

Economics School of Louvain - ESL

Economics School of Namur - ESN

Existe-t-il un potentiel de travail inexploité parmi la population âgée de Belgique ?

Application des méthodes d'estimation de la capacité à travailler

Auteur : Pire Paul

Promoteur : Vincent Vandenberghe

Lecteur : Bruno Van der Linden

Année académique 2018-2019

Master en sciences économiques – 120 crédits – Finalité : spécialisée

Résumé

L'accroissement de l'espérance de vie engendre l'augmentation des dépenses du système de pensions belge et entraîne celui-ci vers une situation de moins en moins soutenable sur le long terme. Une des solutions proposées et déjà en application un peu partout dans le monde est l'augmentation de l'âge légal de départ en retraite, ce qui permettra de réduire les dépenses de la sécurité sociale tout en augmentant ses rentrées. Cette solution, motivée principalement par le besoin ne sera couronnée de succès que si la santé des personnes touchées leur permet de travailler à des âges avancés. C'est la question de départ de ce travail : « Est-ce que la population âgée jouit d'une santé suffisante pour travailler au-delà de l'âge actuel d'éligibilité à la retraite ? ». Nous parcourons la recherche économique actuelle en termes de potentiel de travail à des âges avancés en démarrant notre analyse par une brève revue des facteurs de départ en retraite étudiés par Dora L. Costa dans son ouvrage sur l'histoire de la retraite aux Etats-Unis (Costa, 1998). Nous continuons et nous attarderons sur les méthodes d'estimation de la capacité à travailler. La mise en œuvre de la méthode de Cutler et al. (2012) au moyen de données de l'Enquête de santé nous apprendra qu'il existe en Belgique un potentiel de travail inexploité important parmi ses individus retraités car les prédictions de la capacité à travailler (du fait de la détérioration de leur santé avec l'âge) sont bien inférieures à la baisse de leur taux d'emploi. Nous verrons également que ce potentiel de travail, au-delà de la santé, dépend plus largement de la catégorie d'emploi et de facteurs démographiques comme l'éducation, le genre et la région de résidence.

Mes sincères remerciements

A l'ensemble des personnes qui m'ont permis de rédiger ce travail ; mon promoteur Vincent Vandenberghe sans qui rien n'aurait été possible. Aux différents auteurs, cités ici et aux autres, qui ont été une grande source d'inspiration. A ma famille ; mes parents que j'aime et qui m'ont toujours soutenu à leur manière... Plus particulièrement à ma sœur pour son soutien et ses conseils avisés et à ma Mamie ; pour son soutien également et surtout pour son courage, sa patience et sa tolérance face aux fautes d'orthographe qu'elle m'a permis de corriger. A Solange Dujardin, la secrétaire de l'ESL, qui fait un travail exceptionnel et sans qui je me serais perdu depuis longtemps dans les rouages de ce master. Aussi au point de repère de Louvain-La-Neuve qui m'a mis en contact avec Françoise Voglaire, une psychologue de talent, qui m'a énormément aidé à un moment difficile de mon parcours étudiant. A vous tous, merci beaucoup.

Plan de ce travail

- 1) Motivation du travail
- 2) Littérature introductive
 - 1) Les facteurs de départ en retraite selon Dora I. Costa
 - 2) La capacité à travailler à des âges avancés
- 3) La méthode de Milligan-Wise
- 4) La méthode de Cutler
 - 1) Données
 - 2) La santé subjective
 - 1) Définition binaire de la (mauvaise) santé
 - 2) Analyse multi catégorielle de la santé
 - 3) La santé subjective et les indicateurs de handicap
 - 3) La santé objective
 - 1) Avec un seul indicateur
 - 2) Avec deux indicateurs et plus
 - 3) Avec l'ensemble des indicateurs
 - 4) Avec la santé subjective et objective
 - 1) Toutes les variables en une seule régression
 - 2) Instrumenter la santé subjective par la santé objective
 - 5) Synthétiser la santé en un seul indicateur
 - 6) Les variables de contrôles
 - 1) Le genre
 - 2) L'éducation
 - 3) Les catégories d'emploi
 - 4) Les régions
- 5) Conclusions
 - 1) Résumé des résultats
 - 2) Résumé des limites de ce travail
 - 3) Ce que ce travail m'a apporté
- 6) Bibliographie
- 7) Annexes (Code Stata – version secrétariat)

1. Motivation du travail

Une tendance actuelle des politiques de l'emploi est l'augmentation de l'âge d'accessibilité à la retraite. Ces politiques sont motivées par un déficit croissant de la sécurité sociale lié à l'augmentation de l'espérance de vie. En Belgique, la sécurité sociale fonctionne en deux temps ; durant toute notre carrière nous cotisons, via l'Etat, pour notre retraite tandis que lorsque que celle-ci prend fin, l'Etat nous rémunère pour assurer financièrement nos vieux jours ; c'est le système de pensions « Pay-as-you-go ». La sécurité sociale prélève alors les cotisations parmi les travailleurs et ces cotisations servent à rémunérer les retraités de cette même période. Une phrase tirée d'un article de l'Echo publié en janvier 2018 et écrit par Jean Hindriks nous explique simplement le problème actuel de ce système de pension ;

« Depuis 2003, les dépenses de pension ont augmenté de 80% tandis que les cotisations n'ont augmenté que de 40%. »

Si un déficit constant de la sécurité sociale est loin d'être inenvisageable, un déficit croissant est quant à lui inquiétant parce qu'il met en péril la soutenabilité même du système. Or, comme l'espérance de vie n'a cessé d'augmenter, les dépenses de rémunérations sociales croissent plus vite que les cotisations sociales. Finalement le déséquilibre entre les recettes et les dépenses de la sécurité sociale s'accroît avec le poids d'une population de plus en plus âgée. Dans ce contexte de population de plus en plus âgée et de déficit de la sécurité sociale, une solution des plus simples et efficaces pour assurer la pérennité du système des retraites est d'augmenter le taux d'emploi des personnes âgées. Ceci permettrait de réduire les dépenses de pensions tout en augmentant les recettes de cotisations puisque le nombre de personnes à rémunérer se rétrécit là où le nombre de cotisants s'élargit. Différents travaux ont déjà montré que les différences de taux d'emploi entre les pays à un moment donné sont expliquées majoritairement par des différences entre les programmes de retraite (Gruber, 1999).

Beaucoup de pays ont déjà lancé des réformes de retraite pour faire face au vieillissement de la population. La Belgique ne fait pas exception puisque l'âge légal de départ en retraite de 65 ans sera porté à 66 ans en 2025 et à 67 ans en 2030 (Service fédéral des Pensions, 2019). Selon certains chercheurs ; augmenter l'âge d'éligibilité à la retraite est une approche efficace pour stimuler l'emploi des personnes âgées. Une question se pose naturellement : ces personnes âgées, sont-elles en suffisamment bonne santé pour travailler ? Si elles le sont, est-ce que leur capacité à travailler additionnelle est importante ? Et est-elle similaire pour tout le monde en Belgique ?

La suite de ce document comprend les sections 2 à 5. Dans la section 1 nous passons en revue la littérature économique générale sur les déterminants du départ à la retraite. Nous y précisons la notion, centrale à ce travail, de "capacité à travailler". La section 3 expose une première manière de la quantifier. La section 4 expose la méthode de Culter et al. que nous privilégions et retenons pour son application aux données belges de l'Enquête Santé 2013. La Section 5 expose nos conclusions.

2. Littérature introductive

Nous entamons cette revue par la littérature économique générale sur l'évolution de l'âge de départ à la retraite en relativement longue période. Nous exposerons ensuite les travaux centrés sur la mesure de la capacité à travail.

2.1. Les facteurs du départ en retraite selon Costa (1998)

Dora L. Costa dresse un bilan historique de la retraite en se concentrant sur les Etats-Unis de 1880 à 1990. Dans son livre, « *The Evolution of Retirement : An American Economic History 1880–1990* », elle retrace l'évolution de la retraite, ses facteurs et ses effets sur l'emploi des personnes qui en jouissent, notamment, au sein de la population âgée. Un résumé grossier de son livre constituera l'introduction à notre travail et à ce qu'on pourra attendre des résultats quant à l'importance de l'effet de la santé sur l'emploi.

L'histoire débute en 1880, période à partir de laquelle il nous est possible d'observer le commencement de la diminution continue des taux d'emploi des travailleurs âgés. Le graphique suivant tiré de son livre montre que cette tendance n'est pas spécifique à l'Amérique mais touche également l'ancien continent.

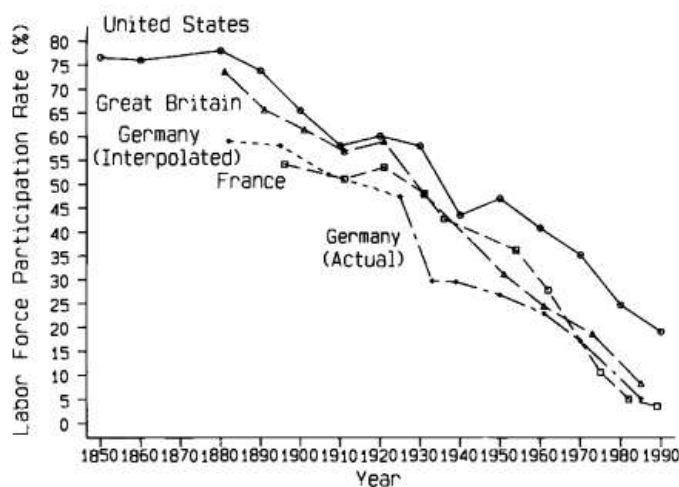


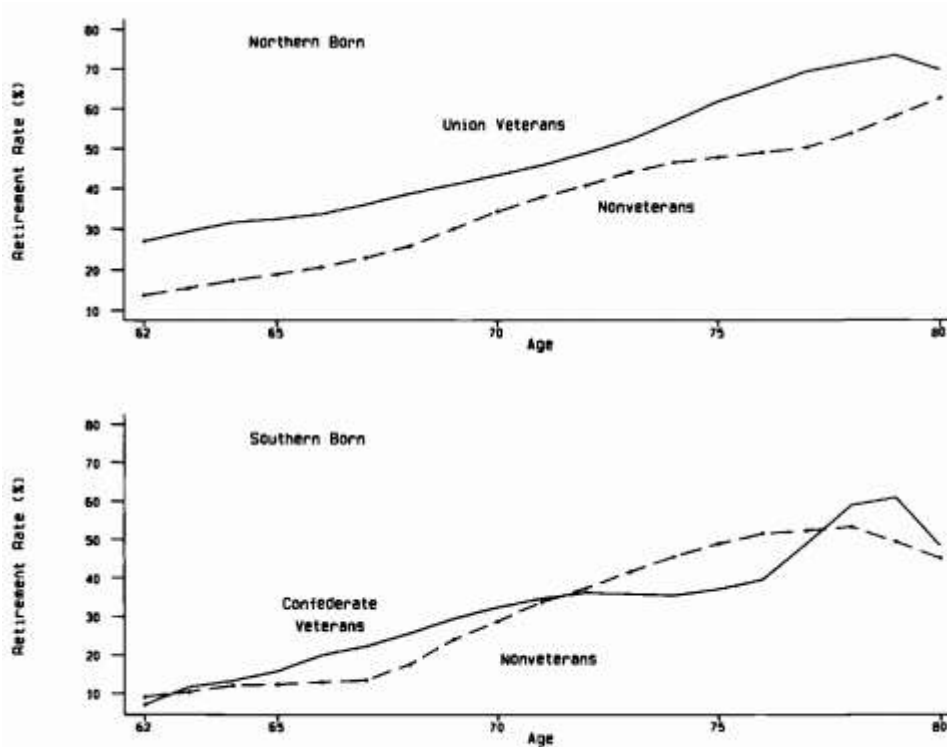
Fig. 2.2 Labor force participation rates of men aged sixty-five and over, 1850–1990, United States, Britain, France, and Germany

p.9 du livre de Costa

En 1883 l'Allemagne de Bismarck associe l'âge de 65 ans comme l'entrée dans l'incapacité à travailler. Cet âge de référence est défini à l'époque comme l'âge à partir duquel un homme commence à voir ses capacités mentales et physiques décliner en entraînant alors son incapacité à travailler. Des études de l'époque et d'aujourd'hui montrent que la productivité des travailleurs âgés décline belle et bien avec l'âge mais d'une façon continue et non pas abrupte comme semble le suggérer cet âge de référence. Costa calcule le taux de probabilité de partir en retraite sans être parti en retraite avant pour chaque âge. Et ce, pour plusieurs périodes s'étalant de 1900 à 1990 ; elle montre que de 1900 à 1920 ce taux croit de façon assez continue mais qu'à partir de la mise en application de l'âge légal de départ en retraite en Amérique à 65 ans, ce taux se met à croître de façon brutale pour les âges proches de l'âge

de 65 ans. Cet argument insinue que la retraite légale incite les individus âgés à quitter le marché du travail parce qu'ils en ont l'occasion.

La raison évoquée à l'époque pour justifier la retraite à 65 ans était que les industries modernes ne voyaient pas d'utilité aux travailleurs âgés et que ceux-ci, sans emploi et sans opportunités d'emploi, ne pouvaient subvenir seuls à leurs besoins. Aujourd'hui, les travailleurs âgés partent en retraite non pas parce qu'ils le doivent mais parce qu'ils le peuvent. Ils jouissent désormais de plus de revenus et ce même sans travailler puisqu'ils reçoivent une retraite. La retraite change les incitations à travailler en offrant des revenus lorsqu'on cesse de travailler. En utilisant la pension accordée aux vétérans de la guerre de sécession comme indicateur d'augmentation des revenus, Costa montre que la plupart des différences de taux d'emploi entre les vétérans de l'Union et le reste de la population s'explique par cette augmentation du revenu.



P.50 du livre de Costa

Fig. 3.6 Retirement rates by veteran status among northern-born and southern-born men aged sixty-two to seventy-nine in 1910

Elle émet également l'hypothèse que ces différences de taux d'emploi puissent provenir de différences de santé due à la guerre. Elle utilise alors les comportements de départ en retraite (graphique ci-dessus) des vétérans de l'union (soldats du Nord, vainqueurs recevant la pension des vétérans) et de ceux des confédérés (soldat du sud, perdants ne recevant pas la pension des vétérans) pour prouver qu'une grande part des différences de comportements de départ en retraite entre les soldats du Nord et du Sud s'explique par les pensions octroyées aux vétérans de l'Union et donc par des différences de revenus et non pas par des différences de santé.

Elle va plus loin en montrant que l'élasticité du départ en retraite avec les revenus décline au cours du temps. Ce dernier résultat est appuyé par les travaux d'autres chercheurs qui avancent comme raison que désormais le choix de départ en retraite n'est plus vraiment lié aux revenus des pensions mais serait plutôt dû soit à un changement des comportements sociétal, soit à l'apparition d'opportunités de loisirs plus variés et moins coûteux qui rentrent en compétition avec le travail.

Costa étudie également la santé des américains au cours du 20^{ème} siècle en comparant la prévalence de différentes maladies entre des américains du début du 20^{ème} siècle avec des données de la fin de celui-ci. Elle montre que d'une manière générale la santé, en termes de maladies chroniques, s'améliore largement durant la période et conclut que la chute des taux d'emploi des personnes âgées modernes intervient **malgré** une amélioration de la santé. Elle utilise ensuite l'indice de masse corporelle comme proxy de la santé (en montrant que l'indice est un bon proxy de la santé) pour représenter ensuite la relation qu'il partage avec le pourcentage de personnes ne travaillant pas. La relation suit une courbe en forme de U pour les américains modernes tout comme pour les vétérans de 1910. Néanmoins la relation entre l'emploi et la santé semble moins importante pour les Américains modernes ; ce qui sous-entend qu'aujourd'hui, contrairement à avant, on n'est plus autant limité par la santé pour travailler.

Costa nous explique aussi qu'en 1880 près de la moitié des personnes âgées vivaient sous le même toit que leurs enfants alors qu'ils ne représentent plus que 5% des individus âgés aujourd'hui. Elle montre que ces changements sont dus à l'augmentation des revenus de ces personnes âgées. Avant, les individus âgés devaient vivre sous la tutelle de leurs enfants parce qu'ils n'avaient pas d'autres choix alors qu'aujourd'hui ils choisissent de vivre séparément. Ce qui montre qu'avant les parents étaient souvent contraints de faire ce choix ; option qu'ils n'auraient pas prise s'ils en avaient eu les moyens financiers.

« Et le pire de tout, mon indépendance est révolue : pour le moment, bien sûr, je devrai vivre avec l'un de mes enfants et je ne sais pas lequel d'entre nous va le détester le plus. »

Livre Costa (p.106), tirée d'Epstein (1992) et traduite par Google trad.

A l'augmentation des revenus viennent s'ajouter une préférence accrue pour les loisirs ; parmi les pensionnés de 1941 à 1951, seuls 3% justifiaient leur départ en retraite pour cette raison. Ce taux passe à 17% en 1963 et à 38%-48% en 1982. Cette augmentation des préférences pour les loisirs ne sont pas le seul fait de l'augmentation des revenus ; aussi parce les loisirs sont désormais bien plus variés et bien moins chers qu'auparavant. Costa nous apprend qu'en 1880 on dépensait environ 2% de ses revenus aux loisirs et près de 75% en nourriture/maison/habits. En 1991, ce premier pourcentage est monté à 5-8% et le deuxième est descendu à 40%. Elle représente également les courbes d'Engels ci-dessous qui se définissent comme l'élasticité du pourcentage de revenus accordé aux loisirs aux couts des loisirs. En 1917, cette élasticité était très forte ce qui symbolise que les loisirs étaient un luxe (il fallait payer cher en % de revenu pour obtenir peu de loisirs en valeur). Aujourd'hui la courbe s'aplatit pour les loisirs chers ce qui symbolise que les loisirs ont perdu de leur luxe (on paye moins cher en % de revenu pour recevoir plus en valeur de loisir).

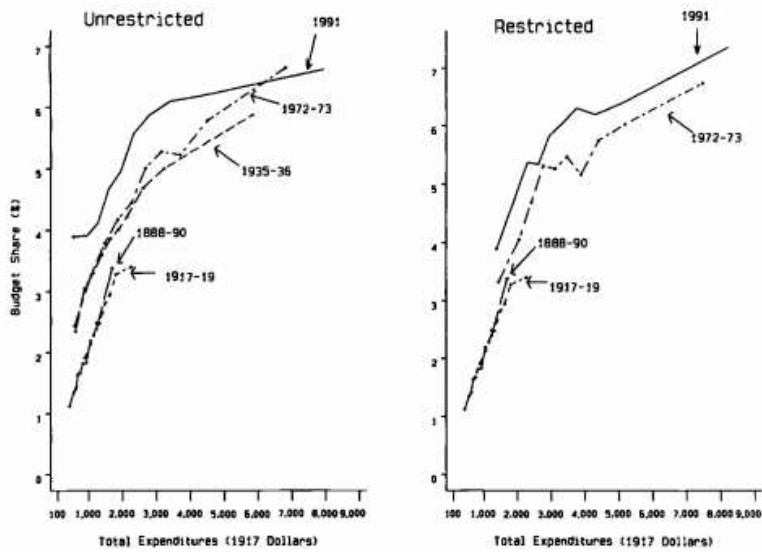


Fig. 7.2 Budget share of recreation and total expenditures in 1917 dollars, 1888-90, 1917-19, 1935-36, 1973-73, and 1991

(p.139 du livre de Costa)

Costa nous apprend ensuite que l'histoire de la pension et celle de la politique (électorale) sont intimement liées ; en 1910, 25% des personnes recevant la pension des anciens vétérans ont plus de 64 ans. Des groupes de pressions font tourner les faveurs politiques vers des accroissements de ces pensions financées par les jeunes, le Sud et les immigrés. Les pensions pour les personnes âgées apparaissent dans les années 1920 et sont financées par les propriétaires ou des personnes n'en jouissant que très peu voir pas du tout.



(p. 163 du livre de Costa)

“En général, l'art du gouvernement consiste à prendre le plus d'argent possible d'une classe de citoyens afin de le donner à une autre.”

Livre de Costa page 160, tirée de Voltaire (1764)

2.2. La capacité à travailler à des âges avancés

Ce mémoire vise à déterminer quelle mesure les individus âgés vivant en Belgique pourraient augmenter leur offre de travail. La réponse à cette question repose sur l'estimation de l'évolution avec l'âge de leur "capacité à travailler". La littérature empirique contient deux méthodes d'estimation de cette capacité ; la méthode de Milligan-Wise (2012) et celle de Cutler et al. (2012). La première consiste à examiner combien de personnes âgées pourraient travailler si elles travaillaient autant que celles ayant le même taux de mortalité que par le passé (méthode Milligan-Wise). La deuxième méthode, que nous utilisons dans la Section 4, consiste à déterminer combien de personnes plus âgées pourraient travailler si elles travaillaient autant que les plus jeunes avec une santé similaire (méthode de Cutler). Bien sûr, d'autres facteurs entrent en compte. Par exemple ; nos compétences ou diplômes restreignent nos opportunités d'emploi et alors, bien qu'on en soit capable au vu de notre santé, on en est incapable dans les faits car le marché du travail n'offre pas les mêmes opportunités d'emploi à tout le monde. Dans ce travail, nous aurons l'occasion d'observer ces différences d'opportunité d'emploi face à l'éducation, le genre, le « type-d'emploi » et les 3 régions principales de Belgique. On verra également comment ces différents facteurs interviennent dans la relation entre l'emploi et la santé.

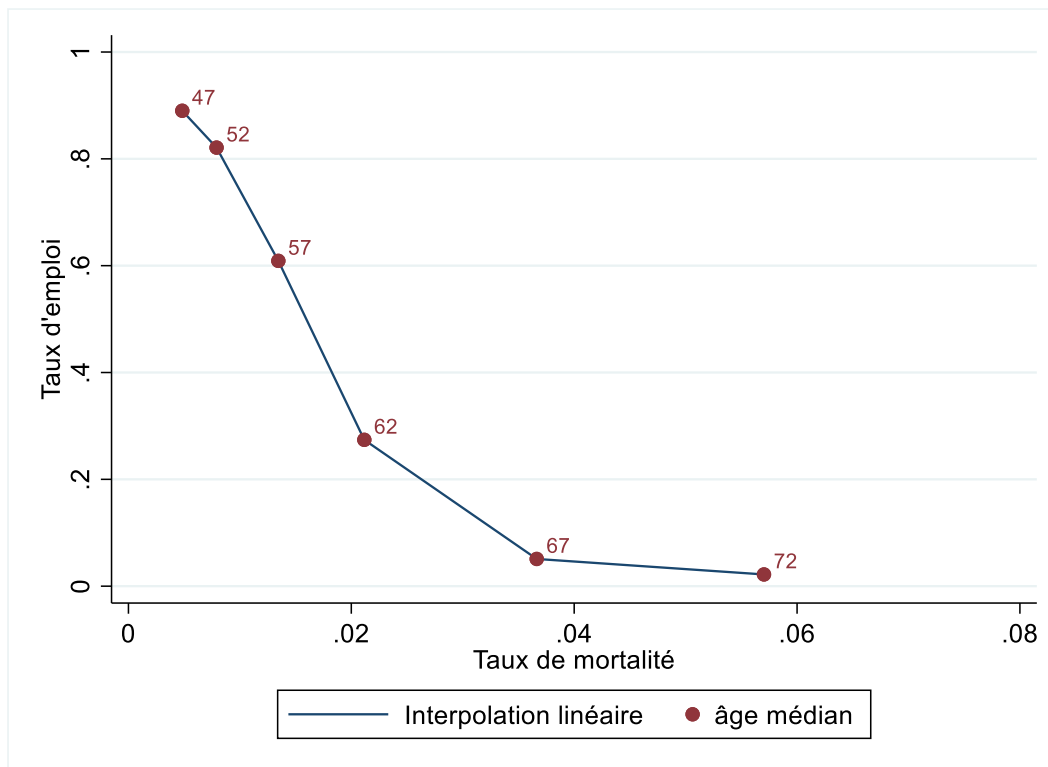
3. La méthode de Milligan-Wise

La méthode de Milligan-Wise fait le lien entre emploi et santé en utilisant le taux de mortalité comme indice de santé. Même s'il est clair que la mortalité constitue un mauvais indicateur de santé, elle présente plusieurs avantages ; d'abord le taux de mortalité laisse moins de doute dans les erreurs de mesure comparé à d'autres mesures de santé plus subtiles pouvant varier entre les périodes de temps, les enquêtes, les pays, ... Ensuite, l'accessibilité d'indicateurs de mortalité est bien plus importante que l'accessibilité d'indicateurs de santé issue d'enquêtes de santé nationales. Finalement, la méthode de Milligan-Wise trouve sa pertinence dans l'accès à une mesure pour un grand nombre de pays et pour un grand nombre de périodes de temps ce qui facilite grandement la comparaison entre les pays (Wise, 2017).

