obèse peut avoir un effet sur les résultats économiques individuels. À ce titre, ce document vise à étudier l'effet d'être obèse sur les salaires au Canada.

Plusieurs études existantes ont abordé cette question en utilisant des données sur la population des États-Unis et de certains pays d'Europe. Cependant, au moment de la rédaction de ce mémoire, il n'existe aucune étude sur l'obésité en lien avec les questions salariales au Canada.

Les premiers résultats issus de la présente étude révèlent qu'il existerait un impact négatif de l'obésité sur le salaire des canadiens et des canadiennes. D'un point de vue qualitatif, ces résultats sont en ligne avec diverses études américaines traitant de ce sujet. Toutefois, l'impact quantitatif apparaît comme étant plus important au Canada.

Mots clés : obésité, obèse, société, salaire, santé

## INTRODUCTION

Selon les données de l'Organisation Mondiale de la Santé (OMS), il y aurait 300 millions d'obèses dans le monde et un milliard d'adultes souffrant d'excès de poids. L'obésité est répandue tant dans les pays développés que dans les pays en voie de développement et elle coexisterait de plus en plus avec la sous nutrition (OMS, 2003).

L'obésité peut avoir des impacts économiques à la fois au niveau macroéconomique et au niveau microéconomique. Au niveau macroéconomique, des impacts considérables sur la dépense totale en frais de santé ont été identifiés. Étant donné que l'obésité accroît la probabilité d'être atteint d'un certain nombre de maladies, la dépense totale en frais de santé liée directement à l'obésité est destinée à augmenter si l'obésité poursuit sa progression. Cette situation est alarmante parce que l'augmentation va créer ou exacerber des tensions sur le budget du secteur de la santé ainsi que sur le budget d'autres fonctions de l'état.

L'évolution des coûts de l'obésité sur le système public doit inclure non seulement les coûts directs sur le système sanitaire mais aussi les coûts indirects sur les programmes sociaux causés par le décès prématuré et l'incapacité de travailler. Birmingham et al (1999) estiment que le coût total direct de l'obésité sur le système de santé du Canada a dépassé 1,8 milliards de dollars, soit 2,4 % du total des dépenses dans le domaine de la santé (en dollars de 1997). Katzmarzyk et Janssen (2004) ont estimé (en dollars de 2001), que l'obésité avait coûté aux Canadiens 4,3 milliards de dollars, dont 1,6 milliards en coûts directs et 2,7 milliards en coûts indirects.

Pour ce qui concerne les effets microéconomiques, notons que l'obésité a des effets sur la santé de l'individu et potentiellement sur sa performance sur le marché du

travail. Il est documenté que l'obésité est liée à plusieurs maladies. Garfinkel (1985), rapporte que le ratio de mortalité chez les obèses, comparé aux individus de poids normal, pour tous les types de cancer, pour le diabète, et les maladies coronariennes était supérieur à un. L'obésité est donc associée à plusieurs types de maladies d'une part, d'autre part, elle peut devenir invalidante lorsqu'elle dépasse un certain niveau (obésité morbide) et elle est liée au décès prématuré des individus. Suite à la croissance du phénomène de l'obésité, plusieurs études économiques ont versé leur attention sur l'analyse des liens entre obésité et résultats économiques individuels. Une relation entre l'obésité et le choix d'accumulation de capital humain est probable. Il se pourrait aussi que des facteurs non-observés conduisent les individus à devenir obèse et à adopter certains comportements sur le marché du travail ou encore qu'il existerait une discrimination salariale ou à l'embauche envers les personnes obèses. En complément des études existantes sur le sujet, le présent document vise à quantifier l'impact de l'obésité sur les salaires au Canada.

Il y a un nombre important d'études sur le lien entre l'obésité et des résultats économiques aux États-Unis et en Europe. Toutefois, pour le Canada, nous avons recensé peu d'écrits économiques sur le sujet. MacMinn, McIntosh et Yung (2007) ont comparé l'obésité à d'autres comportements potentiellement dommageables pour la santé (dont le tabagisme). Par ailleurs, Auld et Powell (2005) ont fait une comparaison Canada/États-Unis sur les caractéristiques des personnes obèses. Dans la lignée des études de Cawley, qui a réalisé plusieurs écrits sur la relation entre obésité et salaires aux États-Unis, le présent document a pour objectif de vérifier s'il existe un lien entre l'obésité et les salaires au Canada et d'en évaluer l'impact.

En adressant ces thèmes, il faudra faire face à un problème éventuel d'endogénéité dû, s'il y a lieu, à des erreurs de mesure, aux variables omises et / ou à la causalité inversée. Un premier enjeu survient, dans les études sur l'obésité, lorsque la taille et le poids des individus y sont auto-déclarés. Comme autre enjeu présent dans ce type d'étude, il existe probablement des variables non-observées qui affectent à la fois le salaire et l'obésité. Également, il se peut que le salaire, à travers l'effet sur le type d'alimentation et le niveau de stress, affecte l'obésité. Pour traiter du premier enjeu, une méthode développée par Cawley et Burkhauser (2006) sera utilisée pour corriger la taille et le poids auto-déclarés.

Pour ce qui est des deux enjeux subséquents, la méthodologie des variables instrumentales sera privilégiée.

Deux types d'instruments servent de base à la présente analyse : un premier type d'instrument basé sur les problèmes de santé des individus et l'autre type plutôt relié aux habitudes de vie qui affectent la dépense d'énergie et le métabolisme. L'utilisation de ce deuxième type d'instrument représente une nouveauté dans la littérature empirique sur l'obésité et incarne une des contributions de ce mémoire. Trois spécifications ont été considérées : initialement en utilisant chacune des deux variables instrumentales séparément et ensuite en prenant les deux variables de manière combinée. Les différentes spécifications considérées mènent à la conclusion que l'IMC a un impact négatif statistiquement significatif sur la détermination du salaire, tant pour les hommes que pour les femmes et les tests effectués confirment la validité des instruments utilisés.

Ce mémoire est organisé de la façon suivante. Le premier chapitre contient une revue de la littérature. Le deuxième chapitre est consacré à la présentation des données utilisées pour l'analyse empirique. Le troisième chapitre présente l'équation du salaire de base ainsi que les résultats des différentes estimations réalisées. La quatrième section résume et conclut.

## CHAPITRE I

### REVUE DE LITTÉRATURE

Cette section, donne la définition de l'obésité suivie d'un portrait de la littérature liée au phénomène de l'obésité.

#### *1.1. Définition de l'obésité*

La définition de l'obésité la plus utilisée est basée sur l'indice de masse corporelle (IMC). Il s'agit du ratio entre le poids en kilogrammes et le carré de la taille . Une personne est définie comme étant obèse si son IMC est supérieur à 30. On dit qu'un individu est en excès de poids, si son IMC est entre 25 et 29,9, et qu'il a un poids normal si son IMC est entre 18,5 et 24,9. Les gens ayant un poids insuffisant ont un IMC inférieur à 18,5 (voir Tableau 1).

Selon Santé Canada, l'utilisation de l'IMC n'est pas une mesure idéale pour évaluer les effets de l'obésité sur la santé, car elle ne tient pas compte de la composition du corps (graisse, muscles, ossature, etc.). Conséquemment, chez les sujets très musclés, les jeunes qui n'ont pas encore atteint leur pleine taille, les personnes âgées de plus de 65 ans, l'interprétation de l'IMC comporte des limites. Cependant, l'IMC reste la mesure d'obésité la plus utilisée dans les études en santé et en économie sur les questions d'obésité.

## 1.2. Les déterminants de l'obésité

Divers facteurs peuvent expliquer l'augmentation du taux d'obésité mentionnée en introduction. D'ailleurs, plusieurs s'entendent pour dire qu'il existe maintes causes à cette problématique. L'augmentation de la consommation d'aliments à haute teneur en calorie, contenant de fortes concentrations de sucre et de gras transgéniques est souvent évoquée. On remarque également la réduction de l'activité physique due, entre autre, aux progrès technologiques (mécanisation, automatisation, etc.) et aux loisirs plus sédentaires, tels que les jeux vidéo, l'Internet et la télévision. Certains expliquent l'accroissement du phénomène de l'obésité par la diminution du prix relatif de la nourriture (occasionnée par la mécanisation de l'agriculture) et par la hausse relative du prix de l'exercice physique. La multiplication des chaînes de restauration rapide ne serait pas étrangère à l'augmentation du nombre d'obèses. Des chercheurs ont aussi démontré que des facteurs génétiques et environnementaux jouent un rôle important sur l'obésité.

Lakdawalla et al. (2002) affirment que 40% de l'accroissement récent du poids des américains serait dû aux innovations dans le domaine de l'agriculture qui ont fait chuter le prix des aliments et que 60% de l'accroissement proviendrait du déclin de l'activité physique à cause des changements technologiques dans nos maisons et sur les chaînes de production.

Anderson et al. (2003) avancent que l'obésité serait liée en partie à la participation accrue des femmes à la vie active. L'obésité a commencé à augmenter fortement à partir des années 70 en Amérique du Nord et à cette même période, les femmes ont amorcé une entrée en force sur le marché du travail. Selon ces auteurs, ce ne serait pas tant le fait que les femmes soient actives sur le marché du travail que le fait qu'elles travaillent à temps plein : elles auraient moins de temps pour « cuisiner des repas santé » à leur famille.

Christakis et Fowler (2007), ont observé un groupe de 12067 personnes ayant un réseau social inter relié. Ils ont utilisé un modèle statistique longitudinal pour voir si le gain de poids d'un individu est lié au gain de poids d'autres individus faisant partie de son réseau social (amis, voisins, conjoint, frères et sœurs). Ils parviennent à établir que si le

conjoint d'une personne devient obèse, la probabilité de devenir obèse pour cette personne augmente de 37%. Cela indique qu'il y a d'importants déterminants sociaux et environnementaux à l'obésité.