Pour reproduire la méthode de Milligan-Wise nous nous inspirons directement du travail de Jousten et Lefebvre (2017) ; tout comme eux, on combine deux bases des données issues de l'enquête sur les forces de travail européennes (EFT-UE) pour les données sur l'emploi et de la « Human Mortality Database » pour les données sur la mortalité. Malheureusement, ces données sur l'emploi ne sont présentes que par tranches d'âges de 5 ans alors qu'on désirerait en avoir pour chaque âge. On décide donc de combler nos taux d'emploi par âge, tout comme Jousten et Lefebvre l'ont fait, par interpolation linéaire afin d'obtenir un taux d'emploi par âge.

L'unité de tous nos graphiques représente presque toujours des taux pour lesquels 1 représente 100%.

Graphique 3.1 : Taux d'emploi des hommes et interpolation linéaire (1983)



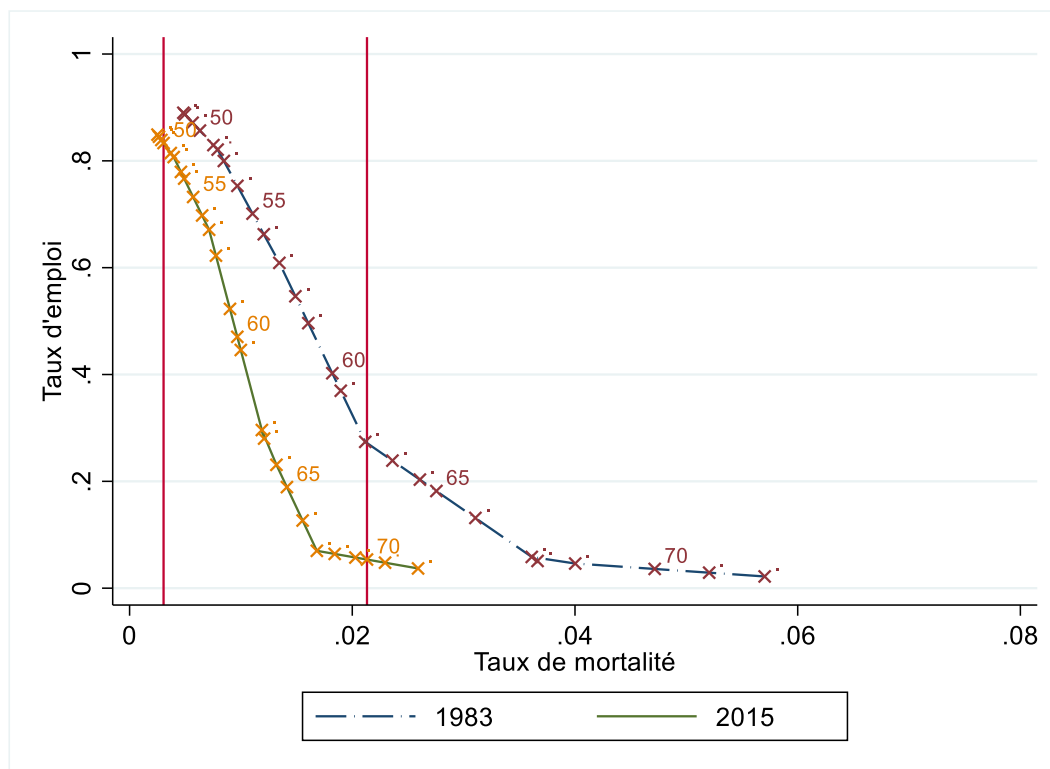
L'opération est décrite par le graphique 3.1 qui présente les taux d'emploi par âges ; chaque point représente une des moyennes de taux d'emploi par tranche d'âge extraite de l'EFT-UE. Les chiffres sur le graphique correspondent à l'âge médian de chaque tranche d'âge auxquelles sont associées leurs moyennes de taux d'emploi respectives. Par exemple, la moyenne du taux d'emploi des personnes de 55-59 ans est associée à l'âge 57. Et les taux d'emploi pour les âges situés entre deux points correspondent à la droite séparant ces deux points.

Désormais, nous disposons de données sur l'emploi et la mortalité par âge et nous sommes prêts à démarrer l'application de la méthode de Milligan-Wise.

Le graphique ci-dessous représente la relation entre l'emploi et la mortalité pour les années 1983 et 2015. On peut observer la relation négative entre l'emploi et la mortalité. Aussi, pour tout taux de mortalité donné, le taux d'emploi moyen par âge de 2015 est inférieur à celui de 1983 ce qui signifie que les personnes de 1983 travaillaient plus pour une même « santé ».

Dans ce travail nous nous intéressons à la capacité à travailler des individus âgés uniquement ; on ne calculera donc la capacité à travailler de 2015 que pour les individus se situant entre les deux droites verticales présentées sur le graphique, c'est-à-dire, ceux de 50 à 70 ans.

Graphique 3.2 : Taux d'emploi contre taux de mortalité des hommes (1983 contre 2012)



La capacité à travailler au sens de Milligan-Wise est le taux d'emploi que la génération présente (d'intérêt) fournirait si elle possédait la même relation emploi/mortalité que la génération passée (de référence). Par exemple, en 2015, le taux de mortalité des hommes de 70 ans était de 2.1% ce qui correspondait approximativement au taux de mortalité des hommes de 61 ans en 1983 ; On estimera donc une capacité à travailler des hommes de 70-74 ans égale à l'intersection de la deuxième droite pointillée verticale avec la courbe emploi/mortalité de 1983. Avec la méthode de Milligan-Wise, on assume que si les gens de 1983 étaient capables de travailler plus pour un même taux de mortalité que ceux de 2015 alors les gens de 2015 sont capables de travailler plus et d'exactement la différence verticale séparant les 2 courbes. Ici, la distance séparant les personnes de 70 ans de 2015 de ceux de même santé de 1983 est légèrement supérieure à 20%. Ces 20% représentent donc la capacité à travailler additionnelle (où inexploitée) des personnes de 70 ans en 2015.

On calcule ensuite cette différence pour chaque âge de 50 à 70 ans qu'on additionnera pour obtenir le nombre d'années de travail inexploité.

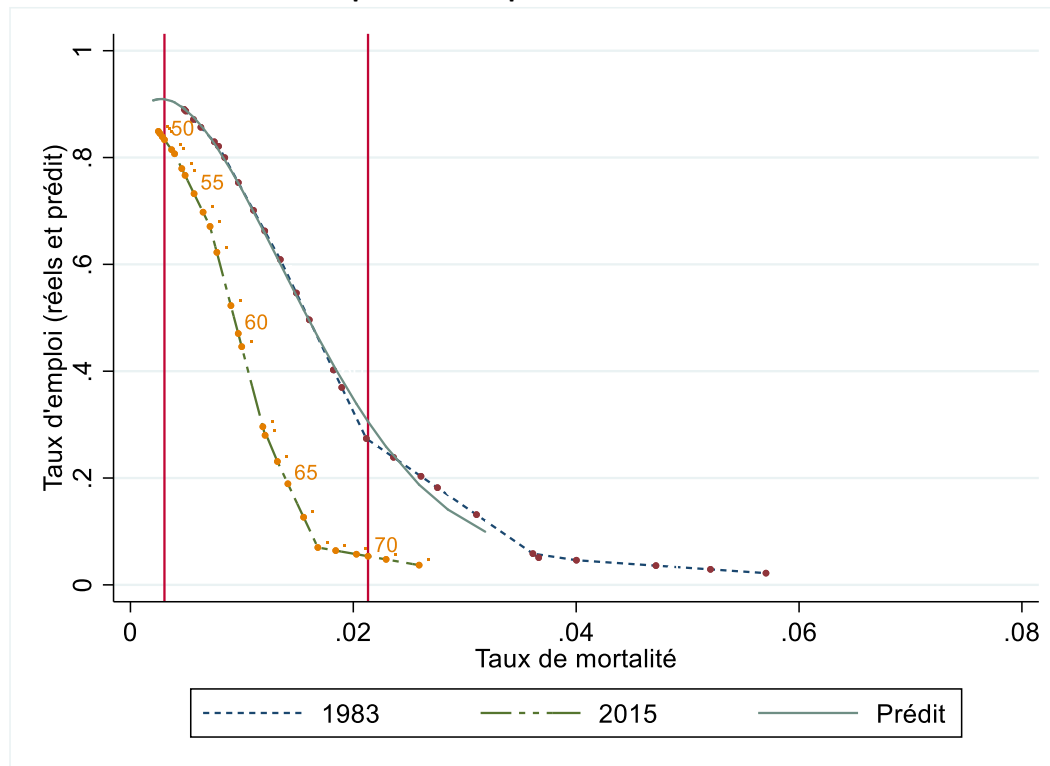
On répètera l'exercice pour chaque année de 1983 à 2015 pour observer l'évolution de ce nombre d'années de travail inexploité au cours du temps.

Un dernier problème s'oppose à nous avant de pouvoir nous lancer ; les taux de mortalité d'une année ne correspondent pas nécessairement aux taux de mortalité d'une autre année. On régresse donc l'emploi de 1983 sur un polynôme¹ de la mortalité pour pouvoir ensuite

¹ Solution offerte par Scott Merryman sur le forum de Stata, discussion disponible sur <https://www.statalist.org/forums/forum/general-stata-discussion/general/1492839-compute-vertical-distance-between-two-line-in-a-twoway-line-graph-seems-easy-not-for-me>

obtenir (prédire) les taux d'emploi de 1983 pour les mêmes taux de mortalité que les personnes de 2015. Ces taux d'emploi modélisés ne sont pas exactement similaires mais restent très proches des taux d'emploi réels comme le montre le graphique suivant. Aussi, la mortalité tel qu'utilisé ici permet d'expliquer plus de 99% des variations de l'emploi dans chacune de nos régressions servant à prédire l'emploi d'une année aux taux de mortalité de 2015.

Graphique 3.3 : Taux d'emploi contre taux de mortalité des hommes (1983 contre 2015) et taux d'emploi de 1983 prédits aux taux de mortalité de 2015



Enfin, le tableau 3.1 présente les taux d'emploi et de mortalité de 2015 par âges auxquels sont associés les taux d'emploi de 1983 aux mêmes taux de mortalité. On présente aussi la différence des taux des deux périodes qui représente l'emploi inexploité pour un âge donné.

La dernière colonne est la somme cumulée de ces différences de taux d'emploi par âge et représente alors le nombre d'années de travail inexploité pour un âge donné ; en prenant 1983 comme année de référence, on calcule 5.164 années de travail inexploitées à 70 ans pour les gens de 2015.

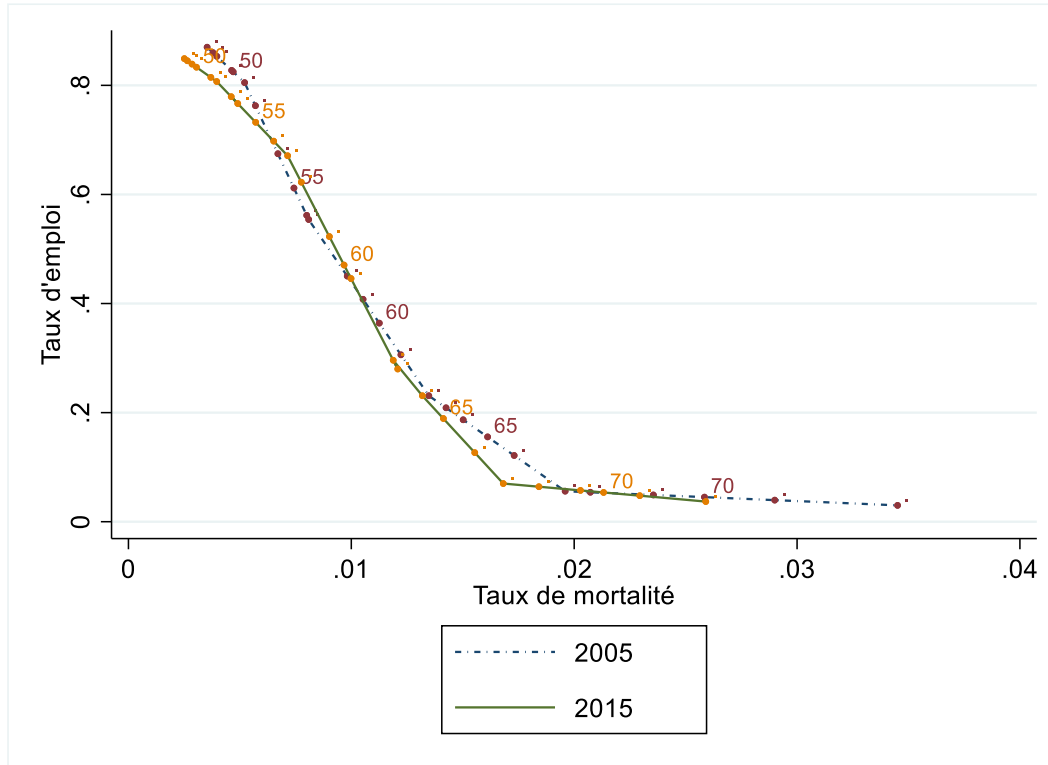
Tableau 3.1 : Taux d'emploi, capacité à travailler additionnelle et nombre d'années de travail inexploitées (1983 contre 2015)

Age	Taux d'emploi 2015	Taux de mortalité	Taux d'emploi 1983	Différence de taux d'emploi par âge entre 1983 et 2015	Nombre d'années de travail inexploitées
50	0.833	0.003	0.909	0.076	0.076
51	0.815	0.004	0.906	0.091	0.167
52	0.807	0.004	0.903	0.096	0.263
53	0.767	0.005	0.890	0.124	0.387
54	0.779	0.005	0.895	0.116	0.503
55	0.732	0.006	0.875	0.142	0.645
56	0.698	0.007	0.855	0.157	0.803
57	0.671	0.007	0.838	0.167	0.970
58	0.622	0.008	0.819	0.197	1.166
59	0.523	0.009	0.777	0.254	1.420
60	0.471	0.010	0.752	0.282	1.702
61	0.446	0.010	0.741	0.295	1.997
62	.28	0.012	0.657	0.377	2.374
63	0.296	0.012	0.666	0.369	2.743
64	0.231	0.013	0.612	0.381	3.124
65	0.189	0.014	0.573	0.384	3.507
66	0.127	0.016	0.515	0.388	3.896
67	.07	0.017	0.464	0.394	4.290
68	0.064	0.018	0.404	0.340	4.630
69	0.057	0.020	0.339	0.282	4.912
70	0.054	0.021	0.306	0.252	5.164

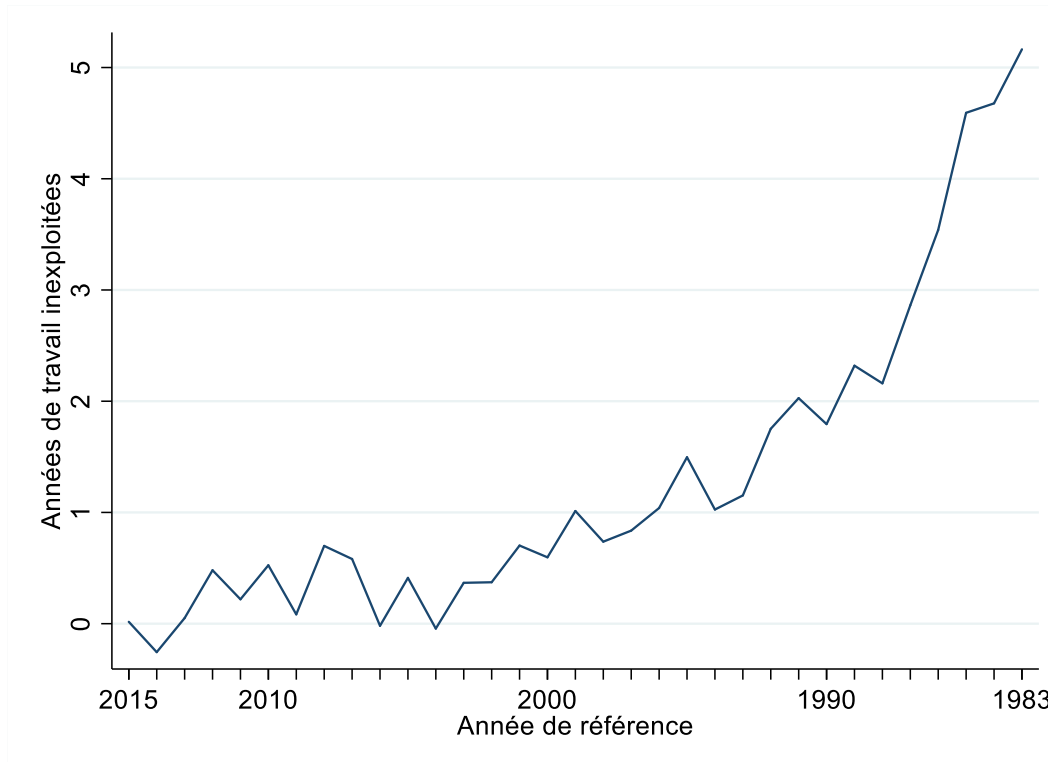
Dans ce cas-ci, en prenant 1983 comme année de référence, on trouve donc une capacité à travailler additionnelle pour les Belges de 50 à 70 ans correspondant à 5.164 années de travail ; qu'en serait-il l'on désirait prendre une autre année de référence ? Faisons l'essai en prenant 2005 comme année de référence ; on peut voir dans le graphique 3.3 que la différence de relation emploi/mortalité n'est plus aussi grande que lorsqu'on prenait 1983 comme année de référence. Ce constat s'explique par l'amélioration importante de la santé que la Belgique a connue au cours des dernières décennies.

Nous faisons ensuite l'exercice pour toutes les années de référence de 1983 à 2015 en conservant toujours 2015 comme année d'intérêt. Le dernier graphique de cette partie indique le nombre d'années inexploitées à 70 ans (de 50 à 70) selon l'année de référence choisie ; on le voit clairement dans le graphique 3.4, le nombre d'années de travail inexploitées qu'on estimera dépend directement de l'année de référence choisie. Plus on remonte dans le temps, plus les gens travaillaient avec de hauts taux de mortalité ce qui explique les résultats positifs en terme de capacités de travail inexploitées pour l'année 2015 lorsqu'on choisit 1983 comme année de référence.

Graphique 3.4 : Taux d'emploi contre taux de mortalité des hommes (2005 contre 2015)



Graphique 3.5 : Nombre d'années de travail inexploitées par année de référence



Enfin, Jousten A. (2017) nous informe que pour bien comprendre le sens du nombre d'années de travail additionnel ; 3 éléments doivent être considérés :

(i) l'équivalence entre la capacité à travailler additionnelle par âge et le nombre d'années de travail additionnel implique l'hypothèse que ces travailleurs additionnels travailleraient le même nombre d'heures, jours et mois que ceux qui travaillent déjà.

(ii) La structure de l'emploi/santé de l'année de référence impactera inévitablement les résultats. Et c'est la raison pour laquelle la capacité à travailler calculée via la méthode de Milligan-Wise se fait toujours en comparant les résultats obtenus à l'aide de différentes années de références.

(iii) le nombre d'années de travail supplémentaires doit être compris comme une construction théorique et interprétée de la même manière, à savoir comme la somme des taux d'emploi additionnel de chaque âge ; la mesure en années ignore les différences de taille et de composition entre chaque âge. Aussi, l'interprétation est compliquée et se fera surtout par comparaison entre les périodes de la mesure en années de travail inexploitées.

Pour conclure, la méthode de Milligan-Wise apporte une manière simple pour estimer et comprendre la capacité à travailler. Le nombre de variables nécessaires pour obtenir des résultats est très restreint (années, âge, emploi et mortalité). Néanmoins, cette simplicité entraîne des limites claires ;

- (i) L'utilisation de la mortalité comme proxys de la santé est très discutable.
- (ii) Les résultats finaux dépendent entièrement des différences de structure emploi/mortalité de l'année de référence avec celle d'intérêt. Par exemple, si on ne dispose pas d'année de référence avec une structure emploi/mortalité suffisamment différente de l'année d'intérêt, on ne sera pas en mesure de calculer une capacité à travailler additionnelle satisfaisante ; en prenant 2005 comme année de référence, on calcule un nombre d'années de travail supplémentaires en 2015 quasi nul. Aussi, le choix d'une période de référence pour déterminer la capacité à travailler est finalement arbitraire et c'est pourquoi on reporte les résultats obtenus via différentes années de références.

Enfin, la méthode de Milligan-Wise rend la notion de capacité à travailler très simple à comprendre. Elle est présentée ici comme une introduction à la méthode principale qui sera appliquée dans ce travail ; la méthode de Cutler. Les deux méthodes partagent le but ; estimer la « capacité à travailler ». Egalement le principe ; observer les différences de taux d'emploi d'individus pour une santé similaire.

4. La méthode de Cutler

Cutler et al. (2012) nous présentent une stratégie différente pour prédire « la capacité à travailler » d'une population. La population de référence n'est plus une génération du passé, comme lors de la méthode de Milligan-Wise, mais les individus de la génération présente qui ne sont pas encore sujet à un éventuel départ en retraite. Ces individus sont les personnes de 50-54 ans pour lesquels, on émet l'hypothèse (lors de la prédiction de leur capacité à travailler) que s'ils ne travaillent pas ; c'est pour des raisons de santé uniquement. Cette tranche d'âge n'est bien sûr pas prise au hasard ; d'abord, ces individus sont proches de l'âge des personnes sujettes à la retraite et devraient donc adapter leur emploi à leur santé d'une façon proche de celle des personnes pour lesquelles on désire quantifier la capacité à travailler. Ensuite, Johansson (2017) nous apprend : « *que les travailleurs de cette tranche d'âge utilisent presque exclusivement le programme d'assurance incapacité à travailler et l'assurance maladie pour quitter le marché de l'emploi.* »²

La méthode de Cutler se déroule en deux étapes : la première consiste à estimer la relation (1) entre l'emploi et la santé pour les individus de 50-54 ans. E_i est une variable binaire indiquant si l'individu i travaille au moment de l'enquête. H_i représente la santé de l'individu i et X_i regroupe les autres variables de contrôle de l'individu i .

$$E_i^{50-54} = \alpha + \beta * H_i^{50-54} + \gamma * X_i^{50-54} + e_i \quad (1)$$

La deuxième étape utilise la relation (1) entre l'emploi et la santé estimée précédemment pour prédire l'emploi (\hat{E}_i^a) des personnes plus âgées et donc, leur capacité à travailler.

$$\hat{E}_i^a = \hat{\alpha}^{50-54} + \hat{\beta}_i^{50-54} * H_i^a + \hat{\gamma}^{50-54} * X_i^a \quad (2)$$

La différence entre la capacité à travailler prédite et le taux d'emploi effectif représente alors la capacité à travailler additionnelle qu'il est envisageable d'atteindre pour la population étudiée. Si celle-ci est positive alors une augmentation de l'âge légal de départ en retraite se justifie par le fait que la santé, à elle seule, ne permet pas de déterminer la chute de l'emploi d'un point de vue macro.

En estimant la relation entre emploi et santé pour des individus de 50-54 ans et en l'utilisant pour prédire l'emploi des personnes de 55 ans et plus, on fait l'hypothèse implicite que la relation entre l'emploi et la santé ne dépend pas de l'âge ; l'effet d'un problème de santé particulier sur la capacité à travailler sera le même pour chaque tranche d'âge. Par exemple, si un cancer ne justifie un arrêt de travail que dans 40% des cas pour les 50-54 ans, il en fera de même pour les 65-69 ans. C'est la raison à l'utilisation d'individus d'âges proches de ceux pour lesquels on désire calculer la capacité à travailler.

Une seconde hypothèse est que le travail, en soi, n'a pas d'effet sur la santé : *la causalité inversée*. Puisque nous observons des individus presque tous au travail pour inférer une potentielle capacité à travailler sur une population qui ne travaille pas ; on suppose alors que la santé évolue indépendamment de la situation professionnelle (au travail ou pas) d'un individu. Par exemple, imaginons que le travail mette notre santé psychologique à rude

² P.314 du livre de Wise ou p.14 du travail de Johansson

épreuve et que plus le nombre d'années de travail avance plus nous nous approchons du « burnout ». Alors l'arrêt du travail (car départ en retraite) signifierait un gain de santé psychologique et donc une capacité à travailler prédite supérieure à ce qu'elle devrait être pour les personnes âgées qui ne travaillent plus (et qui ont donc une meilleure santé psychologique).

Blanks et al. (2017) Synthétise parfaitement l'idée et l'état actuel de la question dans son travail sur la capacité à travailler aux Royaume-Unis ;

« ...nous évoquons la possibilité que le travail et / ou la retraite puissent avoir un effet causal direct sur la santé des individus. C'est une préoccupation potentiellement importante et nos résultats doivent être interprétés en tenant compte de cette mise en garde. Une multitude de raisons théoriques ont été postulées pour expliquer pourquoi le travail, et donc les sorties du travail, pourraient affecter directement la santé et le bien-être physique et / ou mental. Ces raisons théoriques, qui sont examinées plus en détail par Banks, Chandola et Matthews (2015), suggèrent qu'il pourrait exister une hétérogénéité importante dans la manière dont le travail (et les sorties du travail) affectent la santé, le travail pouvant potentiellement avoir des effets positifs et parfois négatifs. Les données empiriques sur les effets de la retraite sur la santé sont mitigées. Un certain nombre d'études ont montré de fortes corrélations entre les sorties du travail et la santé, mais il existe une hétérogénéité lorsqu'il s'agit de déterminer si le fait de se retirer du travail est associé à une meilleure ou une pire santé. Des analyses causales plus robustes sont plus rares dans la littérature et révèlent des effets plus faibles (en particulier sur la santé physique) et des résultats mitigés. »³

Lorsqu'on estime la relation entre santé et emploi via la méthode de Cutler, on utilise un nombre d'indicateurs de santé assez important ; cancer, arthrite, épilepsie, handicap, ... Dès lors, comparer la santé de deux individus (de santé similaire) sera fastidieux, voire impossible. Une méthode développée par Poterba J. (2013) propose une solution pour synthétiser l'ensemble des variables de santé en une seule (le PVW index) ; nous donnant alors un score de santé unique pour chaque individu. L'index se base sur un ensemble d'indicateurs de santé pour lesquels on calcule un score sur base de la première composante principale. Chaque coefficient associé à une maladie particulière est positif signifiant qu'une valeur plus grande de la composante principale première représente une moins bonne santé. On convertit ensuite ce score en centile tout en l'inversant nous donnant le PVW index. Dès lors, une valeur plus grande de cet index représente une meilleure santé et on pourra interpréter le coefficient associé à notre PVW index comme le changement dans le pourcentage moyen d'être à l'emploi suite à l'augmentation d'un centile de l'index de santé.

D'une manière générale, la santé décline avec l'âge. Lorsque nous sommes jeunes nous sommes en assez bonne santé et possédons une forte capacité à travailler. En vieillissant, notre santé décline et certaines personnes ne sont plus capables de travailler parce que leur santé ne le leur permet plus. La méthode consiste finalement à observer quels sont les « soucis de santé » qui déterminent « l'incapacité à travailler » chez une population jeune pour prédire « l'incapacité à travailler » des populations plus âgées avec leur propre santé.

³ Livre de Wise p.332-333 (Traduction par google trad.)

En pratique, la méthode consiste à régresser une variable binaire de l'emploi (1 signifie que l'individu travaille et 0 qu'il ne travaille pas) sur un ensemble de variables de santé objectives (cancer, allergie, haute pression sanguine, épilepsie...) et subjectives (santé subjective, limitations dans les activités quotidiennes...) ainsi que sur des caractéristiques démographiques (éducation, type de travail, année de l'interview...). Avec l'aide du PVW index, on peut régresser l'emploi sur une seule variable de santé et les caractéristiques démographiques réduisant fortement le nombre de coefficients à estimer. Aussi, on pourra directement interpréter l'effet de posséder une santé meilleure en centile.

4.1. Données

Nous disposons de données issues de l'enquête de santé belge qui collecte des informations sur la santé d'individus Belges tous les 4-5 ans (1997, 2001, 2004, 2008 et 2013). Les informations sont collectées en face-à-face directement domicile de la personne interrogée.

Le Registre National de la Population est utilisé comme base de sondage et offre les informations sur l'âge et l'adresse (la région). La méthode d'échantillonnage se déroule en plusieurs étapes ; d'abord, une stratification géographique (par région). Une sélection d'un nombre prédéfini de ménages (cluster) dans chaque région et une sélection d'individus dans chaque ménage.

Nous utilisons ici une version simplifiée de l'enquête nationale de santé en nous concentrant sur un nombre limité de variables de santé ainsi que sur l'utilisation de données individuelles (une observation représente une personne). Omettre la méthode d'échantillonnage dans l'analyse statistique conduit à des estimateurs biaisés. Néanmoins, selon le manuel de l'enquête de santé belge (Scientific Institute of Public Health, 2013), comme le nombre de clusters est grand et comme le nombre d'individus sélectionnés par ménage est petit (max 4) l'effet du plan de sondage (stratification et cluster) sur nos estimateurs reste modéré. Par exemple, selon une pour la variable de *santé subjective* ; La prise en compte des poids, du clustering et de la stratification conduit à des estimateurs 3,17% plus petit (donc très proches). Tandis que l'écart type augmente en termes relatifs de 42% (Stefaan Demarest 2013).

Les tableaux 4.1.a et 4.1.b résument la moyenne (= le % de personnes touchées) de chaque indicateur de santé par tranche d'âge respectivement pour les hommes et pour les femmes et présente également l'emploi moyen. On peut observer que l'emploi des femmes passe de 61.6% à 50-54 ans pour chuter à 4.1% pour les 65-69 ans tandis que l'emploi des hommes passe de 80.1% à 50-54 ans pour tomber à 6.3% à 65-69 ans.

Tableau 4.1.a : Statistiques descriptives des femmes

	50-54 ans	55-59 ans	60-64 ans	65-69 ans	70-74 ans	75-79 ans	80-84 ans	85+ ans
Taux d'emploi	.616	.438	.145	.041	.025	.019	.024	.01
<i>Santé Mentale</i>								
Moyenne GHQ-12 score	2.035	1.627	1.471	1.585	1.594	1.61	1.441	1.748
Détresse psychologique	.344	.308	.28	.317	.294	.321	.288	.319
Troubles mentaux	.227	.179	.168	.163	.171	.173	.14	.183
<i>Santé Subjective et Limitations</i>								
Santé subjective	2.217	2.22	2.248	2.403	2.454	2.516	2.528	2.641
Handicap en mobilité	.027	.039	.054	.086	.114	.165	.206	.384
Sévérité des limitations	.454	.489	.481	.633	.697	.794	.871	.864
Limitations	.246	.263	.277	.341	.388	.426	.443	.418
Limitation sévère	.06	.06	.051	.076	.073	.09	.116	.119
<i>Santé Objective</i>								
Haute pression sanguine	.195	.24	.302	.331	.36	.368	.383	.337
AVC et conséquences	.004	.007	.009	.011	.013	.024	.034	.037
Polyarthrite rhumatoïde	.104	.114	.134	.192	.21	.261	.237	.255
Arthrose	.202	.266	.324	.431	.444	.46	.485	.5
Diabète	.033	.048	.054	.07	.114	.128	.113	.099
Allergie	.157	.154	.12	.149	.136	.133	.135	.131
Ulcère de l'estomac	.042	.061	.052	.054	.06	.06	.058	.074
Cirrhose du foie	.008	.013	.017	.018	.015	.017	.013	.012
Cancer	.034	.031	.041	.033	.052	.032	.04	.03
Maux de tête sévères	.177	.133	.114	.116	.103	.092	.103	.099
Morosité ou dépression	.104	.104	.105	.102	.09	.09	.066	.084
Problèmes de thyroïde	.096	.118	.137	.153	.105	.107	.095	.089
Glaucome	.01	.012	.022	.059	.121	.152	.19	.24
Cataracte	.017	.03	.043	.046	.061	.056	.066	.064
Maladie de Parkinson	.001	.001	.005	.01	.01	.015	.018	.027
Épilepsie	.003	.005	.007	.005	.004	.004	.008	.002
L'ostéoporose	.07	.109	.147	.188	.207	.212	.237	.203
Hanche cassée	0	.002	.006	.007	.015	.021	.026	.03
Trouble de l'intestin	.035	.044	.049	.073	.075	.079	.063	.099
Des pierres dans le rein	.009	.007	.007	.011	.01	.015	.013	.007
Maladie du rein	.009	.014	.007	.01	.013	.009	.024	.01
Cystite chronique	.024	.02	.029	.025	.048	.054	.05	.057
Maladie de peau	.043	.04	.027	.037	.031	.034	.047	.032
Inflammation vés. bilia.	.019	.012	.016	.031	.028	.034	.026	.027
# of obs	1189	1079	822	812	667	467	379	404

Note : les données par vagues sont regroupées

Tableau 4.1.b : Statistiques descriptives des hommes

	50-54 ans	55-59 ans	60-64 ans	65-69 ans	70-74 ans	75-79 ans	80-84 ans	85+ ans
Taux d'emploi	.801	.58	.216	.063	.033	.019	.017	.006
<i>Santé Mentale</i>								
Moyenne GHQ-12 score	1.336	1.249	1.163	.85	1.14	1.242	1.344	1.585
Détresse psychologique	.25	.236	.217	.166	.205	.239	.269	.307
Troubles mentaux	.139	.122	.115	.082	.115	.126	.168	.166
<i>Santé Subjective et Limitations</i>								
Santé subjective	2.086	2.167	2.259	2.29	2.39	2.497	2.503	2.53
Handicap en mobilité	.021	.031	.037	.048	.077	.128	.15	.246
Sévérité des limitations	.384	.477	.534	.52	.632	.758	.792	.805
Limitations	.209	.268	.301	.297	.344	.412	.439	.403
Limitation sévère	.045	.041	.05	.055	.068	.071	.072	.093
<i>Santé Objective</i>								
Haute pression sanguine	.171	.225	.261	.275	.31	.315	.26	.249
AVC et conséquences	.007	.01	.013	.016	.015	.032	.029	.035
Polyarthrite rhumatoïde	.064	.075	.096	.111	.108	.14	.147	.112
Arthrose	.128	.159	.229	.233	.25	.29	.303	.294
Diabète	.046	.061	.096	.106	.128	.13	.107	.093
Allergie	.1	.095	.107	.094	.073	.077	.078	.051
Ulcère de l'estomac	.056	.054	.047	.057	.068	.049	.058	.054
Cirrhose du foie	.011	.013	.007	.019	.007	.01	.009	0
Cancer	.011	.018	.039	.041	.048	.051	.035	.035
Maux de tête sévères	.071	.061	.049	.049	.036	.042	.043	.042
Morosité ou dépression	.068	.062	.065	.04	.055	.051	.052	.032
Problèmes de thyroïde	.026	.03	.025	.033	.024	.042	.023	.016
Glaucome	.006	.013	.018	.044	.077	.103	.124	.153
Cataracte	.015	.019	.028	.023	.045	.056	.043	.058
Maladie de Parkinson	.001	.001	.002	.007	.015	.017	.017	.051
Épilepsie	.005	.008	.003	.005	.008	.003	.006	.006
L'ostéoporose	.011	.019	.025	.023	.02	.039	.04	.07
Hanche cassée	.002	.003	.007	.009	.005	.008	.012	.022
Trouble de l'intestin	.034	.026	.039	.033	.047	.056	.055	.08
Des pierres dans le rein	.012	.013	.012	.019	.025	.012	.003	.006
Maladie du rein	.006	.003	.009	.016	.012	.015	.006	.016
Cystite chronique	.006	.009	.019	.024	.041	.034	.032	.032
Maladie de peau	.032	.026	.035	.039	.033	.045	.043	.045
Inflammation vés. bilia.	.004	.005	.011	.013	.013	.019	.006	.01
Problème de prostate	.024	.05	.092	.146	.169	.162	.243	.23
# of obs	1234	1173	895	821	751	594	346	313

Note : les données par vagues sont regroupées

Les variables de santé sont ici regroupées en trois catégories distinctes : santé subjective, santé objective et santé mentale. Les variables de santé subjective contiennent notamment l'état de santé déclaré ou la sévérité des limitations dues à une maladie ou à un handicap. Les variables de santé mentale sont au nombre de trois avec un score général (GHQ-12 score), un indicateur de détresse psychologique et un dernier de trouble mental. Nous possédons 24 variables de santé objective pour les femmes et 25 pour les hommes puisque pour ceux-ci nous possédons également un indicateur de problème à la prostate.

Les variables de santé indiquent d'une manière générale si un individu est touché par une maladie ou un problème de santé particulier ; une plus grande valeur de ces indicateurs indique donc une moins bonne santé. On peut remarquer que, d'une manière générale, la prévalence d'une maladie particulière s'accroît avec l'âge ; par exemple, 19.5% des femmes de 50-54 ans souffrent d'une pression sanguine élevée contre 33.7% pour celles de plus de 85 ans. L'ostéoporose touche 7% des femmes de 50-54 ans contre 20.3% à 85 ans, ... Etrangement, la « mauvaise santé psychologique » semble suivre une tendance à la baisse, jusqu'à 60-64 ans pour les femmes et 65-69 ans pour les hommes, pour ensuite augmenter progressivement.

4.2. La santé subjective

La santé subjective (ou ressentie) d'un individu peut prendre des dimensions potentiellement infinies. Qu'elle soit physique ou mentale, celle-ci dépend aussi de caractéristiques, démographiques et autres, indénombrables. On pourra par exemple citer le genre de l'individu, son ethnicité, sa vraie santé physique, son état mental général et aussi spécifique au moment de réaliser l'interview servant à enregistrer sa santé subjective... Néanmoins, elle se révèle être un excellent indicateur de santé général puisque celle-ci est un bon prédicteur des indicateurs de santé généraux, et souvent ceux plus spécifiques, comme la mortalité qui, pour rappel, expliquait plus de 99% des variations de l'emploi dans notre partie sur la méthode de Milligan-Wise.

4.2.1. Définition binaire de la (mauvaise) santé

La santé subjective telle que mesurée dans nos données peut prendre 5 valeurs différentes (très pauvre, pauvre, moyenne, bonne et très bonne. Toutes exprimées en variables dichotomiques). Dans un souci de simplicité, consacrons-nous d'abord sur une variable binaire unique indiquant si oui (=1) ou non (=0) l'individu se trouve dans une santé moyenne, pauvre ou très pauvre (= « mauvaise santé »).

Le graphique précédent présente le % d'individu en « mauvaise santé ». Approximativement 25-30 % des personnes de 50-55 ans se disent être en « mauvaise santé ». Ce pourcentage culmine à 80% pour les personnes de 95 ans. Bien sûr, passé 90 ans, si pas déjà avant, on pourra difficilement compter sur nos prédictions étant donné le nombre très restreint d'individus très âgés dans nos données.

Graphique 4.2.1.a : Personnes en moyenne, pauvre et très pauvre santé par âge

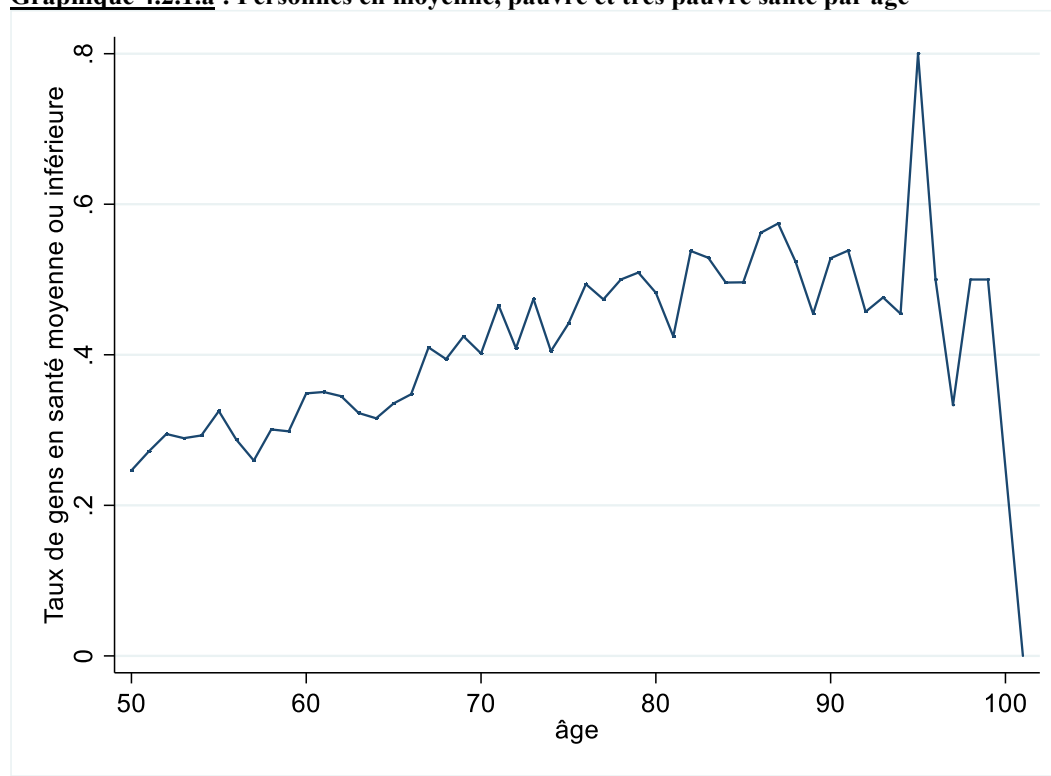


Tableau 4.2.1.a : Régression linéaire de l'emploi des 50-54 ans sur la santé subjective binaire

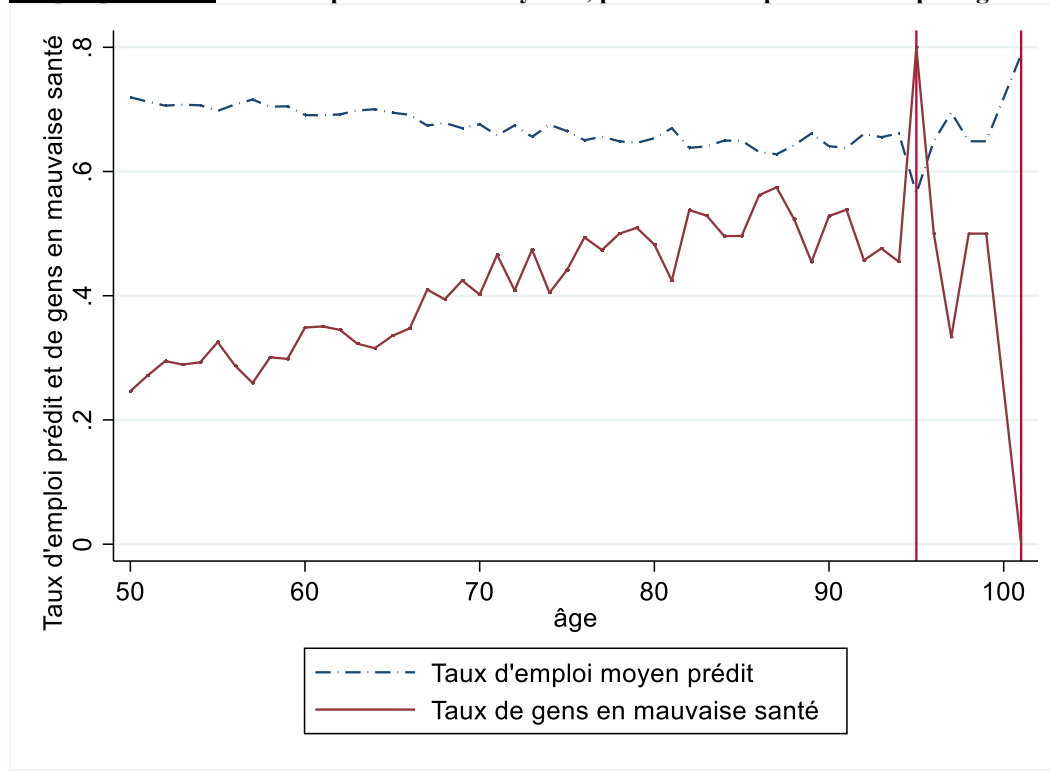
	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Santé moyenne à très pauvre	-0.280	0.022	-12.94	0.000	-0.322	-0.237	***
Constante	0.788	0.010	80.70	0.000	0.769	0.808	***
Mean dependent var	0.711		SD dependent var	0.454			
R-squared	0.076		Number of obs	2423.000			
F-test	200.025		Prob > F	0.000			
Akaike crit. (AIC)	2855.189		Bayesian crit. (BIC)	2866.774			

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

L'estimateur associé à la régression OLS nous indique que pour les personnes de 50-54 ans, augmenter le nombre de personnes en « mauvaise santé » de 0 (personne) à 100% (Tout le monde) entrainerait une diminution de l'emploi de 28 points de pourcentage.

Enfin, observons l'évolution de l'emploi prédit selon la méthode de Cutler avec l'évolution du pourcentage de personne en « mauvaise santé ». Comme l'emploi est prédit avec uniquement cette variable, l'emploi prédit moyen par âge évoluera en sens contraire et à hauteur de 28% de l'évolution du % de personnes en « mauvaise santé » par âge.