Pour Eid et al. (2006), l'obésité n'a pas évolué au même rythme et n'a pas atteint le même niveau, selon le lieu géographique de résidence des individus. Dans le cas des États-Unis, l'obésité a augmenté de 102% en Géorgie alors qu'au Delaware l'augmentation a été de 11% entre 1991 et 1998. Pour expliquer de telles disparités les auteurs ont étudié ce que l'OMS appelle la *théorie de l'étalement urbain*, selon laquelle plus les immeubles sont éloignés les uns des autres dans une ville, plus les individus utilisent leurs véhicules et moins ils marchent.

Des chercheurs, tels que Vuille et Melbin (1979) ont aussi prouvé que les facteurs génétiques et environnementaux jouent un rôle sur l'obésité. Selon eux, il y a des facteurs explicatifs qui diffèrent selon le sexe. Chez les filles, l'hérédité et l'inactivité physiques seraient les facteurs prédominants alors que chez les garçons ce serait l'appétit et l'environnement.

### *1.3. Les effets de l'obésité sur la santé*

L'obésité a toutes sortes d'effets sur la santé, comme en témoigne la vaste littérature sur le sujet. Elle diminue l'espérance de vie, elle est associée au développement de plusieurs types de maladies telles que l'asthme, le diabète, l'hypertension, le cancer et la dépression pour n'en nommer que quelques-unes.

Hubert et al (1983) concluent que chez les femmes l'obésité est la variable la plus importante, après l'âge et la pression sanguine, dans la prévision des maladies cardiovasculaires. Quant à Garfinkel (1985), il rapporte que le ratio de mortalité, comparé aux individus de poids normal, pour tous les types de cancer était de 1,33 chez les hommes obèses et de 1,55 chez les femmes obèses. Il ajoute que pour les maladies coronariennes ce ratio est de 1,95 chez les hommes obèses et de 2,07 chez les femmes obèses. Pour le diabète, le ratio est 5.19 chez les hommes et 7.90 chez les femmes. Les hommes souffrant

d'embonpoint ont des taux de mortalité plus élevés pour le cancer colorectal et de la prostate. Les femmes souffrant d'embonpoint ont, quant à elles, des taux plus élevés de cancer du sein, de cancer des ovaires et d'autres types de cancer.

Selon l'American Society for Reproductive Medicine (2001), le fait d'être obèse ou de souffrir d'embonpoint diminue la fertilité des femmes. Toutefois, le mécanisme reliant l'obésité à la fertilité est pour le moment inconnu. Galtier-Dereure et al. (2000) remarquent que l'obésité cause de nombreuses complications chez les femmes enceintes. La grossesse d'une femme obèse est considérée comme étant une grossesse à haut risque. L'obésité augmente le taux de morbidité fœtal et maternel. Les femmes enceintes ayant un problème d'embonpoint sont plus à risque d'avoir des problèmes de diabète gestationnel et d'hypertension que les femmes de poids normal. Les femmes enceintes souffrant d'embonpoint sont aussi plus susceptibles d'accoucher par césarienne, ce qui augmente l'incidence de complications postopératoires. Il y a des impacts sur l'enfant aussi, puisque le poids de la mère a une incidence sur les problèmes du tube neural de l'enfant. L'obésité maternelle est, en outre, un facteur de risque pour les maladies congénitales.

L'obésité peut aussi diminuer la qualité du sperme chez les hommes. Une recherche présentée lors d'une rencontre de l'European Society of Human Reproduction and Embryology (ESHRE) à Barcelone en 2008, démontre que les hommes obèses ont des niveaux inférieurs de sperme normal. Des échantillons de sperme de plus de 5 000 hommes en Écosse ont été analysés. Les hommes ayant un poids normal (IMC de 20 à 25) avaient un niveau de sperme normal supérieur à ceux ayant un problème de surplus de poids. Les hommes faisant de l'embonpoint avaient 60 % plus de risques d'avoir un volume de sperme bas, selon cette étude ; ils étaient aussi 40 % plus susceptibles d'avoir un sperme anormal. Les chercheurs ont noté par ailleurs que les hommes ayant un IMC faible, avaient le même genre de problème que les obèses au niveau de la fertilité.

#### *1.4. Les effets économiques de l'obésité*

Un premier effet économique de l'obésité serait l'effet sur les coûts de santé publique. Birmingham et al (1999) ont calculé les coûts directs de l'obésité et des comorbidités pour le système de santé du Canada. Les auteurs ont utilisé une définition différente de l'obésité (IMC = 27) et ils ont travaillé à partir de données auto-déclarées issues de l'Enquête Nationale sur la Santé de la Population (ENSP), ainsi que sur des morbidités provoquées par un nombre limité de maladies (maladies pour lesquelles le lien avec l'obésité a été validé). Selon leur estimation, le coût total de l'obésité pour le système de santé du Canada a dépassé 1,8 milliards de dollars, en dollars de 1997, soit 2,4 % du total des dépenses de santé. Les trois facteurs qui ont contribué le plus au coût total étaient l'hypertension (656,6 millions de dollars), le diabète de type 2 (423,2 millions) et la coronaropathie (346,0 millions). Comme autres résultats, les auteurs ont révélé qu'au Canada 50,7 % des coûts du diabète de type 2 étaient attribuables à l'obésité, tout comme 31,6 % de ceux de l'hypertension, pour ne nommer que ces deux maladies. Cette étude n'a pas pris en compte les coûts indirects associés à la perte possible de productivité, à l'absentéisme et aux limitations physiques. L'impact économique aurait été, auquel cas, encore plus élevé.

Plus récemment, Katzmarzyk et Janssen (2004) ont fait un calcul similaire. Les auteurs ont utilisé la définition standard de l'obésité (IMC = 30). Ils ont estimé, en dollars de 2001, que l'obésité avait coûté aux Canadiens 4,3 milliards de dollars dont 1,6 milliards en coûts directs et 2,7 milliards en coûts indirects. Les coûts directs comprennent le coût des soins hospitaliers, les médicaments, les soins de médecins, les soins dans d'autres institutions, et les dépenses directes additionnelles en soins de santé. Les coûts indirects ont été mesurés en fonction de la valeur des années de vie perdues en raison d'un décès prématuré et de la valeur des jours d'activité perdus en raison d'une incapacité à court ou à long terme.

Un autre impact économique de l'obésité se situe au niveau du marché du travail. Cawley et Danziger (2004) établissent un lien entre le nombre de mois passés à l'assistance emploi et le poids. Une personne obèse bénéficiant d'assistance emploi enregistre un délai

plus long pour se trouver un emploi qu'une personne de poids normal. Par ailleurs, Cawley et Burkhauser (2006) démontrent qu'il existe une relation entre le fait d'être obèse et l'âge de la retraite. Les gens obèses ont tendance à prendre leur retraite à un plus jeune âge car ils souffrent d'un plus grand nombre de problèmes de santé que les gens de poids normal.

L'obésité affecte possiblement les salaires par soit une discrimination effective, soit une discrimination statistique. Pour ce qui concerne la discrimination effective, l'on recense une vaste littérature qui analyse la relation entre l'apparence physique et la performance sur le marché du travail. Dans ces études, on cherche à déterminer l'existence d'une prime à la beauté (*beauty premium*), et plusieurs d'entre-elles dont celle de Hamermesh et Biddle (1994), concluent que cette prime existe et que les personnes étant perçues comme ayant une plus belle apparence sont mieux rémunérées que celles étant perçues comme ayant une moins belle apparence. L'obésité aurait un impact à l'embauche mais aussi lors de la progression de la carrière (promotions, etc.). Cette relation est toutefois difficile à prouver et à quantifier. Des poursuites judiciaires contre des employeurs ont eu lieu, invoquant la discrimination à l'endroit d'une personne obèse, comme le démontre la cause *Cook vs. Rhode Island Department of Mental Health, 1993 aux États-Unis*, ce qui suggère l'existence de discrimination à l'endroit des personnes obèses.

La discrimination ne se ferait pas sentir uniquement au niveau du salaire. Rooth (2007) a fait des expériences qui montrent qu'il y a un traitement inégal des obèses à l'embauche sur le marché du travail. Les expériences de Rooth se déroulent en Suède et la manière de procéder est la suivante : des curriculum vitæ (CV) fictifs sont envoyés pour de vrais emplois et on prend soin d'inclure une photo du candidat. En Europe, c'est la norme d'inclure une photographie en postulant pour un emploi. Les photos qui sont envoyées sont choisies au hasard et certaines sont celles d'une personne avec un poids santé et d'autres sont celles de personnes obèses. En tout, 989 photos ont été envoyées.

Les personnes obèses ont reçu 20% moins d'appels que les personnes avec un poids santé. L'expérience a montré que la discrimination est égale pour les hommes et

pour les femmes et elle serait systématique dans certaines professions, particulièrement dans les professions qui demandent un contact direct avec le public.

L'auteur a aussi fait des entrevues avec des gestionnaires suédois pour connaître leurs impressions sur les personnes obèses. Il est ressorti que 9 employeurs sur 10 croient que les personnes obèses sont moins productives que le reste de la population. Un des problèmes rencontrés lors de ces entrevues est le manque de sincérité des employeurs quant à leurs sentiments par rapport aux obèses. Même ceux qui disent être indifférent à l'apparence physique, adoptent des comportements contradictoires. L'auteur a pu évaluer le degré de véracité des réponses d'entrevue en envoyant les CV fictifs. Cette façon de faire donne une information supplémentaire à ce qui se fait au niveau des sondages et des régressions économétriques. Selon cet auteur, il semble que les personnes obèses soient victimes de discrimination.

Becker (1957) soutient que certaines entreprises sont prêtes à payer le prix de la discrimination (introduction d'une variable de préférence pour la discrimination), mais seulement jusqu'à un certain point puisque le contexte de concurrence réduit, selon l'auteur, le bénéfice tiré de la discrimination. Il faut noter que dans cet ouvrage Becker considère seulement la discrimination à l'embauche mais pas celle par rapport au salaire.

Il y a eu un second mouvement à la théorie de la discrimination. Le second courant d'analyse de la discrimination statistique développé principalement par Arrow (1972), Phelps (1972), Stiglitz (1973) et Akerlof (1985) soutient que la non-observabilité de certaines compétences chez les aspirants employés peut être une explication de la persistance de la discrimination. Les employeurs font face à un problème d'information et ils utilisent les connaissances qu'ils ont de la productivité moyenne d'un groupe donné, pour juger de la productivité individuelle d'un candidat.