Graphique 4.2.1.b: Taux de personnes en moyenne, pauvre et très pauvre santé par âge et taux d'emploi



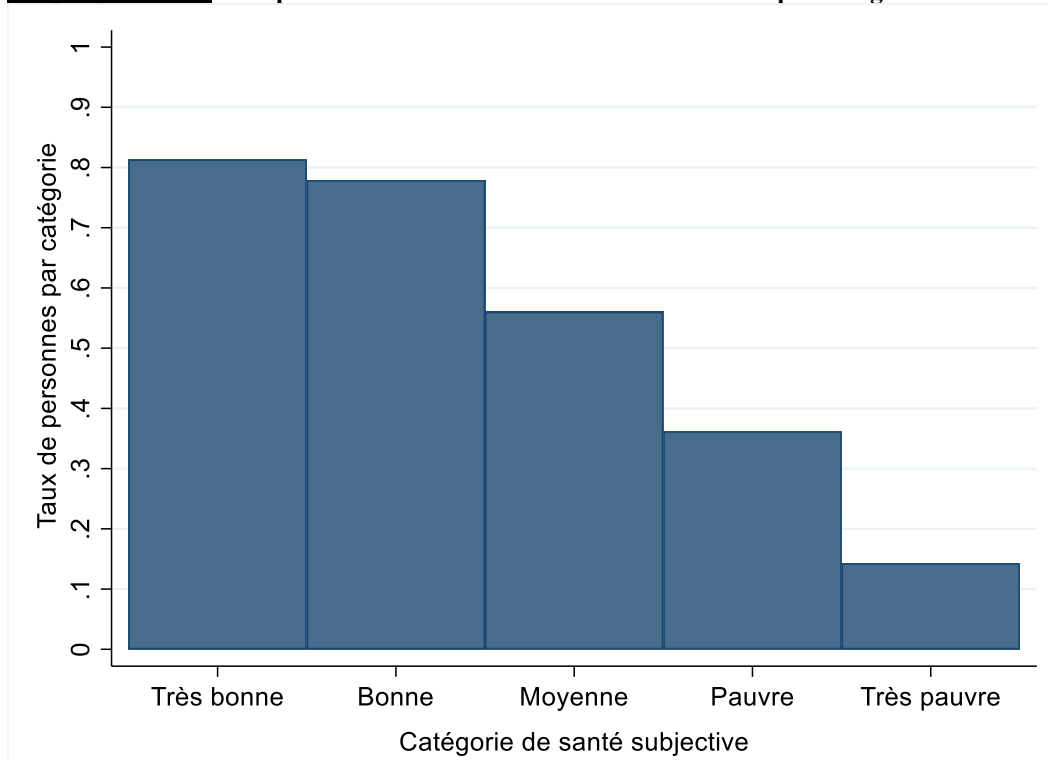
Par exemple, dans le tableau-ci-dessus, on voit que parmi nos personnes de 101 ans, aucune n'est en mauvaise santé (=0% en moyenne à 101 ans). Tandis que chez nos individus de 95 ans, 80% se disent être en mauvaise santé. L'emploi prédit pour un groupe en santé parfaite (% de gens en mauvaise santé=0%) est alors de 78.885% et correspond à l'emploi prédit des individus de 101 ans. Alors pour obtenir l'emploi prédit des personnes de 95 ans (dont 80% sont en mauvaise santé) on calcule $78.885\% - (80\% * (-0.28)) = 56.48\%$.

En réalité, nous ne disposons que d'une seule personne de 101 ans pour laquelle nos données nous indiquent qu'elle est en santé parfaite ; elle se dit être en très bonne santé subjective et, selon nos données, elle ne possède aucun problème de santé particulier. Bien sûr, la capacité à travailler de cette personne ne peut être extrapolée à l'ensemble des personnes de 101 ans puisque nous sommes face à un cas particulier. Aussi, il est important de noter que plus l'âge observé est élevé, moins nous possédons d'individus dans notre base de données. Donc nous pourrions moins compter sur nos prédictions de capacité à travailler puisque ces résultats seront l'agrégation d'un nombre plus restreint de personnes. Nous présenterons donc désormais nos résultats de la capacité à travailler par tranche d'âge.

4.2.2. Analyse multi catégorielle de la santé

Nous répliquons ici l'analyse en préservant chacune des catégories de notre mesure de santé subjective (Très bonne santé, bonne santé, santé moyenne, mauvaise santé, très mauvaise santé). Le graphique 4.2.2.a ci-dessous montre la proportion d'individus au travail par catégorie de santé subjective pour les 50-54 ans ce qui montrera bien l'effet d'une mauvaise santé sur l'emploi que nous prédirons. Le nombre d'individus au travail change fortement et dans le sens attendu en fonction de la santé subjective : (pour la tranche d'âge des 50-54 ans, 13.3% des individus rapportant une santé « très pauvre » étaient au travail, 31.9% pour ceux rapportant une « santé pauvre », 53.6% pour ceux rapportant une « santé moyenne », 74.1% pour ceux rapportant « bonne santé » et 79.1% pour les individus rapportant une « très bonne santé »).

Graphique 4.2.2.a : Proportions d'individus de 50-54 ans au travail par catégorie de santé subjective



Le graphique 4.2.2.b révèle que le nombre moyen d'individus en très bonne et en bonne santé diminue avec l'âge tandis que ceux des individus en moyenne et en mauvaise santé augmente avec l'âge. Comme l'emploi est lié positivement avec le niveau de bonne santé subjective, et que le niveau de bonne santé subjective semble diminuer avec l'âge ; on peut s'attendre à une diminution de l'emploi avec l'âge lorsque l'on prédira l'emploi des individus plus âgés sur base des estimateurs des individus plus jeunes.

Le tableau 4.2.2.a reprend les résultats de la régression de l'emploi sur la variable de santé subjective pour les 50-54 ans. L'interprétation des coefficients des différentes catégories de santé subjective se fait avec la « santé moyenne » comme base. Par exemple ; une personne en très mauvaise santé aura, en moyenne, un taux d'emploi moyen inférieur de 41.8% à une personne en santé moyenne.

Graphique 4.2.2.b : Proportion d'individus dans chaque catégorie de santé subjective par âge

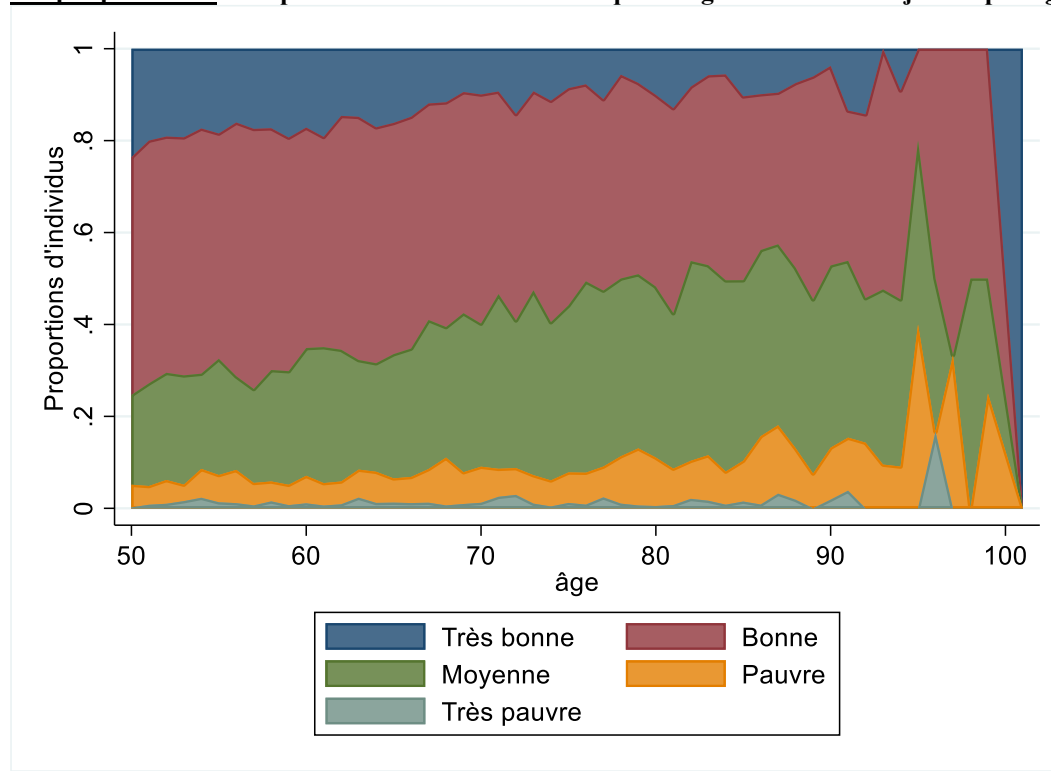


Tableau 4.2.2.a : Régression linéaire de l'emploi des 50-54 ans sur la santé subjective

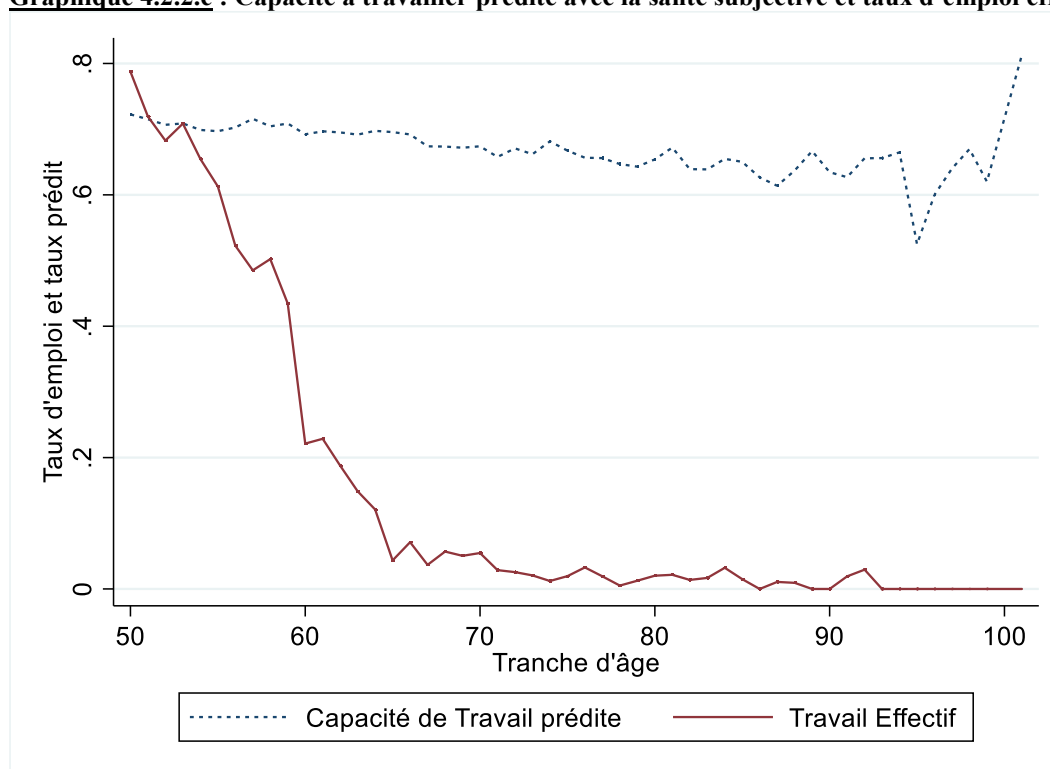
Emploi (dep. Var.)	Coef.	St.Err.	t-value	p-value	[95% Conf Interval]	Sig
Très bonne	0.253	0.028	9.05	0.000	0.198 0.308	***
Bonne	0.219	0.025	8.91	0.000	0.171 0.267	***
Moyenne (Base)	0.000
Pauvre	-0.198	0.050	-4.00	0.000	-0.296 -0.101	***
Très pauvre	-0.418	0.070	-6.00	0.000	-0.554 -0.281	***
Constante	0.560	0.022	25.96	0.000	0.518 0.603	***
Mean dependent var	0.679		SD dependent var	0.467		
R-squared	0.085		Number of obs	2794.000		
F-test	64.736		Prob > F	0.000		
Akaike crit. (AIC)	3436.298		Bayesian crit. (BIC)	3465.974		

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

Enfin, on présente les prédictions de taux d'emploi par catégorie d'âge dans le graphique 4.2.2.c. La capacité à travailler décline lentement passant de 67.9% à 50-54 ans à 60.4% pour les personnes âgées de plus de 85 ans. Pour rappel, la capacité à travailler ne signifie pas la quantité de travail que l'on *devrait* atteindre mais celle que l'on *pourrait* atteindre si seule la mauvaise santé entraînait un arrêt de travail.

Egalement, plus l'âge d'une personne augmente plus celle-ci a de chance de décéder. Il existe donc une *sélection naturelle des individus en meilleure santé* puisque, selon toute logique, ceux-ci vivent plus longtemps que ceux en mauvaise santé.

Graphique 4.2.2.c : Capacité à travailler prédite avec la santé subjective et taux d'emploi effectif par âge



4.2.3. La santé subjective et les indicateurs du handicap

Nous disposons de 4 variables sur le handicap nous informant si l'individu en est touché et dans quelle mesure il est limité dans ces activités quotidiennes. Inclure des indicateurs de handicap physique à la régression montre bien que les variables de santé subjectives prennent une large partie des variations de l'emploi du au handicap physique ; le handicap physique, à lui seul, réduit la probabilité d'être en emploi de 26.1% et les gens en situation d'handicap reporte être en moins bonne santé (subjective) parce qu'elle possède ce handicap.

Tableau 4.2.3.a : Régression linéaire de l'emploi des 50-54 ans sur la santé subjective et le handicap

Emploi	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
<i>Santé subjective</i>							
Très bonne	0.243	0.028	8.68	0.000	0.188	0.298	***
Bonne	0.209	0.025	8.51	0.000	0.161	0.257	***
Moyenne (Base)	0.000	
Pauvre	-0.165	0.050	-3.30	0.001	-0.263	-0.067	***
Très pauvre	-0.382	0.073	-5.22	0.000	-0.526	-0.239	***
Handicap	-0.261	0.059	-4.41	0.000	-0.376	-0.145	***
Constant	0.572	0.022	26.41	0.000	0.529	0.614	***
Mean dependent var	0.711				SD dependent var	0.454	
R-squared	0.121				Number of obs	2423.000	
F-test	41.612				Prob > F	0.000	
Akaike crit. (AIC)	2748.522				Bayesian crit. (BIC)	2800.657	

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

En comparant le tableau de régression ci-dessus avec celui obtenu lorsqu'on utilisait uniquement la santé subjective on voit directement que l'effet de se reporter en très mauvaise ou mauvaise santé subjective diminue largement puisqu'en réalité, beaucoup de cet effet venait du handicap physique. De la même manière, on compare l'effet de la bonne et très bonne santé lorsqu'on prend en compte le handicap ou non ; l'effet diminue également même si cette diminution est très légère. Pourquoi ? J'attribue, à première vue, la diminution de l'effet de la bonne santé subjective au fait que les personnes handicapées mais pour qui ce handicap est moins grave ressentent moins « être en mauvaise santé ». On peut tester cette hypothèse puisque nous disposons également d'indicateurs de sévérité des limitations dans les activités quotidiennes (LAQ). La sévérité des limitations est divisée en 4 catégories (Pas de limitations, limitations légères, limitations moyennes et limitations sévères) et nous montre que ne pas l'inclure conduit à biaiser (surestimation) largement l'estimation des coefficients de la santé subjective et du handicap puisque celle-ci englobait (lors de l'estimation précédente) la dimension des limitations dans les activités quotidiennes. En prenant en compte ces limitations dans les activités quotidiennes et surtout leur sévérité ; les effets sur l'emploi de reporter une mauvaise ou une très mauvaise santé subjective chute de 8.7 points de pourcentage. Le même raisonnement s'applique à la bonne et très bonne santé ; leurs effets diminuent parce qu'en réalité une partie de l'effet de la bonne santé venait du fait de ne pas posséder de limitations dans les activités quotidiennes (où de ne pas en posséder de trop sévères).

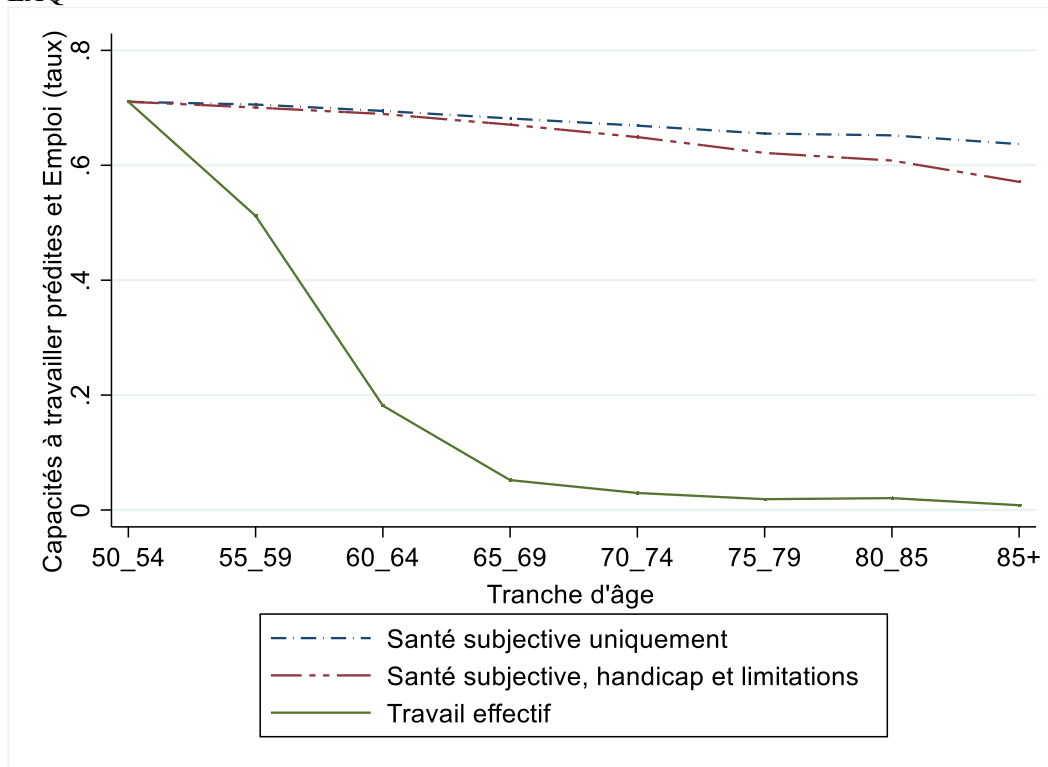
Une analyse plus détaillée de l'effet de l'ajout de variable de santé est vue dans la partie suivante sur les indicateurs de santé objectifs.

Tableau 4.2.2.c : Régression linéaire de l'emploi des 50-54 ans sur la santé subjective, le handicap et limitations dans les activités quotidiennes

	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
<i>Santé subjective</i>							
Très bonne	0.171	0.030	5.77	0.000	0.113	0.229	***
Bonne	0.149	0.026	5.78	0.000	0.098	0.199	***
b. Moyenne	0.000
Mauvaise	-0.078	0.050	-1.58	0.114	-0.175	0.019	.
Très mauvaise	-0.295	0.078	-3.81	0.000	-0.447	-0.143	***
Handicap	-0.185	0.060	-3.10	0.002	-0.302	-0.068	***
<i>Limitations dans les activités quotidiennes</i>							
B. Pas de limitations	0.000
Limitations légères	-0.075	0.034	-2.19	0.028	-0.141	-0.008	**
Limitations moyennes	-0.170	0.039	-4.39	0.000	-0.246	-0.094	***
Limitations sévères	-0.307	0.047	-6.48	0.000	-0.399	-0.214	***
Constant	0.648	0.024	26.85	0.000	0.601	0.695	***
Mean dependent var		0.711	SD dependent var			0.454	
R-squared		0.121	Number of obs			2423.000	
F-test		43.361	Prob > F			0.000	
Akaike crit. (AIC)		2748.522	Bayesian crit. (BIC)			2800.657	

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

Graphique 4.2.3.a : Capacité à travailler estimée (taux) selon la santé subjective avec et sans le handicap et LAQ



Finalement la capacité à travailler, déterminé avec l'aide de nos indicateurs de santé subjectifs, du handicap et des limitations dans les activités quotidiennes, décline plus rapidement que lorsqu'on utilise qu'un seul indicateur montrant l'importance de ceux-ci dans l'estimation.

Remarques sur la santé subjective

La santé subjective utilisée ici nous informe de l'état de santé dans lequel se sent l'individu. Bien que très intéressante comme proxy de la santé générale d'un individu et dans quelle mesure celle-ci intervient pour déterminer l'emploi, cette mesure contient plusieurs biais (Blundell et al., 2017) :

- (i) Erreurs de mesure : l'individu ne donne pas sa santé subjective de manière précise, ou, mauvaise retranscription des données, ...
→ L'estimation de l'effet de la santé est biaisée vers le bas car la somme des variances des différents indicateurs dépasse la variance de la santé (unique).
- (ii) Biais de justification : l'individu se dit être en mauvaise santé pour justifier son « non-emploi » ; i.e. la santé subjective est endogène, ou du moins, pour les personnes sans emploi disant « être en mauvaise santé » pour justifier leur situation de « sans-emploi ». On va alors surestimer l'effet de la mauvaise santé sur l'emploi → Les estimateurs seront biaisés vers le haut.

On peut également mentionner (iii) le biais de simultanéité ; les individus en emploi se sentent en meilleure santé et vice versa. Pour quelles raisons ? Par exemple, le revenu est corrélé positivement avec l'emploi et la santé subjective. Enfin, (iv) la causalité inversée abordée précédemment.

4.3. La santé objective

La santé générale peut également trouver des proxys dans des mesures de santé objective : des indicateurs binaires multiples indiquant si oui ou non l'individu est touché par un problème de santé particulier. Concrètement, nous disposons d'un ensemble d'indicateurs de santé objectifs (voir le tableau résumant les données par tranche d'âge) censés représenter la santé générale des individus.

Selon Blundell et al. (2017), 4 biais de mesures sont à prendre en considération lors de l'utilisation des indicateurs de santé objectifs :

- (i) Nous ne disposons que d'un nombre limité de d'indicateurs ne représentant dès lors que partiellement la santé générale d'un individu. Nous ne serons donc capables d'estimer que partiellement l'effet de la santé (objective) sur l'emploi ; l'effet lié aux indicateurs utilisés.
- (ii) Tous les indicateurs de santé n'ont pas le même effet sur l'emploi ; certains auront des effets parfois forts mais temporaire tandis que d'autres peuvent être faible mais perduré tout au long de la durée de vie.
- (iii) Les effets sur l'emploi d'un indicateur de santé en particulier peuvent être biaisés s'il existe des variables non observées liées à celui-ci.
- (iv) Un problème de santé peu caché différente ampleurs de celui-ci et aura donc plus et moins d'effet si on prend en compte son ampleur ; ce biais se devine lors de notre dernière régression prenant l'ampleur des limitations en compte.

Dans notre base de données, nous disposons de 29 variables de santé objectives (voir tableau 4.1.a et 4.1.b). Régresser l'emploi des 50-54 ans sur l'ensemble de ces variables objectives nous permet déjà d'observer quelles variables possèdent plus ou moins d'importance dans la relation d'emploi/santé (celles qui diffèrent statistiquement de 0 et celles dont l'estimateur est large en magnitude).

Dans un souci de clarté, nous observerons d'abord les prédictions du modèle en utilisant une seule variable (« Polyarthrite rhumatoïde»). Puis avec 2 variables (« Polyarthrite rhumatoïde» + « Glaucome »), puis plus,... Pour nous permettre de comprendre ce qu'implique l'inclusion d'une seule et plusieurs variables de santé objective au modèle.

4.3.1. Avec un seul indicateur

Observons d'abord l'effet sur la capacité à travailler lorsqu'on applique la méthode de Cutler avec l'aide d'une seule variable objective (« Polyarthrite rhumatoïde »). Le coefficient associé au handicap en mobilité est négatif et significatif (-0.099; signifie que si l'on passe d'une population ayant 0% de personne souffrant Polyarthrite rhumatoïde à une population de 100% de personnes souffrant de Polyarthrite rhumatoïde, alors l'emploi devrait diminuer de 0.99 points de pourcentage).

Aussi, le nombre de personnes souffrant de Polyarthrite rhumatoïde croît graduellement pour les hommes comme pour les femmes. Dû à un manque d'observations pour les âges les plus avancés, tous nos individus de 99 ans souffrent de Polyarthrite rhumatoïde tandis qu'aucun de nos individus de 101 ans ne souffre d'ostéo-arthrite. Dès lors, la capacité à travailler estimée sur base de ce modèle sera maximale pour les individus de 101 ans et minimal pour les individus de 99 ans.

Tableau 4.3.1.a : Régression linéaire de l'emploi des 50-54 ans sur la Polyarthrite rhumatoïde

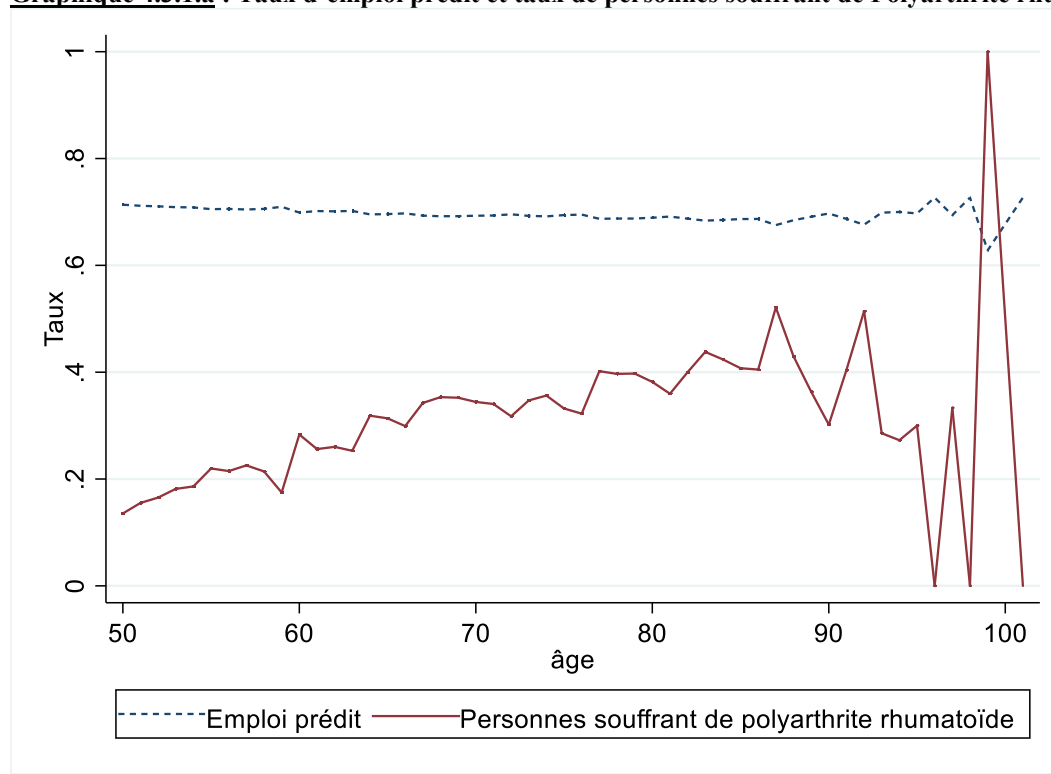
Emploi	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Polyarthrite rhumatoïde	-0.099	0.026	-3.77	0.000	-0.150	-0.047	***
Constant	0.727	0.010	73.39	0.000	0.707	0.746	***
Mean dependent var	0.711		SD dependent var	0.454			
R-squared	0.007		Number of obs	2423.000			
F-test	14.232		Prob > F	0.000			
Akaike crit. (AIC)	3031.704		Bayesian crit. (BIC)	3043.289			

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

Le graphique ci-dessous montre la relation entre l'évolution des prédictions d'emploi (sur base uniquement de l'indicateur de Polyarthrite rhumatoïde) et l'évolution du pourcentage de personnes souffrant de Polyarthrite rhumatoïde. On peut constater que l'emploi prédit en % par âge dépend directement de la proportion d'individus souffrant de Polyarthrite rhumatoïde dans cette même catégorie d'âge. Cette relation est semblable à celle exprimée par la régression ci-dessus ; à savoir, si le pourcentage de personne souffrant de cette maladie augmente de 10 points de pourcentage (entre 2 catégories d'âge) alors le taux de capacité à travailler prédit chutera de 1 point de pourcentage (entre ces deux mêmes catégories d'âge).

Parmi les personnes de 99 ans présentes dans nos données, 100% d'entre elles souffrent de Polyarthrite rhumatoïde ; l'effet négatif de cette maladie sur la capacité à travailler prédite est donc maximal pour cet âge en particulier. La capacité à travailler prédite sera alors la plus petite prédiction possible de ce modèle et est de 62.8%. Et vice-versa, puisque dans notre échantillon, 0% des personnes de 101 ans est touché par la maladie ; le modèle prédit donc une capacité à travailler maximale qui est de 72.7%. Ce qui correspond exactement à $62.8\% + (-100\%)*-0.099 = 72.7\%$ --> l'emploi prédit des personnes souffrant tous de Polyarthrite rhumatoïde moins l'effet négatif (donc plus d'emploi) de la maladie sur l'emploi.

Graphique 4.3.1.a : Taux d'emploi prédit et taux de personnes souffrant de Polyarthrite rhumatoïde.



4.3.2. Avec deux indicateurs et plus

Testons désormais la méthode mais en combinant les 2 variables de santé objectives et observons les prédictions qui en découlent. Le coefficient associé à la Polyarthrite rhumatoïde est toujours significatif et est très similaire à celui trouvé dans la régression précédente. Tandis que le coefficient associé au Glaucome n'est pas significatif mais est bien négatif et donc aura un impact dans le sens espéré sur nos prédictions de capacité à travailler.

Tableau 4.3.2.a : Régression linéaire de l'emploi des 50-54 ans sur la Polyarthrite rhumatoïde et le glaucome

emploi	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Polyarthrite rhumatoïde	-0.097	0.026	-3.70	0.000	-0.149	-0.046	***
Glaucome	-0.108	0.115	-0.93	0.351	-0.334	0.119	
Constant	0.727	0.010	73.38	0.000	0.708	0.747	***
Mean dependent var	0.711		SD dependent var	0.454			
R-squared	0.007		Number of obs	2423.000			
F-test	7.549		Prob > F	0.001			
Akaike crit. (AIC)	3032.638		Bayesian crit. (BIC)	3050.017			

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

Pour observer l'effet potentiel conjoint et individuel de ces deux variables de santé sur l'emploi, nous créons 4 groupes de personnes : le premier groupe ne souffre d'aucune des deux maladies. Le second groupe souffre uniquement des limitations. Le troisième groupe souffre uniquement d'un handicap en mobilité et le dernier groupe souffre des 2 soucis de santé. Ces caractéristiques sont reprises dans le tableau suivant ainsi que l'emploi prédit

associé à chacun de ces groupes (toujours sur base des estimateurs de la régression des 50-54)

Tableau 4.3.2.b : Capacité à travailler en fonction de deux variables objectives

Groupe	% de personne souffrant de polyarthrite	% de personnes souffrant d'un glaucome	Emploi prédit
1	0%	0%	72.7%
2	100%	0%	63%
3	0%	100%	62%
4	100%	100%	52.2%

L'emploi prédit des personnes ne souffrant d'aucune des deux maladies est de 72.7%. L'estimateur associé à la Polyarthrite rhumatoïde est de -0.097. Il y aura donc une différence d'emploi prédit de 9.7 points de pourcentage entre un groupe composé à 100% d'individus souffrant de la maladie et un groupe composé à 0% d'individus souffrant de la maladie. Le même raisonnement s'applique au glaucome. Et l'effet des 2 n'est rien d'autre que la somme de leurs effets individuels ($72.7\% - (0.097 + 0.108) = 52.2\%$).

Lorsqu'on compare les résultats obtenus entre le modèle ne contenant qu'une seule variable objective et celui en utilisant 2, on peut comprendre ce qu'implique le fait d'omettre une ou plusieurs variables de santé ; si un seul indicateur est utilisé, l'effet sur l'emploi de ce souci de santé en particulier sera très certainement biaisé. Pourquoi ? Car il englobera un ensemble de cofacteurs non-observables (les variables omises) ayant (i) un effet individuel sur l'emploi et (ii) corrélé avec notre unique indicateur de (mauvaise) santé.

Pour rappel, un indicateur de santé objectif prend une valeur de 1 si l'individu souffre du problème de santé et 0 si non. Dès lors, en supposant que les soucis de santé sont (tous) liés négativement avec l'emploi et qu'avoir une maladie particulière augmente les chances d'en avoir d'autres (i.e. les indicateurs de santé objectifs sont corrélés positivement entre eux). Alors omettre des indicateurs de santé objectifs entrainera une surestimation des effets des indicateurs présents dans la régression.

Néanmoins, l'effet global de la mauvaise santé sera sous-estimé puisqu'on ne sera capable d'estimer que l'effet partiel (lié aux seuls variables utilisés) de la santé sur l'emploi. Cette sous-estimation de l'effet de la santé globale sur l'emploi à 50-54 ans implique que la capacité à travailler estimée sera surestimée. Par exemple, le taux d'emploi minimal possible (i.e. la capacité à travailler prédite de nos individus en plus mauvaise santé) du premier modèle (un seul indicateur) est de 62.8% tandis que celui du deuxième (avec deux indicateurs) est de 52.2%.

Pour montrer ce qu'on vient de sous-entendre, faisons l'exercice d'estimer la capacité à travailler sur un modèle incluant une seule variable de santé objective, puis deux, puis trois, ...

Et observons l'évolution des résultats obtenus lorsqu'on ajoute ponctuellement de nouveaux indicateurs de santé objectifs. Le tableau ci-dessous reprend le coefficient associé à chaque indicateur pour chaque modèle estimé ; les coefficients associés à chaque indicateur, tout comme les prédictions de capacité à travailler, suivent une tendance à la baisse lorsqu'on

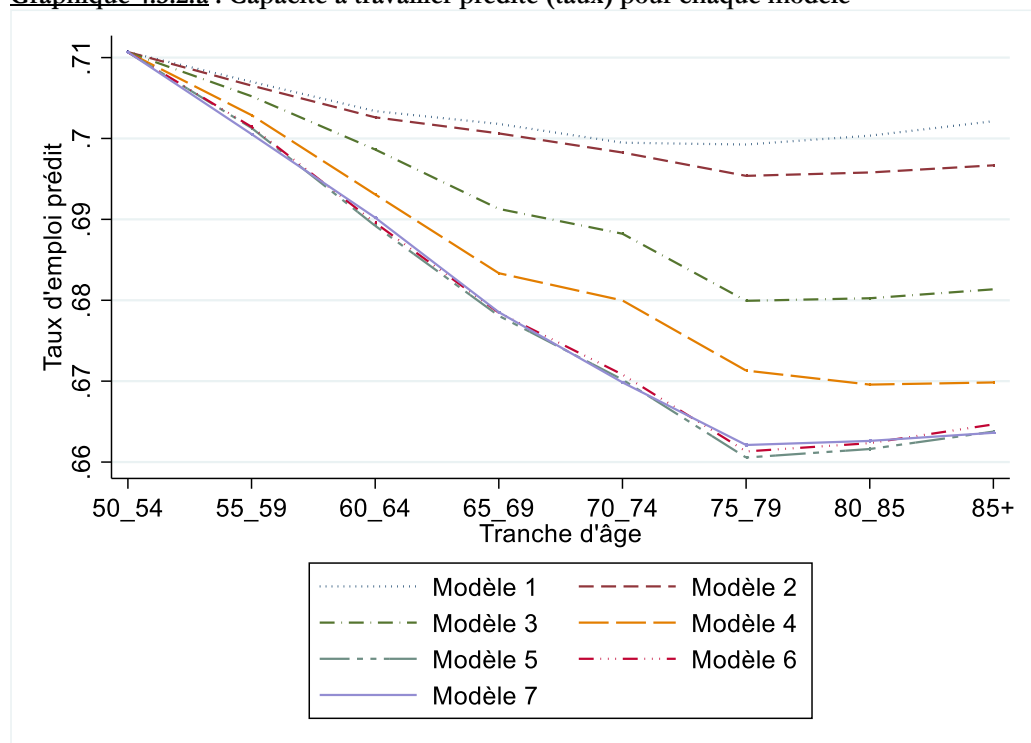
ajoute des indicateurs au modèle. Le graphique, ci-dessous aussi, nous donne les prédictions de capacité à travailler en fonction du modèle ; ajouter des variables objectives au modèle semble faire diminuer la capacité à travailler prédite pour tout âge jusqu'à un certain limite.

Tableau 4.3.2.c : Régressions linéaires diverse montrant l'effet de l'ajout ponctuel de variable de santé

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hypertension artérielle	-0.074*** (0.025)	-0.072*** (0.025)	-0.067*** (0.025)	-0.063** (0.025)	-0.053** (0.025)	-0.053** (0.025)	-0.051** (0.025)
AVC (ou conséquences)		-0.189 (0.133)	-0.182 (0.134)	-0.177 (0.134)	-0.176 (0.132)	-0.173 (0.133)	-0.162 (0.133)
La polyarthrite rhumatoïde			-0.147*** (0.036)	-0.119*** (0.038)	-0.119*** (0.038)	-0.117*** (0.038)	-0.116*** (0.038)
Arthrose				-0.062** (0.027)	-0.058** (0.027)	-0.057** (0.027)	-0.055** (0.027)
Diabète					-0.148*** (0.052)	-0.149*** (0.051)	-0.141*** (0.052)
Allergie						-0.013 (0.028)	-0.006 (0.027)
Ulcère de l'estomac							-0.153*** (0.047)
_cons	0.724*** (0.010)	0.725*** (0.010)	0.736*** (0.010)	0.743*** (0.011)	0.747*** (0.011)	0.748*** (0.011)	0.754*** (0.011)
Obs.	2423	2423	2423	2423	2423	2423	2423
R-squared	0.004	0.005	0.013	0.015	0.019	0.019	0.025

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

Graphique 4.3.2.a : Capacité à travailler prédite (taux) pour chaque modèle



Comme nous ne disposons pas de la variable unique et absolue de santé (comprenant toutes les variations possibles de la santé) mais uniquement d'indicateurs multiples de santé partiel. L'effet estimé de la santé générale sur l'emploi sera sous-estimé pour 2 raisons (Blundell, 2017) :

- (i) L'utilisation d'indicateurs partiels, au lieu de l'indicateur « ultime » de santé, induit une variabilité de la santé plus importante et donc biaise l'effet total de la santé vers 0. --> *Biais d'atténuation*.
- (ii) On est capable d'estimer que l'effet partiel de la santé sur l'emploi puisqu'on ne dispose que partiellement (chacun de nos indicateurs en est une partie) de la santé générale. --> *Biais de la variable omise*.

4.3.3. Avec l'ensemble des indicateurs

Tableau 4.3.3.a : Régression linéaire des individus de 50-54 ans sur l'ensemble des variables objectives

Emploi	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Haut. pression sanguine	-0.037	0.024	-1.54	0.123	-0.085	0.010	
AVC et conséquences	-0.093	0.134	-0.69	0.488	-0.356	0.170	
Polyarthrite rhumatoïde	-0.069	0.037	-1.84	0.066	-0.142	0.005	*
Arthrose	-0.015	0.027	-0.55	0.582	-0.068	0.038	
Diabète	-0.113	0.050	-2.26	0.024	-0.210	-0.015	**
Allergie	0.021	0.027	0.77	0.441	-0.032	0.073	
Ulcère de l'estomac	-0.080	0.046	-1.73	0.084	-0.171	0.011	*
Cirrhose du foie	-0.131	0.097	-1.35	0.177	-0.321	0.059	
Cancer	-0.203	0.068	-3.01	0.003	-0.336	-0.071	***
Maux de tête sévères	-0.069	0.030	-2.26	0.024	-0.128	-0.009	**
Morosité ou dépression	-0.162	0.037	-4.36	0.000	-0.234	-0.089	***
Problèmes de thyroïde	-0.059	0.041	-1.44	0.150	-0.140	0.021	
Glaucome	0.007	0.116	0.06	0.955	-0.221	0.234	
Cataracte	0.039	0.075	0.52	0.607	-0.108	0.186	
Maladie de Parkinson	-0.061	0.352	-0.17	0.863	-0.751	0.629	
Épilepsie	-0.447	0.132	-3.39	0.001	-0.706	-0.189	***
L'ostéoporose	-0.131	0.051	-2.59	0.010	-0.230	-0.032	**
Hanche cassée	-0.106	0.339	-0.31	0.754	-0.771	0.559	
Trouble de l'intestin	-0.160	0.055	-2.92	0.004	-0.268	-0.052	***
Des pierres dans le rein	-0.127	0.097	-1.31	0.190	-0.316	0.063	
Maladie du rein	0.038	0.097	0.39	0.698	-0.152	0.228	
Cystite chronique	-0.166	0.081	-2.06	0.040	-0.325	-0.008	**
Maladie de peau	-0.076	0.052	-1.47	0.142	-0.178	0.026	
Inflammation vés. bilia.	-0.203	0.097	-2.08	0.037	-0.394	-0.012	**
Problème de prostate	0.104	0.079	1.31	0.190	-0.051	0.260	
Constante	0.783	0.011	69.29	0.000	0.761	0.805	***
Mean dependent var	0.711		SD dependent var	0.454			
R-squared	0.070		Number of obs	2423.000			
F-test	7.608		Prob > F	0.000			
Akaike crit. (AIC)	2918.801		Bayesian crit. (BIC)	3069.413			

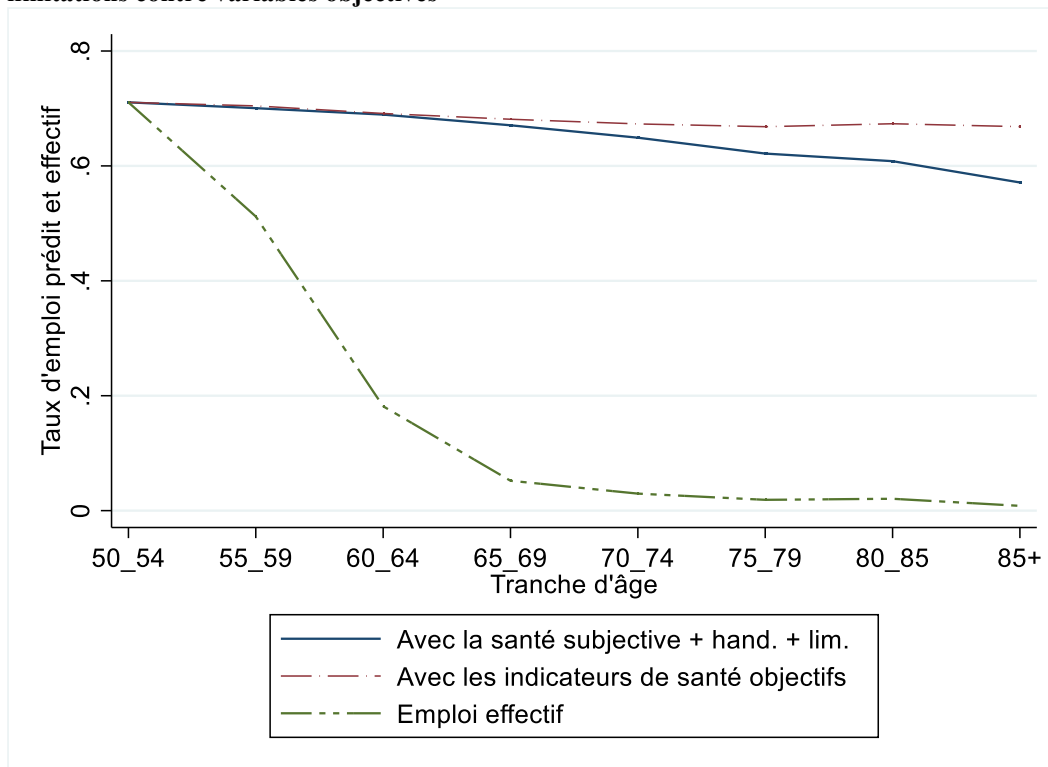
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appliquons désormais la méthode de Cutler en utilisant l'ensemble de nos variables objectives de santé pour ensuite pouvoir comparer ses prédictions avec celles obtenues en utilisant uniquement l'indicateur de santé subjective. Le tableau reprenant les estimateurs associés à chacun des indicateurs est ci-dessus tandis que le graphique des prédictions de ce modèle se trouve juste en dessous. Sur ce même graphique est également représentée la capacité à

travailler prédite du modèle utilisant uniquement la santé subjective pour pouvoir comparer les deux.

On peut voir que, d'une manière générale, la relation entre l'emploi et santé utilisant des indicateurs de santé subjectifs engendre un plus grand déclin de la capacité à travailler que celle utilisant des indicateurs de santé objectifs.

Graphique 4.3.3.a : Taux d'emploi prédit des personnes de 50-54 ans ; variables subjectives, handicap et limitations contre variables objectives



Remarques sur la santé objective

En supposant que chaque indicateur de santé objectif possède une relation négative avec l'emploi des personnes de 50-54 ans et que la proportion d'individus touchés par chaque indicateur objectif augmente avec l'âge. Alors omettre une variable objective au modèle entrainera une sous-estimation de l'effet de la santé générale sur l'emploi ; c'est le *biais de la variable omise* qui peut être potentiellement important lorsqu'on utilise uniquement des indicateurs de santé objectifs dans la régression.

4.4. Avec la santé subjective et objective

4.4.1. Toutes les variables en une seule régression

Tableau 4.4.1.a : Régression linéaire des individus de 50-54 ans sur l'ensemble des variables de santé

Empl	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
<i>Santé Subjective</i>							
Très bonne	0.155	0.031	5.06	0.000	0.095	0.215	***
Bonne	0.133	0.025	5.37	0.000	0.085	0.182	***
b. Moyenne	0.000	
Mauvaise	-0.059	0.048	-1.22	0.222	-0.153	0.035	
Très mauvaise	-0.283	0.086	-3.29	0.001	-0.452	-0.114	***
<i>Handicap et Limitations</i>							
Handicap en mobilité	-0.179	0.061	-2.95	0.003	-0.298	-0.060	***
Sévérité des limitations	-0.094	0.043	-2.20	0.028	-0.178	-0.010	**
Limitations	0.025	0.067	0.37	0.711	-0.106	0.156	
Limitation sévère	-0.016	0.077	-0.21	0.837	-0.167	0.135	
<i>Santé objective</i>							
Haut. pression sanguine	-0.011	0.023	-0.48	0.628	-0.056	0.034	
AVC et conséquences	-0.065	0.115	-0.56	0.573	-0.291	0.161	
Polyarthrite rhumatoïde	-0.033	0.034	-0.98	0.328	-0.100	0.033	
Arthrose	0.033	0.026	1.29	0.198	-0.017	0.084	
Diabète	-0.054	0.046	-1.18	0.240	-0.143	0.036	
Allergie	0.032	0.027	1.20	0.231	-0.020	0.084	
Ulcère de l'estomac	-0.018	0.042	-0.42	0.671	-0.099	0.064	
Cirrhose du foie	0.044	0.091	0.48	0.631	-0.135	0.223	
Cancer	-0.070	0.062	-1.13	0.259	-0.192	0.052	
Maux de tête sévères	-0.023	0.028	-0.83	0.409	-0.077	0.032	
Morosité ou dépression	-0.042	0.035	-1.18	0.238	-0.111	0.028	
Problèmes de thyroïde	-0.040	0.037	-1.09	0.275	-0.113	0.032	
Glaucome	0.017	0.101	0.17	0.865	-0.181	0.216	
Cataracte	0.098	0.071	1.38	0.167	-0.041	0.238	
Maladie de Parkinson	0.032	0.304	0.11	0.916	-0.563	0.627	
Épilepsie	-0.375	0.137	-2.74	0.006	-0.643	-0.106	***
L'ostéoporose	-0.082	0.045	-1.80	0.072	-0.171	0.007	*
Hanche cassée	0.122	0.304	0.40	0.687	-0.473	0.718	
Trouble de l'intestin	-0.083	0.050	-1.67	0.096	-0.181	0.015	*
Des pierres dans le rein	-0.121	0.085	-1.42	0.155	-0.287	0.046	
Maladie du rein	0.137	0.101	1.35	0.176	-0.061	0.335	
Cystite chronique	-0.129	0.075	-1.72	0.086	-0.276	0.018	*
Maladie de peau	-0.048	0.047	-1.03	0.305	-0.140	0.044	
Inflammation vés. bilia.	-0.222	0.083	-2.65	0.008	-0.385	-0.058	***
Problème de prostate	0.108	0.081	1.33	0.183	-0.051	0.266	
<i>Santé psychologique</i>							
Moyenne GHQ-12 score	0.008	0.008	1.06	0.288	-0.007	0.023	
Détresse psychologique	0.058	0.032	1.80	0.072	-0.005	0.122	*
Troubles mentaux	-0.110	0.048	-2.31	0.021	-0.204	-0.017	**
Constant	0.664	0.026	25.73	0.000	0.613	0.714	***
Mean dependent var		0.711	SD dependent var			0.454	
R-squared		0.140	Number of obs			2423.000	
F-test		10.778	Prob > F			0.000	
Akaike crit. (AIC)		2752.469	Bayesian crit. (BIC)			2966.801	

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

4.4.2. En instrumentant la santé subjective par la santé objective

Indicateurs de santé subjectifs comme objectifs souffrent d'erreurs de mesures. L'utilisation des indicateurs de santé objectifs implique aussi le *biais de la variable omise* (qui entraîne une sous-estimation de l'effet de la santé sur l'emploi) tandis que la santé subjective entraîne un *biais de justification* qui nous poussera à surestimer l'effet de la santé (subjective) sur l'emploi.