Phelps croit que l'employeur est gagnant lorsqu'il trie sa main-d'œuvre pour maximiser ses profits. Stiglitz, de son côté, avance que le manque d'information sur les candidats force le recruteur à prendre un risque et que pour réduire son incertitude, le recruteur va utiliser les qualifications moyennes du groupe auquel appartient le candidat.

Becker est aussi l'auteur de la théorie du capital humain. Le capital humain est défini comme étant des connaissances, du talent et de l'expérience acquise par un individu. Plus spécifiquement, il peut s'agir de la formation en milieu de travail (générale et/ou spécifique), de la formation scolaire, de connaissances, et de l'état de santé. Selon cette théorie, le salaire est fonction du capital humain accumulé, donc des efforts que l'individu a faits pour s'améliorer, tant du point de vue physique qu'intellectuel.

En faisant le lien avec cette théorie et celle de la discrimination statistique, si les obèses consacrent moins de temps au travail salarié que le reste de la population (s'ils sont moins souvent embauchés par les entreprises que les personnes non-obèses), ils vont accumuler moins d'expérience et leur productivité n'augmentera pas autant que celle du reste de la population. Par ailleurs, si l'obésité est une cause d'absentéisme, alors elle ferait perdre du capital humain à ceux qui en souffriraient et selon cette théorie cela expliquerait les écarts salariaux. Il semble que les obèses adoptent une attitude différente face à la scolarité. Selon la théorie de Becker, les gens investissent en formation pour bénéficier plus tard de certains avantages et cela explique entre autre pourquoi plus un individu vieillit, moins il investit dans son capital humain (il reste moins de temps pour bénéficier des avantages de l'investissement, lorsqu'on avance en âge) ; il faut souligner que les personnes obèses ont généralement une espérance de vie plus courte que le reste de la population donc suivant cette théorie, elles ont éventuellement un comportement différent du reste de la population par rapport à l'amélioration de leur capital humain. Gortmaker et al. (1993) ont obtenu les résultats dans ce sens. Leur étude compte 370 sujets qui en 1981 souffraient d'embonpoint. Sept ans plus tard, les sujets féminins qui souffraient d'embonpoint avaient complété 0.3 années de moins de scolarité, 20% de moins étaient susceptibles de se marier et 10% de plus vivaient dans un foyer à faible revenu. Bien qu'il soit difficile de comprendre le lien causal, l'embonpoint à l'adolescence a un impact sur les conditions économiques et sociales à l'âge adulte, du moins pour les femmes.

Plusieurs économistes se sont intéressés au lien entre l'obésité et la performance des individus sur le marché du travail, notamment le salaire et la probabilité d'avoir un emploi. Comme nous, ils ont dû faire face inévitablement à la difficulté d'isoler l'effet causal de

l'obésité sur la performance sur le marché du travail. En d'autres mots, il y a plusieurs raisons, présentées précédemment, de croire que l'obésité serait une variable endogène. Ce n'est peut-être pas le fait d'être obèse qui diminue le salaire et peut-être que c'est le salaire qui a un impact sur l'obésité. Les personnes vivant dans la pauvreté s'alimenteraient peut-être moins bien et choisiraient alors des aliments qui sont moins bons pour la santé. Dans la même veine, il pourrait y avoir une relation entre la pauvreté et des problèmes de santé mentale ou de nature émotionnelle qui mèneraient vers des comportements compulsifs tels que la suralimentation et qui entraîneraient l'obésité. Comme autre hypothèse, il existe une possibilité que certains facteurs non-observés induisent les individus à peu investir à la fois dans leur capital humain, dans leur aspect physique et dans leur santé.

Certains ont opté de travailler avec des données sur des sujets « jumeaux » afin de contrer le problème d'endogénéité de l'obésité comme par exemple Berhman et Rosenzweig (2001). Mais pour faire face au problème d'endogénéité de l'obésité, la stratégie généralement utilisée dans la littérature est celle des variables instrumentales. Les instruments utilisés dans la littérature peuvent être classés en trois grandes catégories : la première catégorie d'instrument est constituée de la valeur retardée de l'IMC. Dans ce cas, il est présumé que l'obésité de la période passée n'aura pas de lien avec le salaire actuel. À ce propos, les études utilisent généralement l'IMC retardée de 7 périodes.

La deuxième catégorie est composée de l'IMC d'individus qui ont des gènes très proches, comme des parents, des enfants biologiques ou des sœurs ou frères du sujet observé.

Dans la troisième catégorie, les variables instrumentales sont celles qui ne font pas partie des deux premières catégories utilisées dans la littérature. On retrouve dans cette catégorie des variables telles que, par exemple, certains types de problème de santé, des indicateurs d'estime de soi, etc. Il est toutefois difficile de trouver de bons instruments. D'un côté, un bon instrument doit être corrélé avec la variable instrumentée, d'autre part, il ne doit pas avoir un impact direct sur la variable dépendante. Autrement dit, la variable instrumentale doit avoir un impact sur la variable dépendante uniquement à travers ce qui la lie à la variable instrumentée.

John Cawley s'est intéressé à la question de la relation entre le salaire, la performance sur le marché du travail et l'obésité, dans plusieurs études. Entre autres, Cawley (2000) a fait une étude uniquement sur les femmes et la performance sur le marché du travail selon le poids en utilisant les données du National Longitudinal Survey of Youth (NLSY) et en appliquant la méthode des variables instrumentales. Il conclue que le poids diminue le salaire des femmes mais pas la probabilité d'occuper un emploi.

Plus tard, Cawley (2004) étudiera le même sujet en prenant en compte tant les femmes que les hommes. Il démontre qu'un IMC élevé affecte le salaire des femmes mais de façon différente selon le groupe ethnique auquel elles appartiennent. L'impact négatif de l'obésité sur les salaires était plus élevé chez les femmes de race blanche. Chez les hommes les résultats étaient plus mitigés. Cela s'explique, selon l'auteur, par le fait que les hommes obèses s'orientent vers des emplois pour lesquels l'obésité n'est pas un obstacle et que chez les femmes ce genre d'emplois serait plus rare.

Pagan et Davila (1997) ont aussi travaillé avec les données du NLSY. Ils ont utilisé l'estime de soi et le niveau de pauvreté familial comme variables instrumentales. Il faut noter que les instruments choisis par Pagan et Davila sont pour certains probablement corrélés avec le terme d'erreur dans l'équation de salaire et qu'ils ne sont peut-être pas adéquats en ce sens. Le salaire et le niveau de pauvreté familiale sont étroitement reliés. L'estime de soi peut être reliée à la réussite sociale et indirectement au salaire. Néanmoins, de manière semblable à Cawley (2004), ils concluent que les femmes obèses paient une pénalité sur leur salaire à cause de leur poids mais que ce n'est pas le cas chez les hommes. Ils amènent comme explication qu'il y aurait plus de mobilité professionnelle chez les hommes et que obèses de sexe masculin auraient plus de facilité à se placer dans des emplois où cette pénalité n'existe pas. Les auteurs ajoutent que la pénalité payée par les femmes s'expliquerait par la discrimination.

Conley et Glauber (2007) étudient la question de manière plus large qu'uniquement en terme de salaire : ils observent aussi la probabilité de se marier. Ils utilisent comme instrument l'IMC retardée de l'individu. Ils se différencient par le fait qu'ils utilisent

comme instrument la variable d'IMC retardée de 15 ans (au lieu du traditionnel 7 ans) et leur échantillon est beaucoup plus ancien que celui d'autres études. Néanmoins, ils arrivent à des conclusions similaires : l'impact négatif de l'obésité sur le salaire est plus fort pour les femmes que pour les hommes. De plus, le fait d'être obèse pénaliserait davantage les femmes tant au niveau du salaire individuel qu'au niveau du salaire familial, selon les auteurs. Pour ce qui est de la relation entre l'obésité et la probabilité d'être marié, les conclusions sont les suivantes : la probabilité de se marier est plus faible pour les personnes obèses que pour les personnes de poids normal.

Behrman et Rosenzweig (2001) en utilisant données sur des femmes jumelles du Minnesota arrivent à la conclusion que s'ils contrôlent les variables d'éducation et d'expérience de travail, l'IMC n'était pas une variable significative à la détermination du salaire. Cette méthode permet d'exercer un contrôle sur les facteurs génétiques non-observés, tels que l'habileté, qui affecteraient les salaires. Certains critiques diront de cette étude que l'échantillon était trop petit et ils soulèveront le manque de robustesse des résultats.

Dans les études faites du côté européen les résultats sont moins ambigus : le salaire et le poids sont nettement corrélés de manière négative peu importe le sexe. En utilisant les données de l'*European Community Household Panel*, Brunello et d'Hombres (2005) étudient le rapport entre l'obésité et le salaire horaire. Brunello et D'hombres utilisent des données statistiques de 9 pays européens. Pour instrumenter l'IMC, ils retiennent l'IMC d'un membre de la famille biologique de l'individu observé. Ils ont séparé les échantillons en deux sections : les pays de l'*olive belt* (Espagne, Grèce, Italie, Portugal) et les pays de la *beer belt* (Autriche, Danemark, Irlande, Belgique, Finlande). Ces auteurs concluent que l'effet de l'obésité sur les salaires horaires est négatif tant chez les hommes que chez les femmes et que dans les pays de l'*olive belt* l'effet négatif sur le salaire était plus marqué que dans l'autre groupe de pays.

Une autre étude plus récente sur des sujets européens a été menée par Garcia et Quitana-Domeque (2008). Ces auteurs passent en revue la relation entre revenu familial, IMC, statut d'obésité et catégorie d'IMC, pour les hommes et les femmes de neufs pays

européens. Ils prennent les données de l'European Community Household Panel. Garcia et Quitana-Domeque constatent une relation négative entre revenu familial et obésité. Ils croient en fait que la relation capte la pénalité salariale imposée aux femmes obèses. Malheureusement, ils ne parviennent pas à des résultats statistiquement valables pour les hommes.

## CHAPITRE II

### BASE DE DONNÉE

#### *2.1. Source de donnée*

L'analyse empirique de ce mémoire repose sur les données de l'Enquête sur la Santé dans les Collectivités Canadiennes (ESCC) de 2004 cycle 2.2 collectées par Statistique Canada. Cette enquête est de type coupe transversale et contient des informations reliées à l'état de santé, aux soins et aux déterminants de la santé des canadiens ainsi que certaines variables démographiques et économiques. En outre, le cycle 2.2 contient un sous-ensemble pour lequel la taille et le poids sont à la fois mesurés et déclarés. Pour le reste de l'échantillon, la taille et le poids sont auto-déclarés. Les informations du cycle 2.2 ont été collectées entre janvier 2004 et janvier 2005 dans toutes les provinces canadiennes pour des personnes de tout âge. L'échantillon exclut les individus vivant dans les trois territoires, les individus en institution, les individus qui font partie des forces canadiennes à plein temps, les personnes vivants dans les réserves indiennes et sur les terres de la couronne ainsi que ceux vivant dans les régions éloignées.

Plusieurs études sur l'obésité utilisent des données longitudinales, comme par exemple Cawley (2000) et Pagan et Davila (1997) qui utilisent le National Longitudinal Survey of Youth (NLSY) aux États-Unis, ainsi que, plus récemment, García Villar et Quintana-Domeque (2008), qui utilisent l'European Community Household Panel (ECHP). Ce type de données facilite le contrôle des facteurs individuels non-observés constants dans le temps. De plus, dans le cas spécifique de l'obésité, l'utilisation de données longitudinales permet d'utiliser l'IMC retardée comme variable instrumentale de l'obésité, tel que discuté dans la revue de la littérature.

Il existe des données canadiennes longitudinales sur la santé, telles que celles issues de l'Enquête nationale sur la santé de la population (ENSP) de Statistique Canada. L'ENSP contient plusieurs variables sur la santé, l'activité physique, et des variables économiques. Toutefois, la mesure du poids et de la taille y est auto-déclarée.

Comme première recherche sur le sujet pour le Canada et afin de pouvoir utiliser des variables mesurées pour le poids et la taille des individus, nous avons choisi d'utiliser l'ESCC dans cette étude. Dans la majorité des études sur le sujet, des valeurs auto-codées sont utilisées. Cependant, lorsqu'il y a un écart non négligeable entre le poids réel et celui déclaré par les individus (comme c'est le cas avec les données de l'ESCC), l'utilisation de variables auto-codées peut engendrer des résultats biaisés.

Le fichier de données de l'ESCC contient 35107 observations en tout. L'analyse va se limiter aux sujets âgés entre 18 ans et 64 ans pour lesquels il y a une valeur d'IMC. Les gens qui ont moins de 18 ans et ceux de plus de 65 ans sont retirés de l'échantillon. En effet, les barèmes qui définissent la relation entre IMC et le fait d'être obèse sont affectés par l'âge. Par exemple, il n'y a pas de consensus dans la littérature sur la bonne définition d'obésité chez les enfants. Par ailleurs, les personnes âgées subissent généralement de leur côté un changement au niveau musculaire et osseux en vieillissant qui a pour conséquence de plutôt diminuer leur IMC. L'autre raison pertinente qui justifie le fait d'avoir éliminé deux catégories d'âge est que l'étude se concentre sur les salaires (d'où le choix d'observer les gens qui constituent la majeure partie de la population active). En éliminant les individus âgés de moins de 18 ans et ceux de 65 ans et plus, il y a une perte de 20653 observations, soit 58% de l'échantillon, et il reste donc 14 454 observations. D'ailleurs, pour des raisons qui seront expliquées à la section 2.2, dans l'analyse sur les salaires seuls les gens qui travaillent à temps plein seront pris en compte. Cela réduit l'échantillon à 9806 observations. En outre, les observations qui ont des valeurs manquantes pour les variables utilisées dans les diverses équations, sont aussi éliminées ; ce qui donne un échantillon final de 7903 observations (4282 hommes et 3621 femmes).

## *2.2. Les variables*

Les deux variables d'intérêt dans ce mémoire, soient l'IMC et le salaire seront décrites en premier dans la présente section. Nous poursuivrons la description des variables avec les indicateurs socio-démographiques de l'échantillon. L'indice de masse corporelle (IMC) a été utilisé en tant que variable de mesure de l'obésité, comme il est répandu de le faire dans la littérature. L'IMC est construite à partir des variables de poids (en kilogrammes) et de taille (en mètres). Tel que mentionné précédemment, l'IMC a été corrigé pour une partie des sujets qui avaient une valeur auto-codée d'IMC mais qui n'avaient pas de valeur mesurée d'IMC. La méthodologie retenue pour corriger l'IMC sera expliquée dans la prochaine section.

La variable *SALAIRE*, qui représente le logarithme du salaire horaire, sera la variable dépendante de notre étude. Généralement, les bases de données d'enquête ne contiennent pas de variable de salaire horaire. Typiquement, pour construire la variable de salaire horaire on divise le revenu annuel provenant du travail par le nombre d'heures travaillées dans l'année. Malheureusement, la seule information disponible était le revenu annuel personnel de toutes sources. À cause de toutes ces limites, les observations utilisées dans les régressions furent celles pour lesquelles les individus travaillent à plein temps. En outre, ce revenu individuel était disponible seulement par catégories (voir tableau 2). Il a donc été nécessaire de procéder par interpolation pour assigner un revenu à chaque individu. Pour le calcul des revenus personnels moyens, les valeurs médianes pour lesquels le revenu personnel était codé, ont été utilisées. Pour le dernier groupe, celui composé par les individus ayant un revenu annuel supérieur à 80 000 dollars, la valeur plancher de la catégorie a été utilisée, tel que suggéré par Wooldridge (2002, p509).

De cette manière il y a probablement une légère sous représentation du revenu des individus de notre échantillon. Pour s'assurer que la procédure mène à des salaires horaires réalistes le salaire moyen de l'échantillon a été comparé avec le salaire horaire des canadiens estimé par Statistique Canada. Le salaire moyen de l'échantillon de gens qui ont travaillé à plein temps est de 17,00 dollars. Selon Statistique Canada, le salaire horaire moyen en 2004 était de 17,73 dollars au Canada ; l'écart est de moins d'un dollar avec la valeur « plancher » de 80 000 dollars ce qui donne confiance en la procédure utilisée et sur

la qualité de l'information disponible. Comme alternative, nous aurions aussi pu forcer la moyenne du salaire horaire à 17,73 dollars en mettant la valeur plafond correspondante.

La base de données contient de plus, des variables socio-démographiques usuelles telles que l'âge, la province de résidence, la scolarité, l'état civil et la race. Puisque l'âge est codé en catégorie, une variable *AGE* qui contient l'âge moyen de la catégorie dont fait partie le répondant a été générée. Une alternative aurait été d'utiliser des variables dichotomiques pour les différentes catégories d'âge. Des variables dichotomiques (*NÉCOSSE*, *TNEUVE*, *NBRUNS*, *QUÉBEC*, *ONTARIO*, *MANITOBA*, *SASK*, *ALBERTA*, *COLOMBIEB*), ont aussi été créées pour identifier la province de résidence des sujets selon qu'ils demeurent respectivement en Nouvelle-Écosse, Terre-Neuve, Nouveau-Brunswick, Québec, Ontario, Manitoba, Saskatchewan, Alberta et Colombie-Britannique. Étant donné l'hétérogénéité d'environnement économique et social entre les provinces canadiennes, il s'avérait judicieux d'intégrer la dimension géographique dans les modèles proposés.

La variable dichotomique *RACE\_BLANCHE* prend la valeur 1 si l'individu est de race blanche 0 autrement. De plus, puisqu'il existe des raisons de croire que le bilinguisme au Canada a un effet sur le salaire ( voir Vaillancourt (1980) pour le Québec), nous allons utiliser la variable *BILINGUE*, qui prend la valeur 1 si l'individu maîtrise à la fois la langue française et la langue anglaise et qui prend la valeur 0 si l'individu ne maîtrise qu'une seule des deux langues.

Pour la scolarité, trois variables dichotomiques (*SECONDAIRE*, *POSTSECNON*, *POSTSECOND*) ont été créées ; selon si l'individu a complété des études secondaires, fait des études post-secondaires non-complétées ou s'il a complété des études post-secondaires. Pour ce qui est de l'état civil de l'individu, il est représenté par la variable *MARIE* qui prend la valeur 1 si le sujet est marié et 0 autrement.

Pour résumer, l'échantillon utilisé dans l'analyse empirique inclut les individus âgés de 18 à 64 ans qui travaillent à temps plein pour lesquels il n'y a aucune observation manquante pour toutes les variables utilisées dans les équations. Cela donne 7903 observations dont 4282 hommes et 3621 femmes. De ces valeurs, il y a 5312 observations

(67%) pour lesquelles l'IMC est mesuré et 2591 observations (33%) pour lesquelles l'IMC est corrigé à partir de chiffres auto-déclarés. Ce sous-ensemble, dont les statistiques descriptives sont résumées au tableau 3, sera utilisé pour l'équation du salaire.

### *2.3. Correction de l'IMC*

La base de données retenue a l'avantage d'offrir l'IMC mesurée des individus. Malheureusement, l'IMC n'a pas pu être mesurée pour toutes les observations. Certaines observations ne contiennent que la valeur auto-déclarée de l'IMC.

Pour vérifier l'existence des erreurs de mesure qu'on peut rencontrer lorsqu'on utilise les variables de poids et taille auto-déclarés, une variable dichotomique sur l'obésité basée sur la valeur auto-déclarée de l'IMC puis une autre basée sur l'IMC mesurée ont été créés. Près du tiers des obèses déclarent être non-obèses dans l'échantillon (voir tableaux 4). Plus de 16% des individus de l'échantillon déclarent qu'ils ne sont pas obèses alors qu'ils le sont par contre personne ne déclare être obèse sans l'être réellement. Il est donc important de corriger les valeurs auto-déclarées.