On aimerait obtenir l'effet de la santé sur l'emploi en nettoyant notre santé subjective du biais de justification. Une méthode largement utilisée dans la littérature et notamment dans celle en référence à la méthode de Cutler, est la méthode des variables instrumentales.

En supposant que les biais issus de la santé subjective ne sont pas liés aux variables objectives, alors on peut instrumenter la santé subjective par l'ensemble des indicateurs de santé objectifs pour enlever le biais de justification. En instrumentant la santé subjective avec les variables de santé objectives, on restreint l'analyse aux variations de la santé subjective causées par celles de la santé objective (Catherine, 2015). Donc les personnes rapportant une mauvaise santé subjective mais possédant une bonne santé objective ne seront plus (moins) comprises dans l'estimation de l'effet de la santé sur l'emploi ;

$$\beta_{ols} = \frac{cov(y, x)}{cov(x, x)}$$

$$\beta_{IV} = \frac{cov(y, z)}{cov(x, z)}$$

La méthode des variables instrumentales trouve tout son sens lorsqu'on possède une variable « indépendante » que l'on pense être corrélée avec le terme d'erreur ;

$$E_i = \alpha + \beta * \text{santé} + \varepsilon_i$$

Et $cov(\text{santé}, \varepsilon) \neq 0$

Le choix des instruments utilisés lors de la méthode IV nécessite certaines explications pour assurer que ceux-ci soient valides. Un instrument est valide sous plusieurs conditions ;

- (i) Il doit être suffisamment corrélé avec la variable endogène (santé subjective).

$$cov(z, x) \neq 0$$

- (ii) Il ne peut être corrélé avec la variable dépendante (l'emploi) que via la variable endogène (les variations communes de la santé objective et de l'emploi sont comprises dans la santé subjective).

- (iii) Et doit être indépendant du terme d'erreur de la première équation.

$$cov(z, \varepsilon) = 0$$

En pratique la méthode des variables instrumentales consiste simplement en une étape supplémentaire avant l'estimation des coefficients de la santé ; (i) on estime un modèle de la santé subjective en fonction de tous les indicateurs de santé objectifs (X_i).

$$i. \widehat{SelfReportedHealth} = \alpha + \sum(\beta_i * X_i) + u_i \quad (i)$$

$$E_i = \alpha + i. \widehat{SelfReportedHealth} + \varepsilon_i \quad (ii)$$

On prédit les valeurs de la variable endogène sur base uniquement des variables de santé objectives et on utilise ce nouvel indicateur de santé subjective instrumenté par la santé objective comme indicateur de santé (ii). La santé subjective prédite dépend alors directement de la santé objective ; les individus rapportant une santé subjective mauvaise mais possédant une bonne santé objective ne seront plus pris (ou moins pris en compte) en compte lors de l'estimation de l'effet de la santé sur l'emploi à 50-54 ans.

Tableau 4.4.2.a : Régression par variable instrumentale

Empl	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Très Bonne	0.391	0.184	2.12	0.034	0.030	0.753	**
Bonne	0.634	0.193	3.28	0.001	0.255	1.013	***
Moyenne (base)	0.000	
Mauvaise	0.073	0.270	0.27	0.787	-0.457	0.603	
Très mauvaise	0.106	0.531	0.20	0.842	-0.934	1.146	
Handicap	-0.118	0.071	-1.65	0.098	-0.258	0.022	*
Sévérité des lim.	-0.025	0.063	-0.40	0.690	-0.150	0.099	
Limitation	0.076	0.097	0.79	0.432	-0.114	0.267	
Limitation grave	-0.149	0.098	-1.51	0.130	-0.341	0.044	
Constante	0.300	0.144	2.08	0.037	0.018	0.583	**
Mean dependent var	0.711		SD dependent var	0.454			
R-squared	.		Number of obs	2423.000			
Chi-square	276.841		Prob > chi2	0.000			

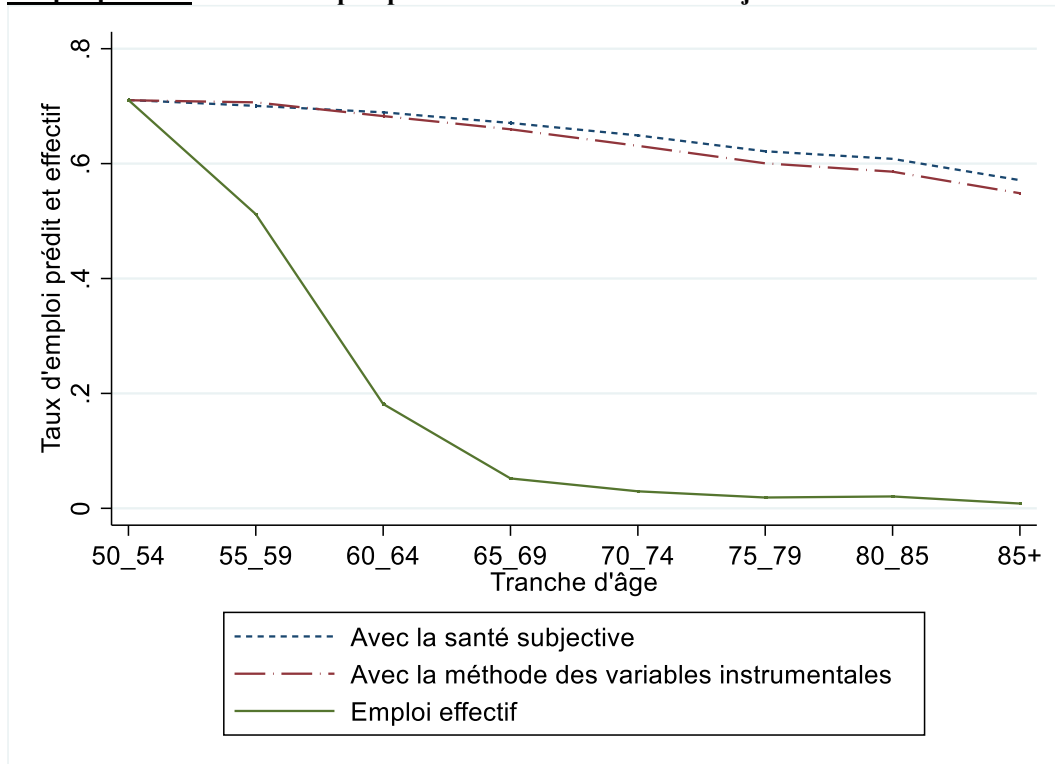
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note : L'ensemble des variables de santé objectifs est utilisé pour instrumenter la santé subjective

Après instrumentation de la santé subjective par l'ensemble des indicateurs de santé objectifs on peut voir que l'effet de se définir en santé « très pauvre », « pauvre » ou « moyenne » sur l'emploi, n'est statistiquement pas différent ce qui semble indiquer la présence du biais de justification.

Enfin on présente les prédictions de la capacité à travailler estimées avec l'aide de la méthode des variables instrumentales pour la comparer avec celle obtenue précédemment avec les variables subjectives et objectives.

Graphique 4.4.1 : Taux d'emploi prédit et effectif : variables subjectives contre IV



Comme on peut le voir graphiquement, la capacité à travailler obtenue via la méthode des variables instrumentales semble prédire une capacité proche mais légèrement inférieure à celle obtenue précédemment lorsqu'on ne prenait pas en compte l'endogénéité de la santé subjective.

Remarque sur l'endogénéité de la santé subjective

La capacité à travailler obtenue en instrumentant la santé subjective par des indicateurs de santé objectifs est pratiquement pareille que lorsqu'on ne contrôle pas pour l'endogénéité de la santé subjective. Certains auteurs (Blundell, 2017) plaident que les erreurs de mesure (sous-estimation de la relation emploi/santé) compensent presque parfaitement le biais de justification (surestimation de la relation emploi/santé) bien que d'autres auteurs estiment que le biais de justification domine.

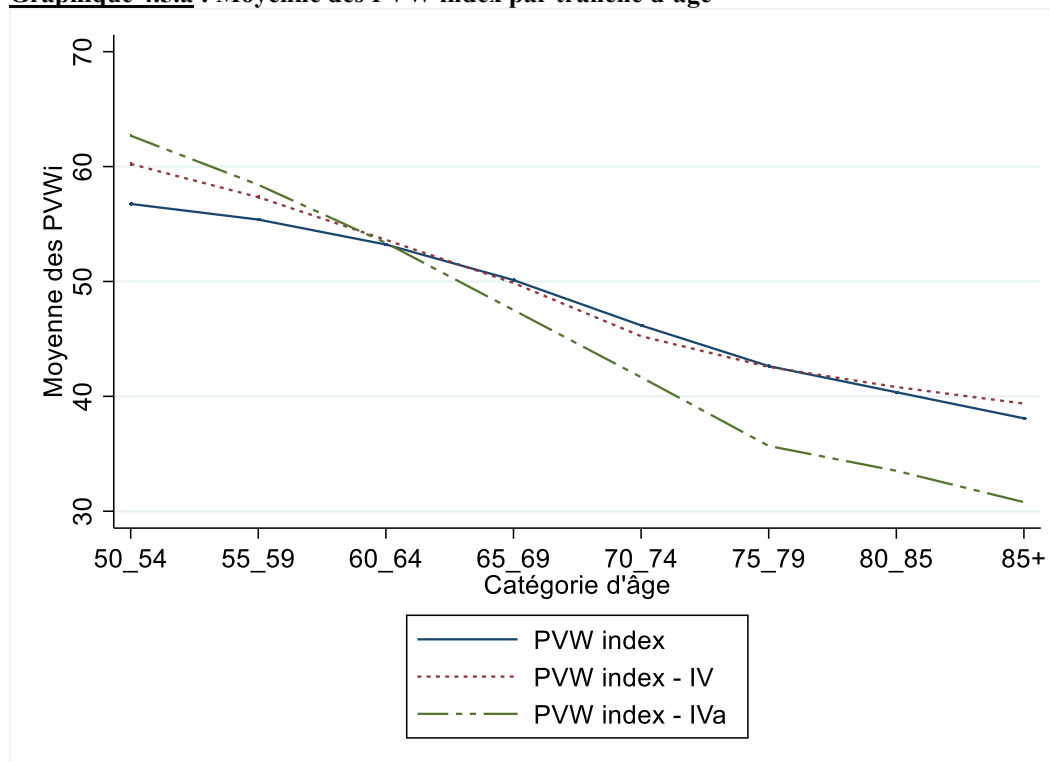
4.5. Synthétiser la santé en un seul indicateur

La méthode de Cutler implique l'utilisation d'un ensemble d'indicateurs de santé assez large. Il existe certaines variables plus importantes que d'autres dans la relation estimée (la santé subjective par exemple) ce qui nous permet de voir l'effet partiel de la variable en question sur l'emploi à 50-54 ans. Néanmoins, déterminer l'effet de la santé « globale » sur l'emploi est impossible puisque nous ne disposons pas d'indicateurs de la santé globale, uniquement d'indicateurs de santé partiel.

Une approche souvent utilisée dans la littérature sur la capacité à travailler est de synthétiser l'ensemble de nos indicateurs de santé en un seul indicateur : le Poterba, Venti et Wise index (PVW index) (Poterba et al.,2013). Pour ce faire, la méthode requiert une analyse en composantes principales que Wise (2017, p.14) définit comme ; « ...la moyenne pondérée des indicateurs pour laquelle les poids sont choisis pour maximiser la proportion de la variance de l'indicateur unique de santé qui peut être expliquée par ces moyennes pondérées ». Bien que la justification mathématique nous échappe, l'idée reste simple ; obtenir les variations communes entre les variables de santé pour nous permettre d'en construire une autre synthétisant leurs variations communes ; c'est l'indice de santé unique ; le PVW index.

L'indicateur unique est alors remis en centile pour lesquels un centile supérieur signifie une meilleure santé. Pour obtenir le PVW index on utilise la technique de l'analyse en composante principale appliquée à l'ensemble de nos indicateurs de santé. On utilise alors la première composante principale pour prédire un score de santé pour chaque individu. Ces scores sont ensuite transformés en centiles pour lesquels une valeur plus élevée signifie une meilleure santé.

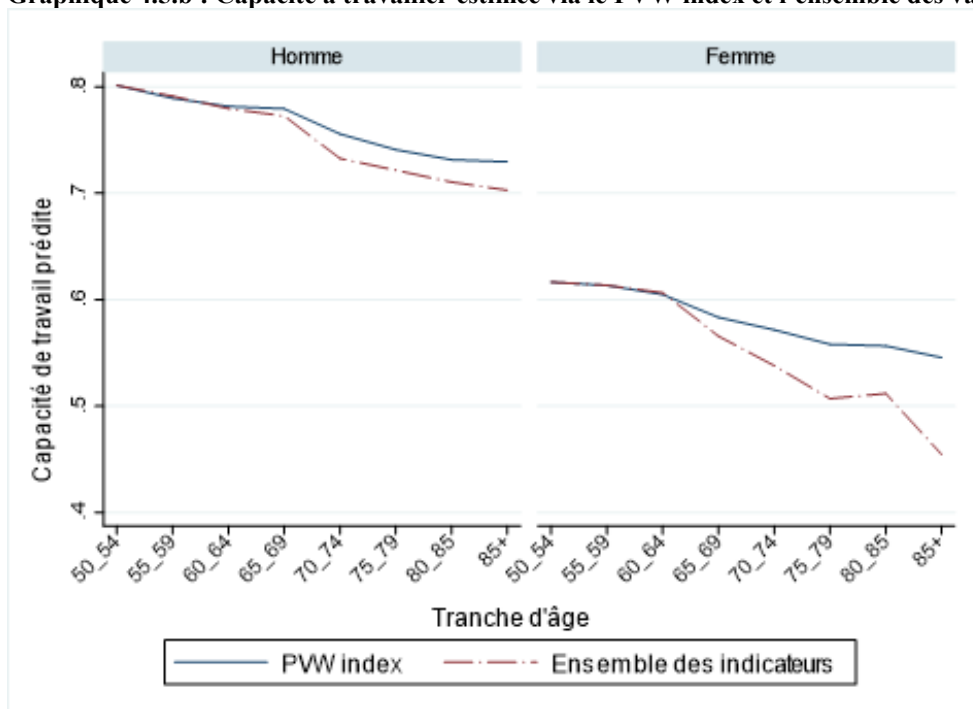
Graphique 4.5.a : Moyenne des PVW index par tranche d'âge



Avec cette approche il est encore possible de contrôler pour le biais de justification présent dans la santé subjective en créant un deuxième PVW index (PVW index – IV) pour lequel la santé subjective est une combinaison linéaire de nos autres variables de santé (c à d ; même principe que la technique des variables instrumentale). Ces deux PVW index sont assez similaires bien que celui instrumenté soit légèrement supérieur pour les deux premières catégories d'âges (graphe 4.5.a) ce qui nous informe que le biais d'endogénéité n'est pas si important mais présent tout de même. Aussi, dans le graphique, se trouve un troisième PVW index (PVW index – Iva) que nous expliquerons plus tard.

Comment déterminer si notre unique indice de santé est bien défini ? C'est-à-dire ; on veut s'assurer que celui-ci englobe bien le plus de variations possibles de la santé telle qu'elle aurait été définie si on avait utilisé l'ensemble de nos variables de santé. Comme dit précédemment, les subtilités mathématiques de l'analyse en composantes principales m'échappent. Je compare donc la capacité à travailler obtenue via notre indice de santé unique en la comparant avec celle obtenue en utilisant l'ensemble des indicateurs de santé. Si les prédictions des capacités de travail des deux méthodes sont proches, alors j'en conclus que mon indice de santé unique est bon. Les résultats sont présentés dans le tableau suivant et montrent que pour les hommes comme pour les femmes, notre index unique de santé est bon... Si on devait s'arrêter d'estimer la capacité à travailler à partir de l'âge de 60-64 ans. Au-delà, les prédictions sont surestimée et surtout pour les femmes.

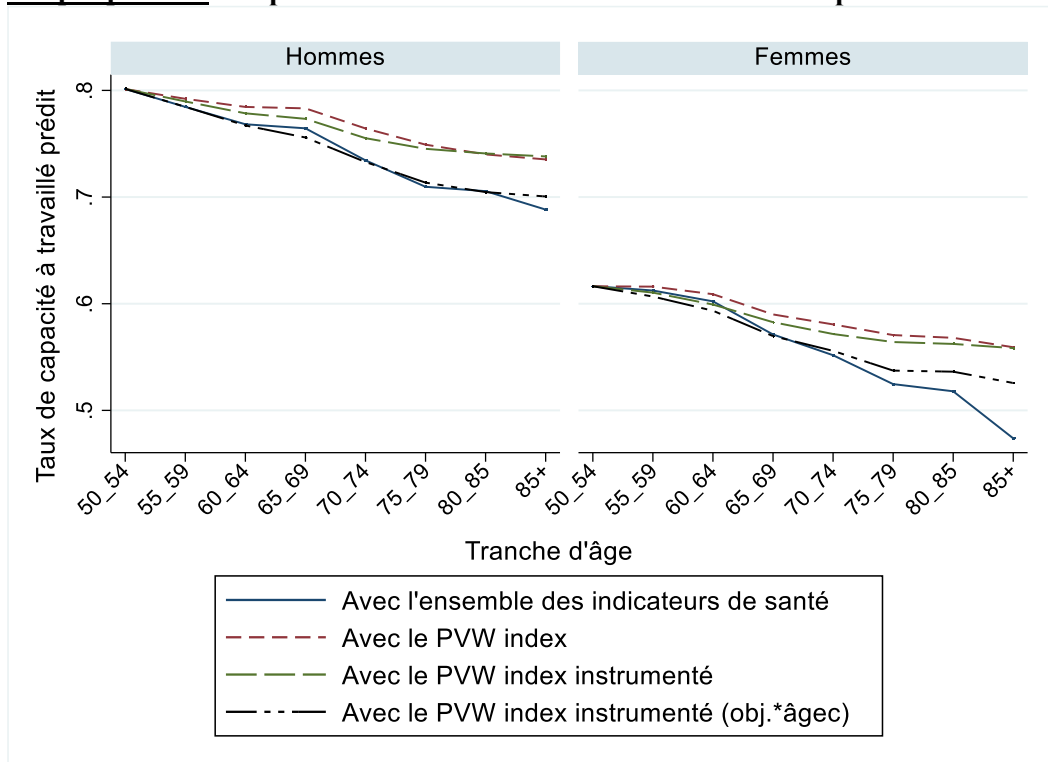
Graphique 4.5.b : Capacité à travailler estimée via le PVW index et l'ensemble des variables de santé



Les résultats du PVW index pour prédire la capacité à travailler tel que l'ensemble de nos indices de santé semblent de moins en moins bons avec l'âge ce qui me semble refléter une des limites de la méthode de Cutler ; l'effet de la santé sur l'emploi est supposé constant avec l'âge. Or, comme on peut le constater ici, les capacités de travail estimées avec l'indice de santé unique sont de moins en moins bonnes à mesure que l'âge des individus avance.

Jürges and all (2017) présentent une version augmentée de l'instrumentalisation de la santé subjective qui permet d'assouplir cette hypothèse en autorisant celle-ci à varier avec l'âge ; concrètement, il ajoute des termes d'interactions de tranches d'âges à chaque indicateur objectif utilisé lors de l'instrumentalisation de la santé subjective. On fait donc de même et remarquons que la capacité à travailler estimée avec l'indicateur unique est désormais très proche de celle estimée avec l'ensemble des indicateurs de santé. Cet unique indicateur de santé pour lequel la santé subjective est instrumentée par les indicateurs de santé objectifs et les tranches d'âges (PVW index-IVa) sera celui utilisé pour la suite de ce travail.

Graphique 4.5.c : Capacité à travailler estimée via les différentes spécifications de la santé



L'estimation de l'effet de la santé synthétisé par le PVW index est reprise dans le tableau suivant. Les coefficients associés aux deux PVW index sont très significatifs et très proches. Le coefficient associé au PVW index sans instrumentation de la santé subjective (première colonne) est de 0.0034 et représente l'effet d'une augmentation d'un centile de notre indice unique santé sur l'emploi à 50-54 ans. En instrumentant des variables de santé objectives, on peut voir que l'effet de la santé sur l'emploi est en réalité plus faible puisqu'il chute à 0.30 (deuxième colonne). En ajoutant des termes d'interaction par tranche d'âge, celui-ci se renforce très légèrement (troisième colonne). Le coefficient associé au PVW index instrumenté par les indicateurs de santé objectif et par les tranches d'âges s'interprète comme tel ; « un groupe d'individu de 50-54 possédant un PVW index (3) de 10 percentiles en plus qu'un autre groupe possèdera en moyenne un taux d'emploi de 3.1 points de pourcentage plus élevé ». En d'autres mots, la différence de taux d'emploi entre une personne de santé maximale et minimale est de 31%.