La procédure de correction utilisée est celle introduite par Cawley et Burkhauser (2006) et utilisée depuis dans la littérature. Il y a eu en premier lieu une prédiction de la taille mesurée en mètres avec la taille auto-déclarée. Pour réaliser la prédiction, les variables utilisées sont la taille en mètres auto-déclarée, la taille en mètres auto-déclarée au carré, l'âge et l'âge au carré (voir tableau 5). Ensuite, il y a eu une prédiction du poids en kilogrammes mesurée en utilisant des variables explicatives telles que le poids auto-déclaré, le poids auto-déclaré au carré, l'âge et l'âge au carré, de la même façon que pour la taille (voir tableau 6).

Une fois que le poids et la taille prédits sont obtenus, l'IMC est généré avec la formule  $IMC = \text{kg/m}^2$  pour les valeurs non codées. La correction a été appliquée uniquement aux observations qui ne présentent que des valeurs codées (non mesurées).

#### 2.4. Les variables instrumentales de l'IMC

L'analyse de la relation entre l'obésité et le salaire présente, tel que présenté dans l'introduction, un problème d'endogénéité. Il peut exister une corrélation entre la variable d'obésité et le terme d'erreur de l'équation due : (i) à des erreurs de mesure, (ii) aux variables omises, qui peuvent affecter à la fois le salaire et l'obésité, et / ou (iii) à la causalité inversée, c'est-à-dire que c'est le salaire qui pourrait avoir un impact sur l'obésité. La méthode des variables instrumentales permet d'adresser ce problème, dans la mesure où les instruments sont bons. Pour que les instruments soient pertinents, ils doivent être corrélés avec la variable explicative potentiellement endogène ; de plus, pour être valides, ils ne doivent pas être corrélés avec le terme d'erreur. C'est-à-dire que le seul effet que l'instrument a sur la variable dépendante doit être l'effet indirect à travers sa corrélation avec la variable endogène de droite.

Tel que mentionné dans la section 1.3, plusieurs études sur l'obésité prennent comme instrument l'IMC de l'individu retardé d'un certain nombre d'années. Dans ces études, on utilise souvent sept années de retard. Cette stratégie repose sur le fait que l'IMC est corrélée au cours du temps mais que le salaire actuel ne dépend pas directement du fait que l'individu était obèse par le passé. Étant donné que nous utilisons une coupe transversale des données, nous n'avons pas de données historiques sur l'individu au niveau de son IMC, nous ne pouvons donc pas adopter cette stratégie.

D'hombres et Brunello (2005) instrumentent l'IMC individuel avec celui de la famille biologique de l'individu catégorisé selon s'il s'agit d'un parent, d'un enfant ou de la fratrie. Pour un individu ayant plus d'une personne dans sa famille biologique, ils ont fait la moyenne des IMC des 3 catégories : ainsi, pour une personne ayant deux parents et plusieurs frères et sœurs on calculera la moyenne d'IMC des parents et celui de la fratrie séparément. Ensuite, on fera la moyenne non pondérée des deux moyennes pour l'assigner à l'individu, donc pour une personne qui a beaucoup de frères et sœurs, cela confère plus de poids aux parents de l'individu.

Pagan et Davila (1997) ont procédé de manière différente en combinant dans l'équation de salaire l'IMC à une variable d'interaction entre l'IMC et la variable d'occupation. Ils se sont intéressés à deux items : l'effet de l'obésité sur le salaire et sur les choix occupationnels. Dans l'équation de salaire ils utilisent l'estime de soi et la limitation sur la santé comme instruments. Cependant, lorsqu'ils testent l'exogénéité, ils ne peuvent pas refuser l'hypothèse d'exogénéité de l'IMC.

Morris (2005) utilise des instruments de nature géographique. Il utilise d'abord l'IMC moyen de la région où réside l'individu. Ensuite, il utilise la prévalence de l'obésité dans la région où réside l'individu. De cette manière il capte les influences sociales et environnementales qui affectent les attitudes et les comportements face à la nourriture, à l'activité sportive, et aux soins pour l'aspect physique. Cette méthode a été utilisée dans plusieurs études qui ne traitaient pas nécessairement de questions d'obésité.

Deux instruments sont proposés dans le présent document. Un premier type d'instruments consiste (comme Pagan et Davila (1997)) en des variables représentant des problèmes de santé qui sont corrélés avec l'obésité. Notamment, parmi ce type de variables, la variable dichotomique *DIABETE* (selon que les individus souffrent de diabète ou pas) sera le choix retenu. Ce choix est motivé par le fait que le diabète est hautement héréditaire et génétique et qu'il modifie directement le métabolisme de l'individu ainsi que sa capacité d'assimiler les différents types d'aliments. De plus, il est abondamment documenté dans la littérature qu'il existe un lien fort entre ce problème de santé et l'obésité ; ce lien est aussi confirmé dans notre base de données en regardant la corrélation simple (voir tableau 7) et cela en fait vraisemblablement un bon instrument.

Le deuxième type d'instruments vise à représenter les habitudes de vie des individus qui peuvent affecter le fait d'être obèse. Dans cette catégorie, la dépense énergétique quotidienne des individus (*ENERGIE*) a été choisie. La dépense énergétique est une variable disponible dans la base de donnée et calculée en utilisant la durée et la fréquence des activités physique combinées au coût énergétique de chacune des activités exprimées en kilocalories (indicateur de l'intensité de l'activité physique). Cette variable pourrait être un bon instrument car les comportements au niveau de l'exercice affectent

encore une fois le métabolisme et sont des déterminants clés de l'obésité (voir James (1995)). Cette catégorie d'instruments n'a jamais été utilisée dans la littérature et revêt donc un caractère nouveau et pourrait être pris en compte dans le développement de politiques publiques.

Toutefois, il faut se rappeler que ces instruments ne sont valides que sous l'hypothèse où ils n'affectent pas directement, ni ne sont affectés par le salaire. Autrement dit, cette hypothèse implique que la dépense énergétique n'affecte pas directement le salaire et qu'elle n'est pas affectée par ce dernier. De même pour le cas du diabète, cela implique que le diabète n'affecte pas directement le salaire et qu'il n'est pas affecté par ce dernier. Ces hypothèses pourraient être mises en doute, par exemple dans le cas du diabète. Le salaire est souvent utilisé dans la littérature comme déterminant de la santé en général, et on pourrait penser qu'il en serait de même dans le cas spécifique du diabète. Similairement, l'énergie dépensée de façon journalière pourrait avoir des liens directs avec le salaire. Cela est le cas si les travaux à bas salaire sont aussi ceux qui font dépenser plus d'énergie à cause de leur caractère manuel. Ou encore si les travaux à hauts salaires affectent la culture de la condition physique en poussant les individus à plus dépenser d'énergie en même temps qu'à mieux performer sur le marché du travail.

Malgré ces limites, étant donné les défis que présentent la recherche de bons instruments et compte tenu du fait que les résultats des tests économétriques (qui seront présentés dans les paragraphes qui suivent) soient tout de même très concluants, nous irons de l'avant avec ces deux instruments.

## CHAPITRE III

### ANALYSE EMPIRIQUE

#### 3.1. L'équation de salaire

L'objectif de notre analyse est de savoir si l'obésité affecte le salaire. Nous avons modélisé l'équation du salaire horaire correspondant de la manière suivante :

$$\text{SALAIRE}_i = \beta_0 + \beta_1 X_i + \beta_2 \text{IMC}_i + \varepsilon_i \quad (\text{A})$$

Où :

- L'indice  $i$  est pour dénoter l'individu
- $\text{SALAIRE}_i$  est le logarithme du salaire horaire
- $X_i$  est une matrice contenant les différentes variables explicatives exogènes introduites à la section 2.2
- $\text{IMC}_i$  est l'indice de masse corporel, qui est potentiellement endogène
- $\varepsilon_i$  est le terme d'erreur
- $\beta_0$  est le terme constant,  $\beta_1$  est le vecteur des coefficients des variables exogènes contenues dans  $X_i$  et  $\beta_2$  est le coefficient de l'IMC, qui est l'objet de notre analyse

L'exécution des régressions a été faite séparément pour les hommes et pour les femmes. Ceci est en accord à la fois avec la littérature en économie du travail et celle sur l'obésité. Il est bien connu en économie de travail que les équations des salaires pour les hommes et les femmes sont différentes, soit parce qu'il y a de la discrimination, soit parce que certaines caractéristiques ont une valeur différente selon qu'elles viennent d'un homme ou d'une femme. De la même façon, l'obésité est vécue de manière non identique entre les

hommes et les femmes. Les hommes et les femmes ont une constitution physique différente, ainsi, en principe, le seuil pour juger une femme comme obèse pourrait être différent du seuil pour les hommes. Cependant, à cause de manque d'études qui permettent de mesurer cette différence, dans la littérature sur l'obésité on utilise une mesure commune pour les hommes et les femmes.

L'analyse empirique du présent mémoire se base sur une première régression avec la méthode des moindres carrés ordinaires (MCO) et, afin de se charger du problème d'endogénéité de l'IMC, sur des régressions avec la méthode des variables instrumentales (VI).

Les résultats des régressions ainsi que des différents tests sur les instruments feront l'objet d'une discussion dans la section qui suit.

### 3.2. Résultats

Les résultats des régressions sont rapportés dans les tableaux 10 et 11, pour les hommes et les femmes respectivement. L'option « robuste » est utilisée dans toutes les régressions pour contrôler les problèmes potentiels d'hétéroscédasticité dans les termes d'erreur. La première colonne contient les résultats de la régression par moindres carrés ordinaires (MCO). Le coefficient de l'IMC est positif chez les hommes et négatif chez les femmes. Ce résultat s'apparente à ce qu'on retrouve, entre autres, dans Cawley (2004) et Garcia et al (2006). Cependant, dans les deux cas le coefficient est très faible et statistiquement non-significatif.

La deuxième et troisième colonne contiennent les résultats de la régression avec la méthode des variables instrumentales (VI) lorsqu'on considère le premier et le deuxième instrument respectivement (*ENERGIE*, et *DIABETE*). La quatrième colonne contient les résultats de la régression avec la méthode VI lorsqu'on considère les deux instruments ensemble.