Tableau 4.5.a : Régression de l'emploi des 50-54 ans sur nos différents PVW index

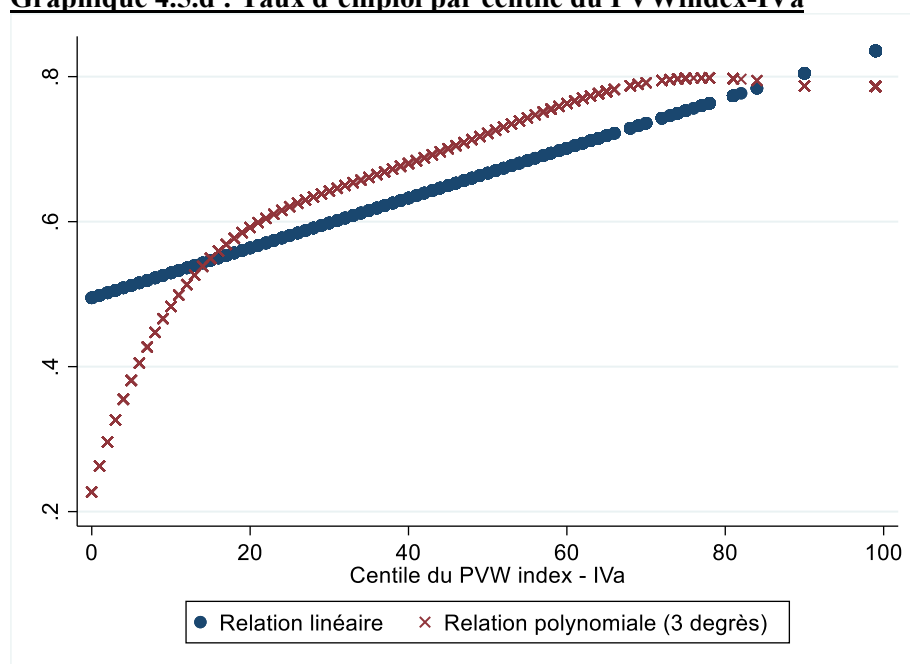
	(1)	(2)	(3)
PVWindex	0.0034*** (0.0003)		
PVWindex-IV		0.0030*** (0.0003)	
PVWindex-IVa			0.0031*** (0.0003)
Femme	-0.1549*** (0.0178)	-0.1547*** (0.0179)	-0.1542*** (0.0179)
_cons	0.5944*** (0.0216)	0.6061*** (0.0224)	0.5949*** (0.0227)
Obs.	2423	2423	2423
R-squared	0.0927	0.0829	0.0858

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

Comme nous disposons désormais d'un unique indice de santé prenant des valeurs de 0 à 99, il nous est possible de représenter graphiquement la relation emploi/santé des 50-54 ans estimée ci-dessus.

Le graphique suivant représente la relation estimée dans la troisième colonne du tableau précédent. Celle-ci est linéaire puisque ce modèle estime une relation linéaire. Néanmoins, il nous est possible de tester si cette relation est bel et bien linéaire ou non en utilisant une forme polynomiale du PVW index-IVa lors de la régression. On le voit aisément ; lorsqu'on estime une relation linéaire on obtient l'effet moyen de la santé sur l'emploi tandis qu'en autorisant la relation à être non linéaire on constate qu'en réalité la relation emploi/santé est concave ; la santé a un effet relativement fort sur l'emploi pour un bas niveau de santé et son effet s'affaiblit à mesure que la santé s'améliore. Aussi, lorsque la santé est bonne (supérieure au septantième centile), l'effet de celle-ci sur l'emploi est plus ou moins nulle.

Graphique 4.5.d : Taux d'emploi par centile du PVWindex-IVa



Comme on estimera la relation sous sa forme linéaire, l'effet de la santé sur l'emploi sera sous-estimé pour les individus possédant une santé très pauvre comme on peut le voir sur le graphique précédent ; l'élasticité de la santé à l'emploi est constante sous la forme linéaire utilisée dans la suite de ce travail pour estimer la capacité à travailler. Pour la forme non-linéaire, celle-ci nous révèle qu'en réalité l'élasticité de l'emploi à la santé est très forte pour les individus en très mauvaise santé (ceux pour lesquels le PVW index prend une valeur approximativement inférieure au 20^{ième} centile) et pratiquement nulle pour les individus en très bonne santé (supérieure au 70^{ième} centile). Pour les individus entre le 20^{ième} et 70^{ième} centile ; l'élasticité est plus ou moins constante et égale à la forme linéaire.

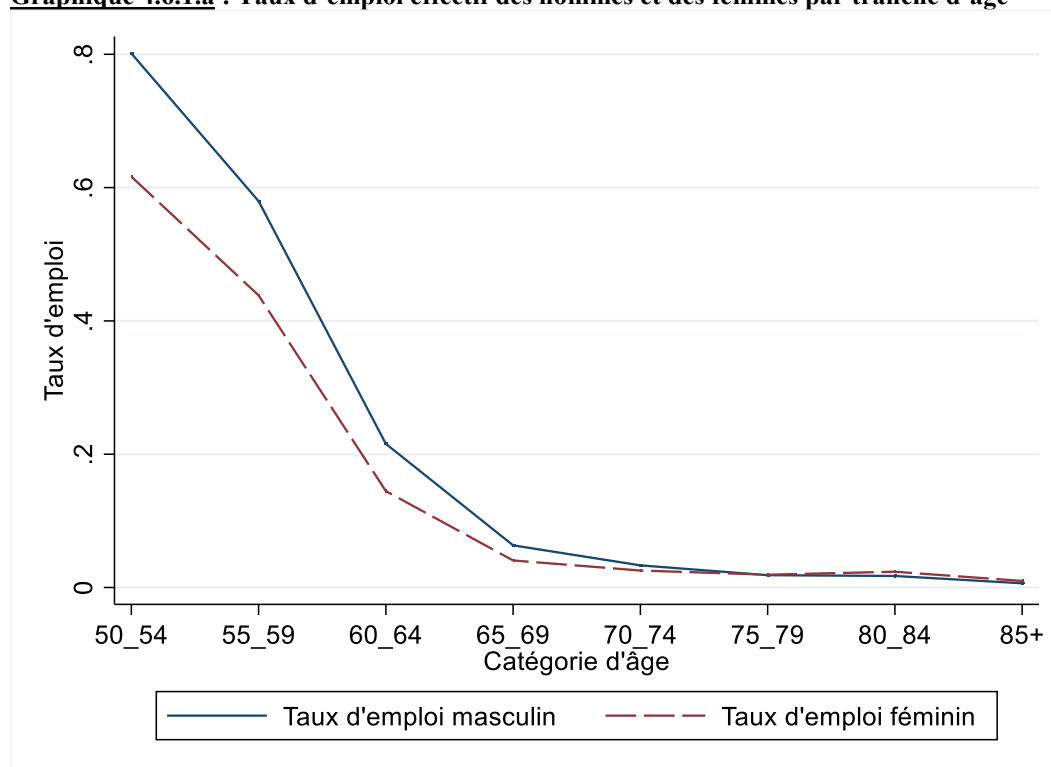
4.6. Les variables de contrôles

4.6.1. Le genre

La méthode de Cutler conduit à des résultats allant tous dans le même sens ; il existe une capacité à travailler additionnelle importante parmi la population retraitée. Néanmoins, le choix des variables et de la spécification du modèle conduit l'estimation de la capacité à travailler additionnelle vers des résultats légèrement différents.

Pour le moment, nous nous sommes concentrés sur des données de santé uniquement. Il existe pourtant d'autres facteurs influençant la santé et le travail ; omettre ces facteurs peut conduire à biaiser la relation estimée d'une manière similaire à celle que nous avons vue lors des modèles où on testait l'ajout ponctuel de variables objectives (une, puis deux, ...).

Graphique 4.6.1.a : Taux d'emploi effectif des hommes et des femmes par tranche d'âge



Quels sont ces autres facteurs influençant l'emploi et la santé ? Un premier facteur est le genre de la personne ; homme ou femme. Le marché de l'emploi des hommes et des femmes Belges est différent. Une simple observation des taux d'emploi féminin et masculin permet de le confirmer ; le taux d'emploi féminin est inférieur à celui des hommes pour toutes catégories d'âge avant 75 ans.

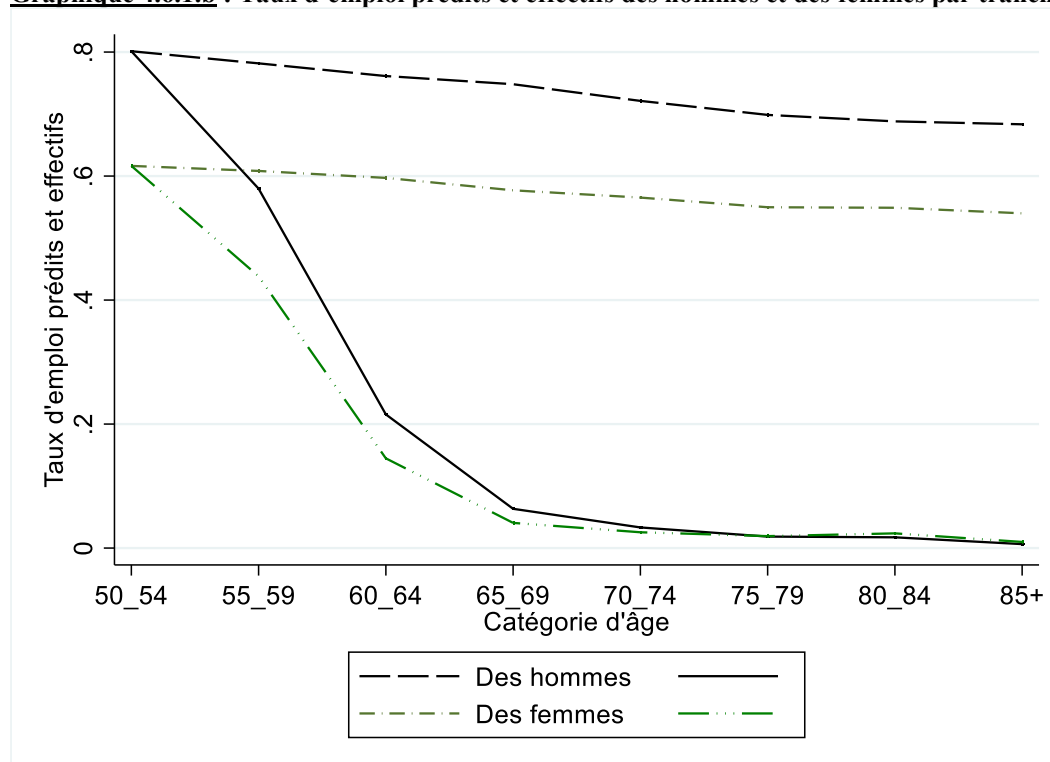
L'homme et la femme ne sont pas égaux sur le marché du travail. De plus, on peut facilement imaginer qu'il en est de même pour la santé générale. C'est deux arguments nous poussent à rechercher des relations emploi-santé différentes en fonction du genre et justifient donc une estimation séparée, par genre, de celles-ci. Le tableau suivant regroupe les résultats de la régression de l'emploi des individus de 50-54 ans sur l'unique indice de santé PVW index. La première colonne correspond à l'estimation de la relation emploi-santé où les hommes et les femmes sont regroupés et où on contrôle pour le genre. La seconde et la troisième colonne sont l'estimation de la relation respectivement pour les hommes et pour les femmes.

Tableau 4.6.1.a : Tableau de régressions de l'emploi des 50-54 ans sur nos différents PVW index

	Homme et femme	Homme	Femme
PVWindex-IVa	0.0031*** (0.0003)	0.0036*** (0.0004)	0.0026*** (0.0004)
Femme	-0.1542*** (0.0182)		
_cons	0.5949*** (0.0249)	0.5600*** (0.0327)	0.4682*** (0.0292)
Obs.	2423	1234	1189
R-squared	0.0858	0.0737	0.0289

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

Graphique 4.6.1.b : Taux d'emploi prédits et effectifs des hommes et des femmes par tranche d'âge



La première régression nous confirme que le taux d'emploi féminin est largement et significativement inférieur à celui des hommes tandis que les deuxièmes et troisièmes régressions nous informent que la relation emploi-santé est susceptible de changer en fonction du genre ; la relation est plus forte pour les hommes que pour les femmes (i.e. la santé des femmes semble avoir moins d'effet sur l'emploi que celle des hommes). Aussi, les variations de l'emploi expliquées par la santé sont plus de deux fois moins fortes pour les femmes que pour les hommes ce qui signifie que les autres facteurs influençant l'emploi ont plus d'importance pour les femmes que pour les hommes.

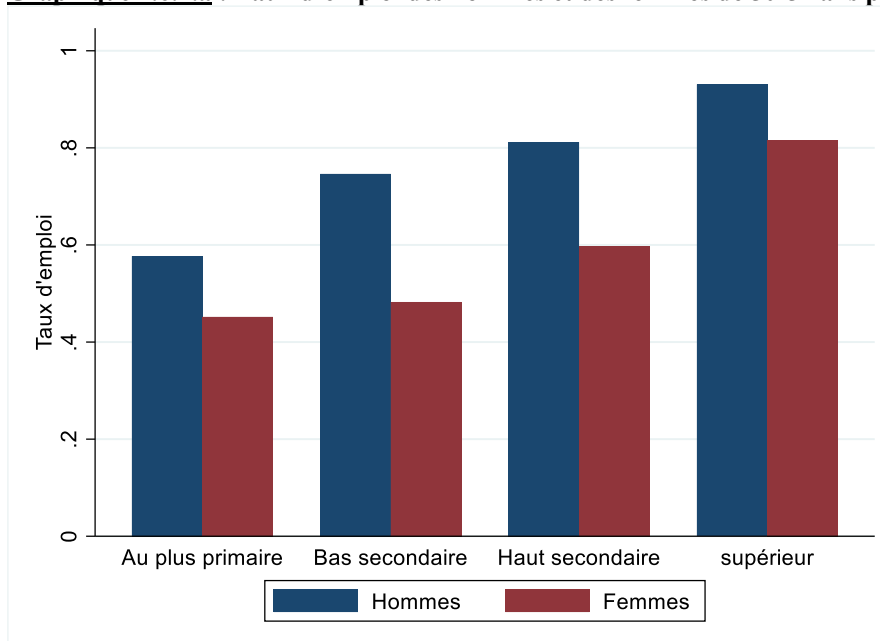
4.6.2. L'éducation

Une seconde source de variation potentiellement importante dans la relation emploi-santé se trouve dans les différences d'éducation. Il est intuitif de penser que l'emploi moyen partage une relation positive avec l'éducation ; un individu plus éduqué jouit, théoriquement, de plus d'opportunités d'emploi qu'un individu sans éducation.

C'est ce que le tableau ci-dessous montre ; le pourcentage d'homme de 50-54 ans au travail pour chaque niveau d'éducation est, respectivement du plus petit au plus grand niveau d'éducation, de 57.7%, 74.6%, 81.2% et 93%. Pour les femmes ces pourcentages sont de 45.2%, 48.3%, 59.8% et 81.6%. Le niveau « Bas secondaire » regroupe les individus qui ont démarré des études secondaires mais se sont arrêté en 1^{ère} ou en 2^{ème} année.

Aussi, il existe de nombreux liens entre l'éducation et la santé. D'une manière générale les individus mieux éduqués sont en meilleure santé et sont en meilleure condition pour combattre d'éventuels soucis de santé ; plus d'éducation signifie de meilleurs revenus (et donc de se faire mieux soigner), de meilleures opportunités d'emploi (travail intellectuel contre physique),...

Graphique 4.6.2.a : Taux d'emploi des hommes et des femmes de 50-54 ans par niveau d'éducation

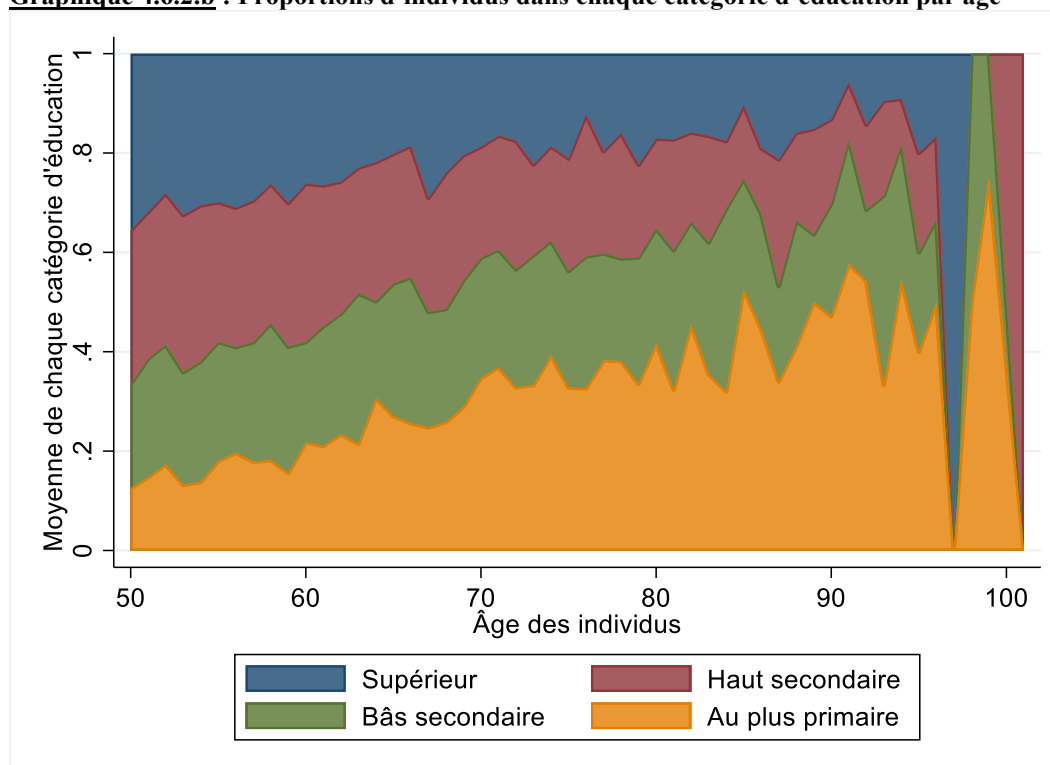


Enfin, la proportion d'individus par niveau d'éducation change en fonction de l'âge des individus. Nos individus jeunes sont en moyenne mieux éduqués que nos individus âgés puisqu'ils ont fait leur éducation plus récemment.

Si l'on devait prédire la capacité à travailler uniquement en fonction de l'éducation (donc simplement estimer la relation de l'emploi avec l'éducation pour faire nos prédictions) on verrait celle-ci décliner avec l'âge puisque l'éducation moyenne décline avec l'âge. Or ce déclin de la capacité à travailler ne sera pas dû à un déclin de santé mais plutôt au fait que, dans notre base de données, les individus âgés viennent de périodes antérieures. Autrement dit ; si quelqu'un reproduisait l'exercice dans 30 ans, ses individus de 80-84 ans seraient autant éduqués que notre génération de 50-54 ans actuellement. Alors, en supposant que l'éducation moyenne par tranche d'âge n'évolue plus d'ici là, il ne verra pas sa capacité à travailler décliner avec l'âge puisque l'éducation ne déclinera plus avec l'âge.

Ceci pour dire qu'inclure l'éducation en tant que contrôle pour estimer la capacité à travailler moyenne des hommes et des femmes (sans différencier la capacité à travailler moyenne par niveau d'éducation) entrainera une sous-estimation de celle-ci puisque l'éducation décline avec l'âge. Dès lors, lorsqu'on veut estimer la capacité à travailler en contrôlant l'éducation, il nous faudra également séparer chaque estimation de la relation emploi-santé par niveau d'éducation et faire nos prédictions et interprétations de la capacité à travailler par niveau d'éducation également.

Graphique 4.6.2.b : Proportions d'individus dans chaque catégorie d'éducation par âge



Le tableau suivant reprend les résultats des régressions où l'on contrôle pour l'éducation.

Il existe une relation claire entre le niveau d'éducation et l'emploi ; plus une personne est éduquée plus elle a de chance d'être au travail à 50-54 ans. Pour les hommes, avoir une éducation « Bas secondaire » offre des taux d'emplois moyen supérieur de 16.54 points de pourcentage par rapport au niveau « Primaire ». Les hommes ayant obtenu leur diplôme de secondaire ont 22 points de pourcentage en plus que le plus bas niveau d'éducation. Et le plus haut niveau d'éducation semble offrir aux hommes des taux d'emploi moyen supérieur de 32 points de pourcentage par rapport au niveau d'éducation le plus bas.

Pour les femmes la situation est légèrement différente ; l'éducation constitue toujours un facteur important du travail bien que pour elles, il n'y a statistiquement pas de différence entre avoir une éducation « primaire » et « bas secondaire ». Aussi, le diplôme de secondaire semble être nettement moins intéressant qu'il l'était pour les hommes (puisque'il n'augmente leur taux d'emploi que de 13.5 points de pourcentage par rapport au primaire). Enfin, l'éducation supérieure offre des taux d'emploi de 34.33 points de pourcentage en plus que le niveau d'éducation le plus bas.

Tableau 4.6.2.a : Régression de l'emploi des 50-54 sur la santé (par genre avec contrôle de l'éducation)

	Homme	Femme
PVWindex-Iva	0.0031*** (0.0004)	0.0020*** (0.0004)
Primaire (base)		
Bas secondaire	0.1642*** (0.0440)	0.0304 (0.0481)
Haut secondaire	0.2175*** (0.0412)	0.1347*** (0.0463)
Supérieur	0.3207*** (0.0386)	0.3433*** (0.0439)
_cons	0.3815*** (0.0436)	0.3442*** (0.0443)
Obs.	1234	1189
R-squared	0.1417	0.1041

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

Enfin il nous reste à discuter la relation entre l'emploi et la santé à 50-54 ans par niveau d'éducation ; rien ne garantit que cette relation soit similaire pour tout niveau d'éducation.

Pour les hommes, la relation emploi-santé est plus forte pour les 2 niveaux d'éducation du secondaire et est la moins forte pour ceux avec la meilleure éducation.

Etrangement, les femmes du premier niveau de secondaire possèdent la relation emploi-santé la plus la plus forte de toute avec un taux d'emploi moyen passant de 26.36% pour les femmes possédant une santé nulle à un taux d'emploi moyen de 66.36% pour celles possédant une santé parfaite.

Table 4.6.2.b : Régression de l'emploi des 50-54 sur la santé par niveau d'éducation (hommes)

	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0029*** (0.0011)	0.0038*** (0.0008)	0.0037*** (0.0008)	0.0021*** (0.0007)
_cons	0.3977*** (0.0750)	0.5024*** (0.0651)	0.5627*** (0.0605)	0.7747*** (0.0549)
Obs.	182	276	372	404
R-squared	0.0382	0.0796	0.0755	0.0548

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

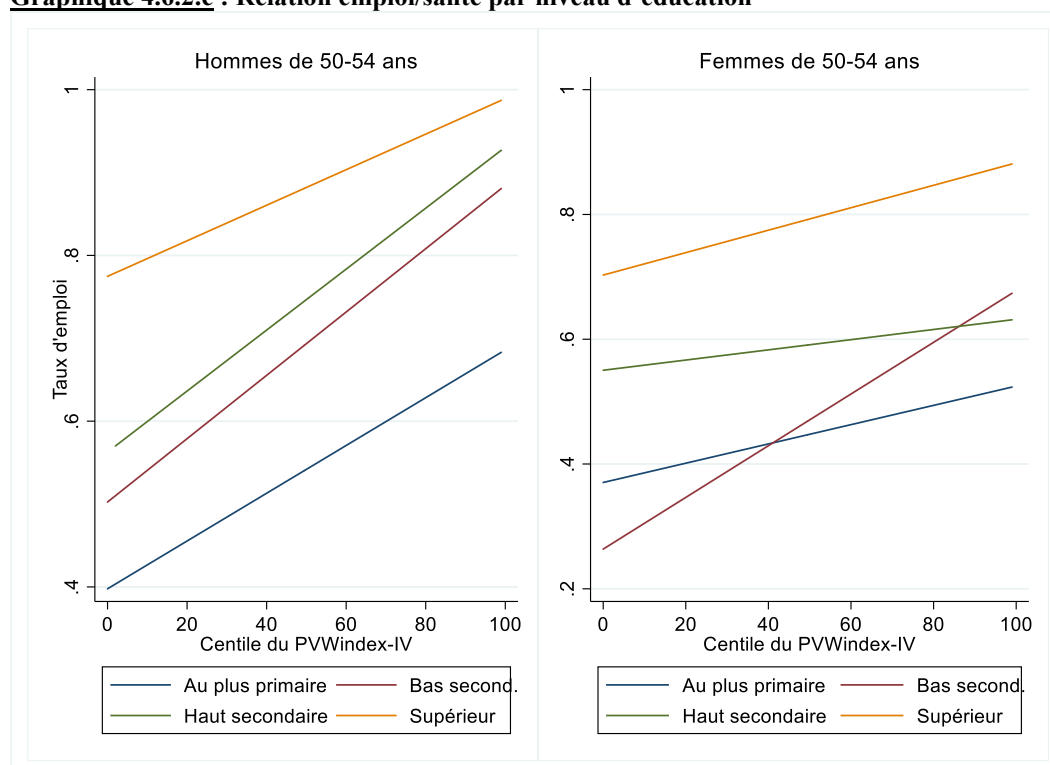
Table 4.6.2.c : Régression de l'emploi des 50-54 sur la santé par niveau d'éducation (femmes)

	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0015 (0.0011)	0.0041*** (0.0008)	0.0008 (0.0008)	0.0018** (0.0008)
_cons	0.3704*** (0.0698)	0.2636*** (0.0504)	0.5503*** (0.0523)	0.7028*** (0.0568)
Obs.	166	286	373	364
R-squared	0.0109	0.0757	0.0030	0.0177

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

Le graphique suivant représente graphiquement ces différentes relations emploi/santé des hommes et des femmes par niveau d'éducation.