La méthode des variables instrumentales considère la meilleure combinaison linéaire d'instruments. En théorie, lorsqu'on a plusieurs instruments il est optimal de tous les utiliser. Cependant, lorsque les instruments sont faibles, à cause des considérations de la section 2.4 combinées au fait que la corrélation simple entre chaque instrument et l'IMC n'est pas assez élevée (voir les corrélations simples dans le tableau 7 ainsi que les résultats de la première étape dans les tableaux 8 et 9), l'utilisation d'un nombre élevé d'instruments génère une perte d'efficacité et il est suggéré dans ce cas (voir entre autres Staiger et Stock (1997)) d'en utiliser un nombre limité. À cause de cette considération, les régressions furent réalisées en choisissant parmi les variables disponibles dans la base de données celles qui nous semblaient être les meilleurs instruments. Dans un premier temps les deux d'instruments sont utilisés séparément et ensuite, les instruments sont utilisés ensembles.

Un test d'exogénéité a été réalisé pour vérifier la validité de l'utilisation de la méthode des variables instrumentales. En effet, lorsqu'il n'y a pas de problème d'endogénéité, l'utilisation des moindres carrés ordinaires est préférée à la méthode des variables instrumentales, car bien que toutes deux appropriées, c'est la méthode des moindres carrés ordinaires qui est la plus efficace. Cependant, dans le cas où la variable de droite serait endogène, la méthode des moindres carrés ordinaires n'est pas adéquate. Par contre, il est adéquat d'utiliser la méthode des variables instrumentales dans ce cas.

Pour faire cette vérification, la fonction *ESTAT Endogeneous*, (qui est disponible dans le logiciel STATA) a été utilisé. Cette fonction effectue un test de spécification de Hausman adapté à la situation où les erreurs standards sont robustes. Dans le test de Hausman, l'hypothèse nulle est telle que les coefficients de la régression simple (OLS) sont égaux aux coefficients du modèle avec variables instrumentales. Dans toutes les spécifications présentées, à exception du modèle 2 pour les femmes, les p-value obtenues sont très proches de zéro. L'hypothèse nulle du test de Hausman est donc rejetée à un niveau de confiance standard. Comme conclusion, la méthode d'estimation appropriée est la méthode des variables instrumentales.

Il faut noter que le test de Hausman sur l'égalité des coefficients des deux estimateurs est équivalent au test d'exogénéité proposé par Wooldridge (2002). Ce test

consiste à faire une première régression (MCO) contenant toutes les variables exogènes et tous les instruments, suivie d'une récupération du résidu de cette régression. Ce résidu est ensuite utilisé avec les variables exogènes de la section 2.2 dans la régression (MCO) du salaire. Une analyse des statistiques t du résidu des équations est subséquentement réalisée, permettant ainsi de juger de l'exogénéité des variables. Il faut néanmoins noter que ces deux tests d'exogénéité ne sont valides que lorsque les instruments sont fiables.

Les résultats montrent qu'une fois instrumentée, l'obésité s'avère être une variable ayant un pouvoir explicatif sur le salaire. Dans tous les cas, l'IMC s'avère être en relation inverse avec le salaire tant chez les hommes que chez les femmes.

Les résultats des régressions (voir table 8) sont tels qu'une augmentation d'un point d'IMC correspond pour les hommes à une réduction de salaire de 9%, 2,5% et 2,9% pour les modèles 1, 2, et 3 respectivement. Parallèlement, pour les femmes, une augmentation d'un point d'IMC affecte le salaire de manière telle qu'il y a diminution de 6%, 2,7% et 4,7% pour les modèles 1, 2, et 3 respectivement. Cependant, le coefficient du modèle 2 pour les femmes n'est pas significatif. Les résultats sont donc quantitativement assez différents pour les hommes et pour les femmes.

Les résultats du paragraphe précédent sont en ligne avec ceux de Cawley (2004), qui considère des régressions séparées par race. Pour les sujets de race blanche, il trouve que le coefficient de l'IMC de son équation salariale par moindres carrés ordinaires est  $-0.001$  et  $-0.01$  pour les hommes et les femmes respectivement. En instrumentant l'IMC, les coefficients de Cawley sont  $-0.013$  pour les hommes et  $-0.017$  pour les femmes. Alors, bien que qualitativement proches des résultats de Cawley, les résultats obtenus avec des données canadiennes semblent indiquer qu'au Canada l'obésité a un impact beaucoup plus élevé sur les salaires qu'aux États-Unis.

La littérature antérieure à Cawley (2004) trouve un effet soit nul soit faiblement positif de l'obésité sur les salaires des hommes et dans toute la littérature l'impact de l'obésité sur le salaire chez les hommes est systématiquement inférieur à l'impact chez les femmes. Les résultats présentés révèlent que les coefficients ont le même signe tant chez

les hommes que les femmes et que la magnitude est soit la même soit plus importante chez les hommes que chez les femmes.

Tel que mentionné précédemment, pour que les instruments soient considérés comme étant conformes, la première condition consiste à avoir des instruments corrélés avec la variable instrumentée, c'est-à-dire que  $E(\text{imc}, \text{instrument}) \neq 0$ . Pour faire la vérification de cette condition, il faut considérer la première étape de la régression, en d'autres mots, la régression de l'IMC sur toutes les variables exogènes du modèle incluant les variables instrumentales de l'IMC. En particulier, il faut observer la statistique F correspondant à l'hypothèse que les coefficients des variables instrumentales soient conjointement zéro. Lorsqu'on utilise une seule variable instrumentale cela équivaut à regarder la statistique  $t$  du coefficient de la variable en question.

Dans la littérature empirique, il est généralement accepté comme condition nécessaire pour que les instruments soient « bons » que la statistique F de la première étape soit au moins supérieure à 10 (voir, Staiger et Stock (1997)). Les statistiques F correspondantes à nos modèles sont reportées aux tableaux 10 et 11. Elles sont toutes supérieures à 10. Le résultat de ce test renforce notre confiance quant à la qualité des instruments.

Bien qu'il soit nécessaire que les instruments soient corrélés avec la variable à instrumenter, cette condition n'est pas suffisante pour affirmer que l'on a de bons instruments. Une deuxième condition s'impose à l'effet que les instruments ne doivent pas être corrélés avec le terme d'erreur. Malheureusement, il n'y a pas une façon de tester directement cette condition. Cependant, lorsqu'on a un nombre d'instruments supérieur au nombre des variables endogènes, il est possible de procéder à un test de sur-identification pour tester davantage la validité des instruments.

Étant donné qu'il y a deux instruments dans le modèle 3, il est possible de faire en complément un test de Hausman pour la sur-identification. Ceci permet de tester l'hypothèse qu'un ou plusieurs instruments soient corrélés avec le terme d'erreur. L'évaluation de la sur-identification est réalisée avec un test issu de l'application STATA

(fonction *ESTAT override*). Le principe général de ce test consiste à comparer la régression réalisée avec tous les instruments à des régressions faites avec de multiples sous-ensembles d'instruments. L'hypothèse nulle est que l'espérance du produit des instruments au résidu de l'équation de salaire faite avec les instruments et les variables exogènes du modèle sont les mêmes. Dans le cas d'un rejet de l'hypothèse nulle, il faut reconsidérer la validité des instruments. Tant chez les hommes que chez les femmes, les résultats confirment que l'hypothèse nulle ne peut être refusée. Donc sur la base des tests réalisés, nos instruments semblent se qualifier comme des bons instruments.

## CONCLUSION

Ce mémoire vise à évaluer l'impact de l'obésité sur les salaires au Canada en utilisant les données de l'ESCC 2004. À ce propos des régressions des salaires avec la méthode des moindres carrés ordinaires (MCO) ainsi qu'avec la méthode des variables instrumentales (VI), pour tenir compte de l'endogénéité de l'obésité, ont été conduites. Pour la méthode des variables instrumentales, deux instruments ont été considérés : un premier basé sur les problèmes de santé de l'individu et l'autre relié à ses habitudes de vie qui affectent la dépense d'énergie et le métabolisme. L'utilisation de ce deuxième type d'instrument constitue une nouveauté dans la littérature empirique sur l'obésité et en fait donc une des contributions de ce mémoire. Trois spécifications ont été considérées : initialement en utilisant chaque variable instrumentale séparément et ensuite en les utilisant ensemble. Différents tests confirment la validité des instruments utilisés.

Les régressions avec la méthode MCO illustrent un IMC présentant un coefficient positif chez les hommes et négatif chez les femmes. Cependant, aucun de ces coefficients n'est statistiquement significatif. Par contre, une fois instrumentée, l'IMC devient une variable significative dans la détermination du salaire, tant pour les hommes que pour les femmes et peu importe la manière dont l'obésité est instrumentée. Dans tous les scénarios évalués avec variables instrumentales, l'IMC a un impact négatif sur le salaire.

La relation entre l'IMC et le salaire varie, selon les instruments utilisés, entre  $-2,5\%$  et  $-9\%$  pour une augmentation d'un point d'IMC pour les hommes et entre  $-2.6\%$  et  $-6\%$  pour les femmes. Le fait que les résultats soient sensibles au type d'instrument utilisé semble indiquer que chaque instrument capture un effet sur des segments différents de la population. L'impact de l'obésité sur le salaire est plus grand lorsqu'on instrumente avec

des indicateurs liés à la dépense d'énergie que lorsqu'on utilise des instruments liés aux problèmes de santé. Étant donné les limites des résultats de tests économétriques, et compte tenu des hypothèses sur lesquelles repose l'analyse, il faut demeurer modéré face à l'interprétation des résultats quantitatifs. En outre, dans l'équation de salaire, nous faisons abstraction de l'enjeu du biais de sélection. L'obésité pourrait affecter la probabilité d'occuper un emploi. Cet aspect n'est pas traité dans le présent document, et d'ailleurs il ne l'est pas dans la plupart des travaux empiriques existants à date sur la relation entre obésité et salaire. Cette absence doit être prise en considération dans l'interprétation des résultats et devrait éventuellement être incorporée dans des analyses futures.

Les résultats de ce mémoire sont qualitativement en ligne avec les résultats issus des études faites, entre autres, par Garcia et al (2006) pour l'Europe et par Pagan et Davila (1997) pour les États-Unis. Néanmoins, les résultats se distinguent de ceux obtenus aux États-Unis parce que selon notre analyse la relation négative entre la performance sur le marché du travail et l'obésité n'affecte pas seulement les femmes mais aussi les hommes, avec un impact même plus fort chez ces derniers. La magnitude des coefficients pour le Canada est plus élevée que celle obtenue dans des études américaines comparables, ce qui pourrait signifier que la pénalité de l'obésité serait plus grande sur le marché du travail canadien que sur le marché du travail américain.

Pour terminer, il faut ajouter que l'analyse accomplie dans ce mémoire ne permet pas d'évaluer les différents déterminants de l'effet de l'obésité sur le salaire. De plus, la méthodologie utilisée ne prend pas en compte le choix de participation dans le marché du travail, pas plus que les choix occupationnels des individus. Ce manque est commun à la plupart des articles dans la littérature sur le sujet. Pourtant l'inclusion de ces éléments serait une importante avenue de recherche future et pourrait servir entre autre à évaluer la robustesse des résultats obtenus dans le présent document.

## BIBLIOGRAPHIE

AKERLOF G. «Discriminatory, Status-Based Wages Among Tradition-Oriented Stochastically Trading Coconut Producers», *Journal of Political Economy*, 93(2), 1985, p. 265-276

American Society for Reproductive Medicine ASRM «Fact sheet: Weight and fertility», 2001, <http://www.asrm.org/Patients/FactSheets/weightfertility.pdf>

ANDERSON P.M., K.F. BUTCHER, et P.B. LEVINE. «Economic Perspectives on Childhood Obesity», *Economic Perspectives*, 3Q/2003, 2003, p.30-48

ARROW K. «Some Mathematical Models of Race Discrimination in the Labor Market», *Racial Discrimination in Economic Life*, Lexington Books : DC Heath, 1972, chapitre 6, p. 187-203

ARROW K. «The Theory of Discrimination», in *Discrimination in Labor Markets*, Princeton University Press, 1973, p. 3-33

AULD C. ET L.M. POWELL «The Economics of Obesity: Research and Policy Implications from a Canada-U.S. Comparison», *Draft paper from Institute for Health Research and Policy at the University of Illinois at Chicago*, 2005

AVERETT S. et S. KORENMAN. «Black and White Differences in Social and Economic Consequences of Obesity», *International Journal of Obesity*, 23(2), 1999, p. 166-173

BECKER G. S. *The Economics of Discrimination*, University of Chicago Press, 1957

BEHRMAN J. et M. R. ROSENZWEIG. «The Returns to Increasing Body Weight », *Penn Institute for Economic Research (PIER) Working Paper 01-052*, 2001

BIRMINGHAM C.L. et al. «The Cost of Obesity in Canada», *Journal de l'Association Médicale Canadienne*, 160(4), 1999, p. 483-488

BRUNELLO G. et B. D'HOMBRES. «Does Obesity Hurt your Wages more in Dublin than in Madrid », *IZA Discussion Paper No. 1704*, 2005

CAWLEY J. et R.V. BURKHAUSER «The importance of Objective Health Measures in Predicting Early receipt of Social Security Benefits : The Case of Fatness», *University of Michigan Working paper WP148*, Michigan Retirement Research Center, 2006

CAWLEY J. et S. DANZIGER. «Obesity as a barrier to the transition from welfare to work», *National Bureau of Economic Research NBER working paper No. W10508*, 2004, p. 1-22

CAWLEY J. «The impact of obesity on Wages», *Journal of Human Resources*, 39(2), 2004, p.451-474

CAWLEY J. «Body Weight and Women's Labor Market Outcomes » *National Bureau of Economic Research NBER Working Papers No 7841*, 2000

CHRISTAKIS N.A.et J.H. FOWLER «The Spread of obesity in Large Social Network over 32 Years», *The New England journal of medicine*, 357, 2007, p. 370-379

CONLEY D. et R. GLAUBER. «Gender, Body Mass and Socioeconomic Status : The new evidence from PSID », *National Bureau of Economic Research NBER Working Paper No. W11343*, 2007

EID J., H. OVERMAN, D. PUGA, et M. TURNER «Fat City : The relationship between urban Sprawl & obesity», *Journal of Urban Economics* 63, 2006, p. 385-404

GALTIER-DEREURE Florence et al. «Obesity and pregnancy : complications and cost», *American Journal of Clinical nutrition*, 71(5), 2000, S1242-S1248

GORTMAKER S.L., A. MUST, J.M. PERRIN, A.M. SOBOL, et W.H. DIETZ, «Social and economic consequences of overweight among adolescents and young adults.» *New England Journal of Medicine*, 329, 1993, p.1008-12

GARCIA J. et C. QUINTANA-DOMEQUE. « Income and Body Mass Index in Europe», *Economics Working Papers 1001, Department of Economics and Business, Universitat Pompeu Fabra*, 2006 révisé en mai 2008

GARFINKEL L. «Overweight and cancer», *Annals of Internal Medicine*, 103(6.2), 1985, p. 1034-1036

HAMERMESH D.S. et J.E. BIDDLE. «Beauty and the Labor Market», *The American Economic Review*, 84(5), 1994, p. 1174-1194

HAVET N. «Écarts salariaux et disparités professionnelles entre sexes: développements théoriques et validité empirique», *L'Actualité économique, Revue d'analyse économique*, 80(1), 2004, p. 6-16

HUBERT H.B. et al. «Obesity as an independent risk factor for cardiovascular disease: a 26 year follow up of participants on the Framingham heart study», *Circulation (American Heart Association)*, 67, 1983, p. 968-977

JAMES W.P. «A public health approach to the problem of obesity», *International Journal of Obesity and Related Metabolic Disorders*, 19(3), 1995, S37-S45

KATZMARZYK P. et al. «The Economic Burden of Physical Inactivity in Canada», *Journal de l'Association Médicale Canadienne*, 163(11), 2000, p. 1435-1440

KATZMARZYK P.T. et I. JANSEN . «The Economic Costs Associated With Physical Inactivity and Obesity in Canada: an Update», *Revue Canadienne de Physiologie Appliquée*, 29, 2004, p. 90-115

LAKDAWALLA D. ET T PHILIPSON, «Technological Change and the Growth of Obesity: A Theoretical and Empirical Examination», NBER *National Bureau of Economic Research Working Paper No 8446*, 2002

MACMINN W., J. MCINTOSH et C. YUNG «How much does obesity matter? Results from the 2001 canadian community health survey» ,*Advances in Health Economics and Health Services Research*, Volume 17, 2007, 337–369

MICHAUD P.-C., A.H.O. VAN SOEST, et T. ANDREYEVA. «Cross-Country Variation in Obesity Patterns among Older Americans and Europeans», *Forum for Health Economics & Policy*, 10(2), 2007, Article 8

MORRIS S. «Body mass index and occupational attainment», *Tanaka Business School Discussion Papers*, 2005, TBS/DP05/41

Organisation Mondiale de la Santé [OMS]. (2003) «Obesity and overweight», *rapport de l'OMS*, 2003, [http://www.who.int/dietphysicalactivity/media/en/gsf\\_s\\_obesity.pdf](http://www.who.int/dietphysicalactivity/media/en/gsf_s_obesity.pdf)

PAGAN J.A. et A. DAVILA «Obesity, occupational attainment and earnings», *Social Science Quarterly*, 78, 1997, p.756-770

PHELPS E. «The Statistical Theory of Racism and Sexism», *American Economic Review*, 62(4), 1972, p. 659-661

RAINE K.D. «Le surpoids et l'obésité au Canada : une perspective de la santé de la population», *Rapport de l'Institut Canadien d'Information sur la Santé*, 2004

ROOTH D.O. «Evidence of Unequal Treatment in Hiring Against Obese Applicants – A Field Experiment», *IZA Discussion Paper No. 2775*, 2007

SANTÉ Canada. « Lignes directrices canadiennes pour la classification du poids chez les adultes », Ottawa, Santé Canada, 2003

SHAYEB A.G.et S. BHATTACHARYA. «An exploration of the association between male body mass index and semen quality» *European Society of Human Reproduction and Embryology ESHRE meeting 2008*

STAIGER D. et J.H. STOCK «Instrumental variable regression with weak instruments», *Econometrica*, 65(3), 1997, p. 557-586

STIGLITZ J.E.«The Theory of 'Screening', Education, and the Distribution of Income», *Cowles Foundation Discussion Papers No 354*, 1973.

VAILLANCOURT F. «Differences in earning by language groups in Quebec, 1970 : an economic analysis», Centre international de recherche sur le bilinguisme de l'Université Laval, 1980, ICRB-B-90

VUILLE J.-C. et T. MELBIN. «Obesity in 10-Year-Olds: An Epidemiologic Study», *Official Journal of the American Academy of Pediatrics*, 64(5), 1979, p. 564-572

WOOLDRIDGE J.M. *Econometric analysis of cross-section and panel data*, MIT Press, Cambridge, Massachusetts, 2002

ZIVOT E., C.R. NELSON, R. STARTZ. «Valid Confidence Intervals and Inference in the Presence of Weak Instruments» , *Discussion Papers in Economics at the University of Washington 96-15*, 1996

## ANNEXE

**Tableau 1 : classification du risque pour la santé en fonction de l'indice de masse corporelle (IMC)**

$IMC = \text{poids(kg)/taille(m)}^2$		
Classification	Catégorie de l'IMC (kg/m <sup>2</sup> )	Risque de développer des problèmes de santé
Poids insuffisant	< 18,5	Accru
Poids normal	18,5 - 24,9	Moindre
Excès de poids	25,0 - 29,9	Accru
Obésité, classe I	30,0 - 34,9	Élevé
Obésité, classe II	35,0 - 39,9	Très élevé
Obésité, classe III	>= 40,0	Extrêmement élevé

Source: Santé Canada. Lignes directrices canadiennes pour la classification du poids chez les adultes. Ministère des Travaux publics et Services gouvernementaux du Canada, 2003

**Tableau 2 : revenu personnel généré par interpolation pour individu de 18 à 64 ans**

Revenu	Fréq.	Pourcentage	Cum.
7 500	1 068	13,51	13,51
22 500	2 146	27,15	40,67
40 000	2 535	32,08	72,74
65 000	1 609	20,36	93,10
80 000	545	6,90	100,00
Total	7 903	100,00	

Tableau 3 : statistiques descriptives des personnes âgées de 18 à 64 ans

Variables	Moy.	E.Type		Moy.	E.Type		Moy.	E.Type
Marié	0,4120	0,4922		0,4250	0,4944		0,3966	0,4893
Race blanche	0,8733	0,3326		0,8739	0,3320		0,8727	0,3334
Age	40,4149	12,4196		40,6209	12,4451		40,1714	12,3867
Salaire_h	16,9987	9,4991		18,3139	9,5770		15,4433	9,1670
Enfant	0,1442	0,3514		0,1338	0,3405		0,1566	0,3635
Bilingue	0,1888	0,3914		0,1822	0,3860		0,1966	0,3975
Nécosse	0,0487	0,2153		0,0399	0,1958		0,0591	0,2358
Tneuve	0,0544	0,2268		0,0603	0,2380		0,0475	0,2127
Nbruns	0,0496	0,2171		0,0474	0,2125		0,0522	0,2225
Ipe	0,0491	0,2161		0,0448	0,2070		0,0541	0,2263
Québec	0,1470	0,3542		0,1604	0,3671		0,1312	0,3376
Ontario	0,2686	0,4433		0,2709	0,4445		0,2659	0,4419
Manitoba	0,1272	0,3332		0,1238	0,3294		0,1312	0,3376
Sask	0,0549	0,2278		0,0516	0,2213		0,0588	0,2353
Alberta	0,1063	0,3082		0,1077	0,3100		0,1047	0,3062
Columbieb	0,0941	0,2920		0,0932	0,2907		0,0953	0,2936
Secondaire	0,1865	0,3895		0,1908	0,3930		0,1814	0,3854
Postsecnon	0,0967	0,2955		0,0922	0,2894		0,1019	0,3026
Postsecond	0,5786	0,4938		0,5460	0,4979		0,6172	0,4861
Primaire	0,1382	0,3451		0,1709	0,3765		0,0994	0,2993
Diabète	0,0381	0,1914		0,0453	0,2080		0,0295	0,1694
Hautepression	0,1034	0,3045		0,1123	0,3158		0,0928	0,2902
Fruits légumes	3,8920	2,1361		3,6447	2,1383		4,1844	2,0964
Énergie	1,7247	1,7765		1,7649	1,8773		1,6772	1,6485
Imc final	27,5216	5,4986		27,8348	4,8557		27,1513	6,1533
Total	7 903			Homme	4 282		Femme	3 621

Tableau 4 : personnes obèses auto-déclarées versus personnes obèses mesurées

obèse_aut	obèse mesuré		Total
	non	oui	
non	366	28	394
oui	0	85	85
Total	366	113	479

obèse\_aut :obèse selon données auto-déclarées

obèse\_mesuré :obèse selon données mesurées

Tableau 5 : coefficients de l'équation de correction de la taille (en mètres) auto-déclarée (aut)

Taille mesurée	Hommes	Femmes
Taille_aut	1,287 (1,19)	-0,711 (-0,58)
Taille_aut_carré	-0,011 (-0,36)	0,489 (1,32)
Age	-0,002 (-1,06)	-0,000 (-0,15)
Age_carré	0,000 (1,06)	0,000 (0,17)
_cons	-0,15 (-0,16)	1,486 (1,47)
Observations	271	212
R-carré	0,7782	0,7829

1) Statistiques t entre parenthèses

2) Les astérisques indiquent le niveau de signification statistique:

\*\*\* p<0,01, \*\* p<0,05, \* p<0,1

Tableau 6 : coefficients de l'équation de correction du poids (en kilogrammes) auto-déclaré (aut)

Poids mesuré	Hommes	Femmes
Poids_aut	0,834*** (5,99)	1,039*** (9,63)
Poids_aut_carré	0,001 (1,24)	-0,000 (-0,11)
Age	0,2577* (1,86)	0,110 (0,85)
Age_carré	-0,002 (-1,46)	-0,001 (-0,89)
_cons	2,747 (0,41)	-1,661 (-0,36)
Observations	271	210
R-carré	0,9315	0,9562

1) Statistiques t entre parenthèses

2) Les astérisques indiquent le niveau de signification statistique:

\*\*\* p<0,01, \*\* p<0,05, \* p<0,1

Tableau 7 : corrélation entre l'IMC et instruments de l'équation de salaire

	Imc
Diabète	0,1718
Énergie	-0,1062

Tableau 8 : régression variable salaire avec IMC, première étape - échantillon hommes

	MODÈLE 1	MODÈLE 2	MODÈLE 3
Énergie (coef)	-0,1283868		-0,1067005
T-stat	-3,19		-2,71
P-value	0,001		0,007
Diabète (coef)		4,600428	4,560428
T-stat		9,82	9,77
P-value		0	0
R carré	0,0787	0,1138	0,1153
F-test	10,18	96,47	51,58

Chaque régression inclut les autres variables exogènes Xi.

Tableau 9 : régression variable salaire avec IMC, première étape - échantillon femmes

	MODÈLE 1	MODÈLE 2	MODÈLE 3
Énergie (coef)	-0,4154041		-0,4017835
T-stat	-7,64		-7,42
P-value	0		0
Diabète (coef)		3,333592	3,174854
T-stat		4,74	4,5
P-value		0	0
R carré	0,0724	0,0689	0,0798
F test	58,38	22,47	39,62

Chaque régression inclut les autres variables exogènes Xi.

Tableau 10 : régression variable salaire avec IMC - échantillon hommes

Variabes	MCO	MODÈLE 1	MODÈLE 2	MODÈLE 3
Imc	0,0016 (0,82)	-0,0906 (-1,82)*	-0,0257 (-2,45)**	-0,0289 (-2,78)***
Race_blanche	0,1940 (6,46)***	0,2696 (5)***	0,2164 (6,84)***	0,2190 (6,9)***
Secondaire	0,2184 (7,32)***	0,1872 (4,51)***	0,2092 (6,78)***	0,2081 (6,71)***
Postsecnon	0,1792 (4,42)***	0,1756 (3,54)***	0,1782 (4,29)***	0,1780 (4,27)***
Postsecond	0,3686 (13,92)***	0,3418 (9,5)***	0,3607 (13,26)***	0,3597 (13,16)***
Enfant	0,0297 (1,14)	0,0294 (0,89)	0,0297 (1,11)	0,0296 (1,10)
Marié	0,1491 (7,31)***	0,2068 (5,28)***	0,1661 (7,66)***	0,1681 (7,75)***
Age	0,0788 (14,06)***	0,1027 (7,03)***	0,0859 (13,42)***	0,0867 (13,54)***
Age_carré	-0,0008 (-11,76)***	-0,0010 (-7,1)***	-0,0009 (-11,62)***	-0,0009 (-11,69)***
Bilingue	0,0760 (3,16)***	0,0441 (1,27)	0,0666 (2,67)***	0,0655 (2,61)***
_cons <sup>1</sup>	0,0652 (0,52)	2,0735 (1,89)*	0,6586 (2,57)**	0,7279 (2,85)***
Dummies provinces <sup>2</sup>	oui	oui	oui	oui
Endogénéité (Robust score chi <sup>2</sup> <sup>3</sup> )	--	4,9915 (p =0,0255)	7,2952 (p =0,0069)	9,41236 (p =0,0022)
Overid <sup>4</sup>	--	--	--	2,28891 (p =0,1303)
Observations	4 282	4 282	4 282	4 282
Instruments :				
Diabète			x	x
Énergie		x		x

<sup>1</sup> Constante

<sup>2</sup> Dummies provinces : variables dichotomiques de provinces de résidence

<sup>3</sup> Robust score chi<sup>2</sup> : Statistique khi carré robuste

<sup>4</sup> Overid : Test de suridentification

Tableau 11 : régression variable salaire avec IMC - échantillon femmes

Variabiles	MCO	MODELE 1	MODELE 2	MODELE 3
Imc	-0,0024 (-1,25)	-0,0599 (-3,38)***	-0,0268 (-1,31)	-0,0464 (-3,45)***
Race_blanche	0,1337 (3,95)***	0,1544 (3,95)***	0,1425 (3,98)***	0,1496 (4,02)***
Secondaire	0,2588 (6,07)***	0,2111 (4,19)***	0,2386 (5,1)***	0,2223 (4,7)***
Postsecnon	0,1959 (4,03)***	0,1819 (3,34)***	0,1899 (3,84)***	0,1851 (3,57)***
Postsecond	0,5247 (13,95)***	0,4910 (11,51)***	0,5104 (12,89)***	0,4989 (12,38)***
Enfant	0,0107 (0,34)	0,0393 (1,09)	0,0229 (0,69)	0,0326 (0,96)
Marié	-0,0095 (-0,4)	-0,0297 (-1,1)	-0,0181 (-0,72)	-0,0250 (-0,97)
Age	0,0904 (13,35)***	0,1129 (11,1)***	0,1000 (9,52)***	0,1076 (12,17)***
Age_carré	-0,0009 (-11,29)***	-0,0012 (-10,18)***	-0,0010 (-9,08)***	-0,0011 (-10,91)***
Bilingue	0,0515 (1,85)*	0,0456 (1,44)	0,0490 (1,71)*	0,0470 (1,55)
_cons <sup>5</sup>	-0,0698 (-0,48)	1,0379 (2,75)***	0,4004 (0,96)	0,7779 (2,58)**
Dummies provinces <sup>6</sup>	oui	oui	oui	oui
Endogénéité (Robust score chi2 <sup>7</sup> )	--	12,6772 (p=0,0004)	1,41884 (p=0,2336)	11,7634 (p=0,0006)
Overid <sup>8</sup>	--	--	--	1,49202 (p=0,2219)
Observations	3 621	3 621	3 621	3 621
Instruments :				
Diabète			x	x
Énergie		x		x

<sup>5</sup> Constante

<sup>6</sup> Dummies provinces : variables dichotomiques de provinces de résidence

<sup>7</sup> Robust score chi2 : Statistique khi carré robuste

<sup>8</sup> Overid : Test de suridentification