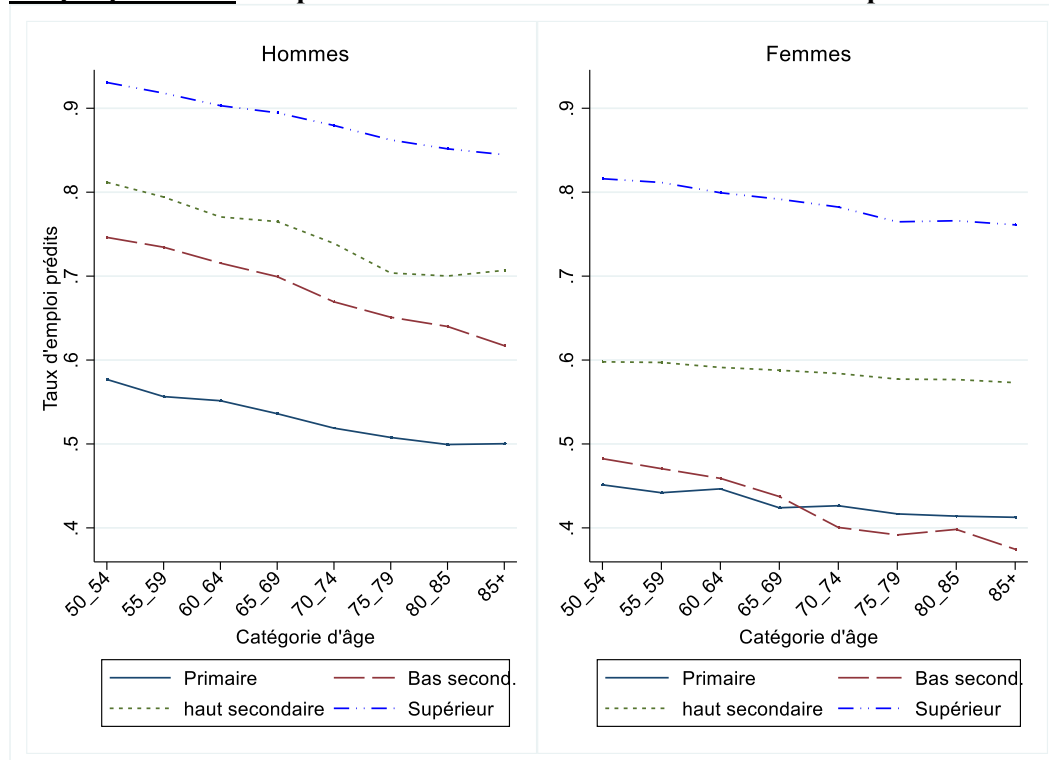
Graphique 4.6.2.c : Relation emploi/santé par niveau d'éducation



Pour les hommes ce graphique est sans surprise ; plus l'éducation et la santé sont bonnes, plus les taux d'emploi sont élevés. Pour les femmes on retrouve le fait étrange cité précédemment ; la forte relation emploi/santé des femmes de niveau de bas secondaire. Les femmes de santé pauvre de cette catégorie d'éducation possèdent des taux d'emploi inférieurs à ceux du niveau d'éducation primaire. Néanmoins, leur relation emploi/santé est si forte que leur taux d'emploi vient dépasser ceux du niveau d'éducation du haut secondaire pour les femmes en très bonne santé.

Enfin, on représente la capacité à travailler estimée par ces modèles pour chaque niveau d'éducation et par genre. Comme on s'y attend, la capacité à travailler estimée est très fortement dépendante du niveau d'éducation.

Graphique 4.6.2.d : Capacité à travailler des hommes et des femmes par niveau d'éducation

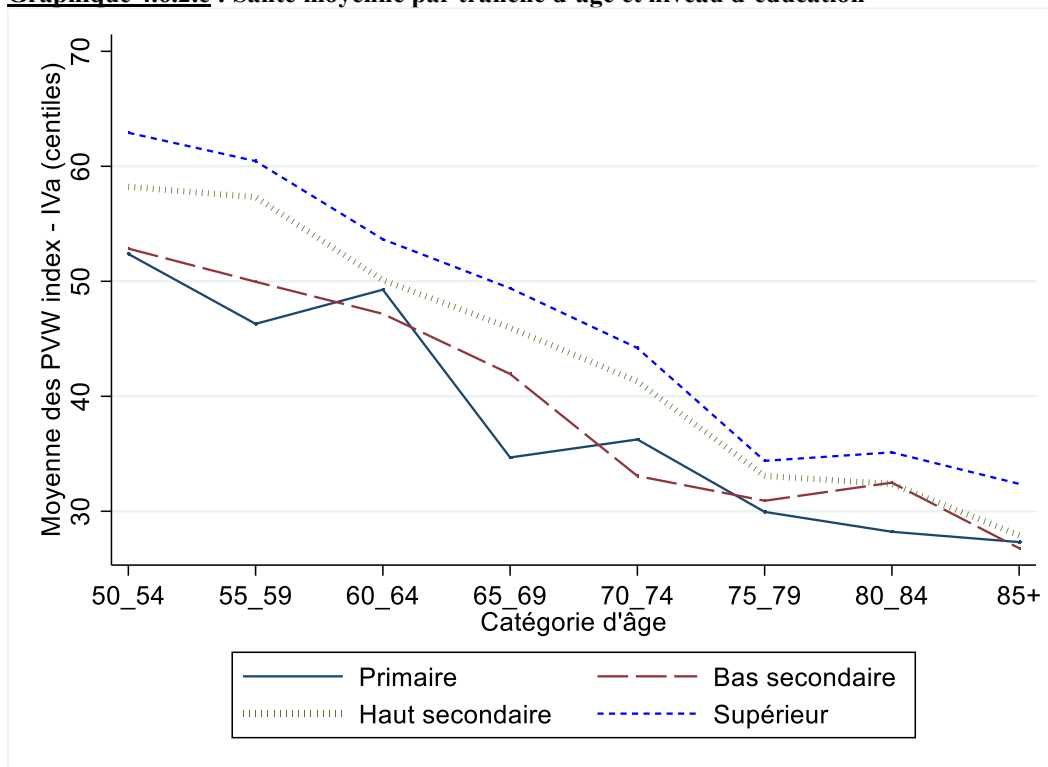


Les prédictions de la capacité à travailler en contrôlant pour l'éducation suivent exactement le raisonnement évoqué dans cette section ; la capacité à travailler est très dépendante du niveau d'éducation des individus principalement parce que le niveau d'éducation restreint les opportunités d'emploi.

Pour les hommes, les deux niveaux d'éducation du secondaire présentent le moins de différences de taux d'emploi prédits tandis que pour les femmes, il n'y a pratiquement pas de différence de capacité à travailler entre une éducation « primaire » et « bas secondaire ». Aussi, les femmes du niveau bas secondaire démarrent avec une capacité à travailler supérieure au niveau primaire mais finissent avec une capacité à travailler inférieure car elles dépendent plus de leur santé. Il est possible également que celles-ci voient leur santé se dégrader plus rapidement que celles du niveau primaire. Pour tester cette hypothèse nous représentons graphiquement la moyenne par âge des santés de chaque catégorie d'éducation dans le graphique suivant.

L'observation du graphique nous permet de dire que le déclin plus rapide de la capacité à travailler des femmes du bas secondaire est principalement dû à la forte dépendance relative (par rapport aux autres niveaux d'éducation) de leur emploi par rapport à leur santé. Aussi, on peut remarquer que le déclin de santé est assez similaire entre chaque niveau d'éducation, qu'une meilleure éducation signifie également une meilleure santé moyenne et que l'évolution de la santé des femmes du plus bas niveau d'éducation est plus erratique que les autres.

Graphique 4.6.2.e : Santé moyenne par tranche d'âge et niveau d'éducation



4.6.3. Les catégories d'emploi

Il nous reste encore une variable démographique importante pour estimer la capacité à travailler selon la méthode de Cutler ; la catégorie d'emploi. Tous les emplois ne sont pas égaux en terme de besoins de santé pour être capable de travailler ; une personne travaillant derrière un bureau sera sans doute moins gênée par une jambe cassée qu'un joueur de foot.

Notre base de données sépare nos individus en dix catégories d'emploi différentes mais on ne pourra en étudier la relation emploi/santé que pour neuf puisque la dernière, celle des militaires, souffre d'un manque d'observations.

On commence par inclure les types d'emploi en contrôle à notre dernière régression. On peut voir dans les régressions suivantes que les catégories d'emploi offriront des opportunités d'emplois parfois largement différents l'une de l'autre. Ce qui nous intéresse c'est de vérifier si contrôler pour le type d'emploi change les résultats de l'effet de la santé sur l'emploi par niveau d'éducation calculé dans les tableaux de régressions précédents où on ne les contrôlait pas. Ils changent très peu ; ce qui montre que l'éducation joue un rôle important dans la

relation emploi/santé indépendamment du type d'emploi.

Tableau 4.6.3.a : Régression de l'emploi des 50-54 sur la santé par niveau d'éducation en contrôlant pour le type d'emploi.

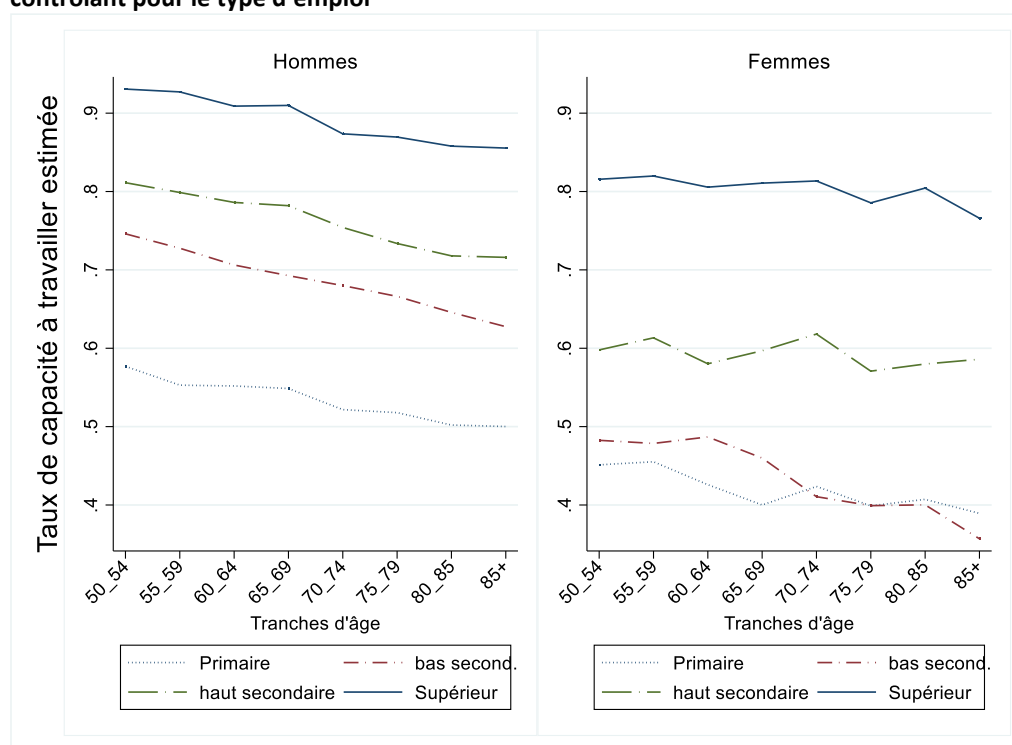
	Hommes				Femmes			
	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0030*** (0.0010)	0.0035*** (0.0009)	0.0035*** (0.0008)	0.0019*** (0.0006)	0.0016 (0.0011)	0.0040*** (0.0008)	0.0009 (0.0008)	0.0015** (0.0008)
Managers								
Professionnels	0.3628 (0.3285)	-0.2951 (0.1989)	-0.0648 (0.0753)	0.0578 (0.0357)	0.1975 (0.3754)	-0.4381*** (0.1493)	0.1026 (0.1146)	-0.0415 (0.0589)
Techn./Associates	-0.2714 (0.4584)	-0.0727 (0.1315)	-0.0836 (0.0657)	0.0449 (0.0535)	-0.0426 (0.3538)	-0.4174** (0.1797)	-0.0922 (0.1217)	-0.2451*** (0.0851)
Employé de bureau	0.1257 (0.3434)	-0.1930* (0.1105)	-0.0662 (0.0520)	0.0733* (0.0414)	0.3219 (0.3120)	-0.3717*** (0.0808)	-0.0212 (0.1025)	-0.1283* (0.0731)
Services et ventes	-0.0496 (0.3430)	-0.0840 (0.0984)	-0.1616*** (0.0590)	0.0909*** (0.0334)	0.1253 (0.2834)	-0.4440*** (0.0775)	-0.1753* (0.1057)	-0.2240*** (0.1139)
Agri, forêt, pêche	-0.0378 (0.3859)	0.0769 (0.0853)	-0.1165 (0.1264)	-0.2849 (0.2780)	0.4854 (0.3232)	-0.1047* (0.0561)	0.3669*** (0.1028)	0.0513 (0.0542)
Artisans et commerçants	-0.2000 (0.3325)	-0.2134** (0.1046)	-0.1489** (0.0622)	-0.2232* (0.1344)	-0.1007 (0.2879)	-0.6714*** (0.1058)	-0.2833* (0.1446)	
Opérateur/assembleurs	0.0080 (0.3341)	-0.0966 (0.0978)	-0.2695*** (0.0984)	0.0140 (0.0995)	-0.0193 (0.2945)	-0.4410*** (0.1086)	-0.5047*** (0.1481)	-0.4712** (0.2275)
Métiers élémentaires	-0.1766 (0.3437)	-0.1620 (0.1128)	-0.2009* (0.1120)	-0.3585** (0.1693)	0.2423 (0.2800)	-0.4497*** (0.0827)	-0.0606 (0.1211)	-0.3460* (0.1801)
Militaires	0.2988 (0.3277)	0.0659 (0.0819)	0.0674 (0.0561)	0.1045** (0.0521)		-0.7986*** (0.0668)	0.2813*** (0.0942)	
_cons	0.4728 (0.3281)	0.6538*** (0.1009)	0.6873*** (0.0703)	0.7601*** (0.0597)	0.2216 (0.2805)	0.7112*** (0.0798)	0.6286*** (0.1045)	0.8236*** (0.0753)
Obs.	182	276	372	404	166	286	373	364
R-squared	0.1132	0.1157	0.1086	0.1584	0.0925	0.1066	0.0673	0.0822

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

Les estimations de capacité à travailler (ci-dessous) sur base de ces régressions révèlent des différences significatives à ce que nous avons trouvé pour les femmes lorsqu'on ne contrôlait pas pour le type d'emploi.

Les femmes de niveau d'éducation supérieur ou égale au secondaire subissent en réalité moins d'effet de la santé sur l'emploi. Tandis que les femmes de niveau de bas secondaire ne voient plus leur capacité à travailler descendre sous celles de niveau primaire avec l'âge ; leur capacité à travailler vient plutôt **rejoindre** celui des femmes de niveau primaire au plus. Ce dernier résultat nous confirme que l'éducation et les types d'emploi partagent des interactions non négligeables et nous prouve que ne pas contrôler pour le type d'emploi vient biaiser (légèrement et pour certains seulement) la capacité à travailler estimée par niveau d'éducation.

Graphique 4.6.3.a : Capacité à travailler estimée des hommes et des femmes par niveau d'éducation en contrôlant pour le type d'emploi



On décide aussi d'observer plus précisément ces différences d'opportunité d'emploi par catégorie d'emploi avec l'aide des régressions 4.6.3b et 4.6.3c.

Tableau 4.6.3.b : Régression de l'emploi des 50-54 sur la santé par type d'emploi (hommes)

	Managers	Professio.	Techn. et professi. associés	Employés de bureau	Service et vente	Agri. Forêt et pêche	Artisans, ouvriers de métier	Opérat., assembl.	Métiers élément.
PVWindex-IVa	0.0017 (0.0012)	0.0015* (0.0008)	0.0021* (0.0012)	0.0037*** (0.0010)	0.0064*** (0.0011)	0.0039 (0.0029)	0.0026** (0.0011)	0.0047*** (0.0012)	0.0033* (0.0016)
_cons	0.7821*** (0.0973)	0.8228*** (0.0651)	0.6986*** (0.1014)	0.6056*** (0.0811)	0.3601*** (0.0963)	0.5552** (0.2414)	0.4606*** (0.0784)	0.4336*** (0.0896)	0.4202*** (0.1241)
Obs.	143	221	88	177	157	31	198	132	80
R-squared	0.0247	0.0287	0.0374	0.1000	0.2190	0.0904	0.0290	0.1110	0.0495

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

Tableau 4.6.3.c : Régression de l'emploi des 50-54 sur la santé par type d'emploi (femmes)

	Managers	Professio.	Techn. et professi. associés	Employés de bureau	Service et vente	Agri. Forêt et pêche	Artisans, ouvriers de métier	Opérat., assembl.	Métiers élément.
PVWindex-IVa	0.0057*** (0.0019)	0.0007 (0.0008)	0.0019 (0.0016)	0.0025*** (0.0009)	0.0018* (0.0010)	0.0034 (0.0040)	0.0022 (0.0017)	0.0006 (0.0017)	0.0037*** (0.0011)
_cons	0.4390*** (0.1399)	0.7994*** (0.0571)	0.5031*** (0.1147)	0.5163*** (0.0638)	0.3956*** (0.0612)	0.6769 (0.3577)	0.1640* (0.0971)	0.3497*** (0.1105)	0.3386*** (0.0688)
Obs.	56	241	100	250	238	8	60	65	169
R-squared	0.1640	0.0032	0.0135	0.0302	0.0142	0.1064	0.0270	0.0021	0.0599

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

Ces régressions nous montrent plus précisément les différences d'opportunités d'emploi par catégories d'emploi. Pour les hommes comme pour les femmes la catégorie d'emploi qui offre les meilleures opportunités d'emploi est celle des professionnels avec des taux d'emploi de base proche des 80%. Parmi les catégories offrant le moins d'opportunités d'emploi on retrouve, chez les hommes, celle des services et ventes. Pour les femmes la situation est différente ; celles travaillant dans le secteur des artisans et ouvriers de métiers démarrent avec des taux d'emploi de base ridiculement bas (16.40%). On peut également noter que les secteurs des métiers élémentaires, opérateurs/assembleurs et Services et ventes présentent des taux d'emploi de base inférieur à 40%.

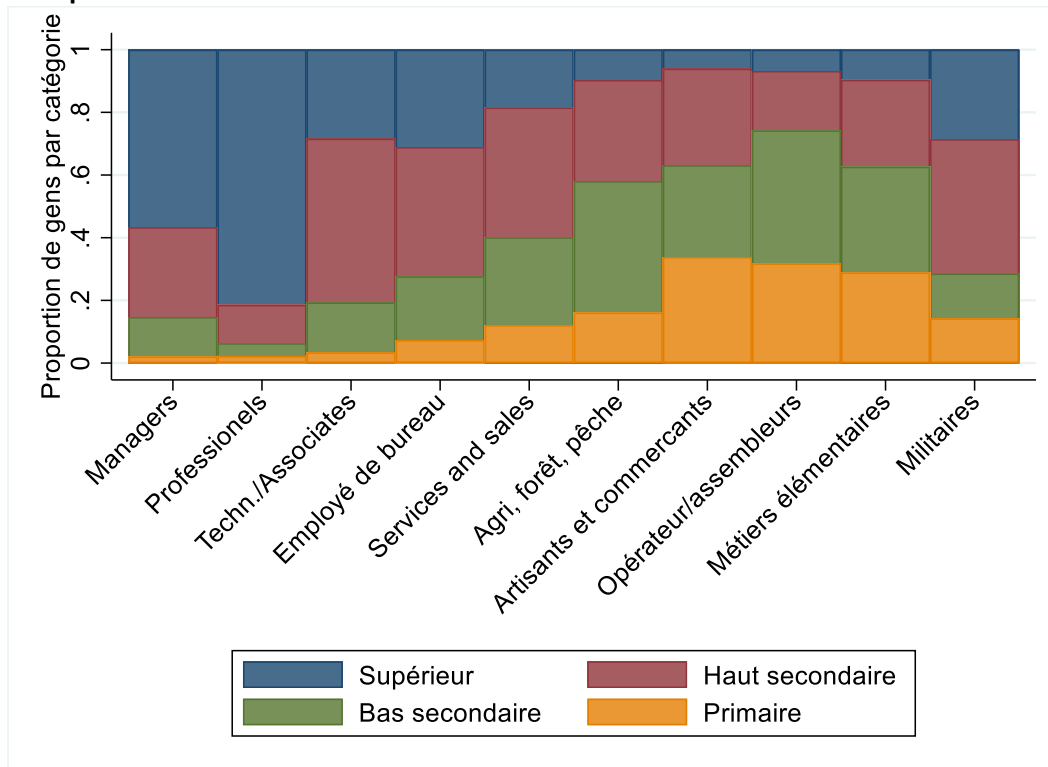
Concernant la dépendance à la santé pour travailler dans chacune de ces catégories, on retrouve, là encore, une forte hétérogénéité ; les hommes du secteur des services et des ventes voient des différences de taux d'emploi de 64 points de pourcentage entre ceux de santé parfaite et ceux de la pire santé. Pour les opérateurs/ assembleurs c'est 47 points de pourcentage et pour les employés de bureau on est à 37 points de pourcentage de différence. Pour les femmes c'est la catégorie des managers qui semble dépendre le plus de sa santé pour travailler avec une différence de 57 points de pourcentage entre celles en meilleure et en pire santé.

Comme on ne contrôle pas pour l'éducation, ces différences n'en sont peut-être que le reflet ; il est possible que certaines catégories d'emploi cumulent des personnes de certains niveaux d'éducation et que ce qu'on observe n'est pas tant des différences entre catégorie d'emploi mais plutôt des différences entre l'éducation des personnes travaillant dans chacune de ces catégories. On décide donc d'illustrer cette idée en représentant graphiquement le pourcentage de personnes de 50-54 ans de chaque niveau d'éducation par type d'emploi dans les graphiques 4.6.3.a et 4.6.3.b.

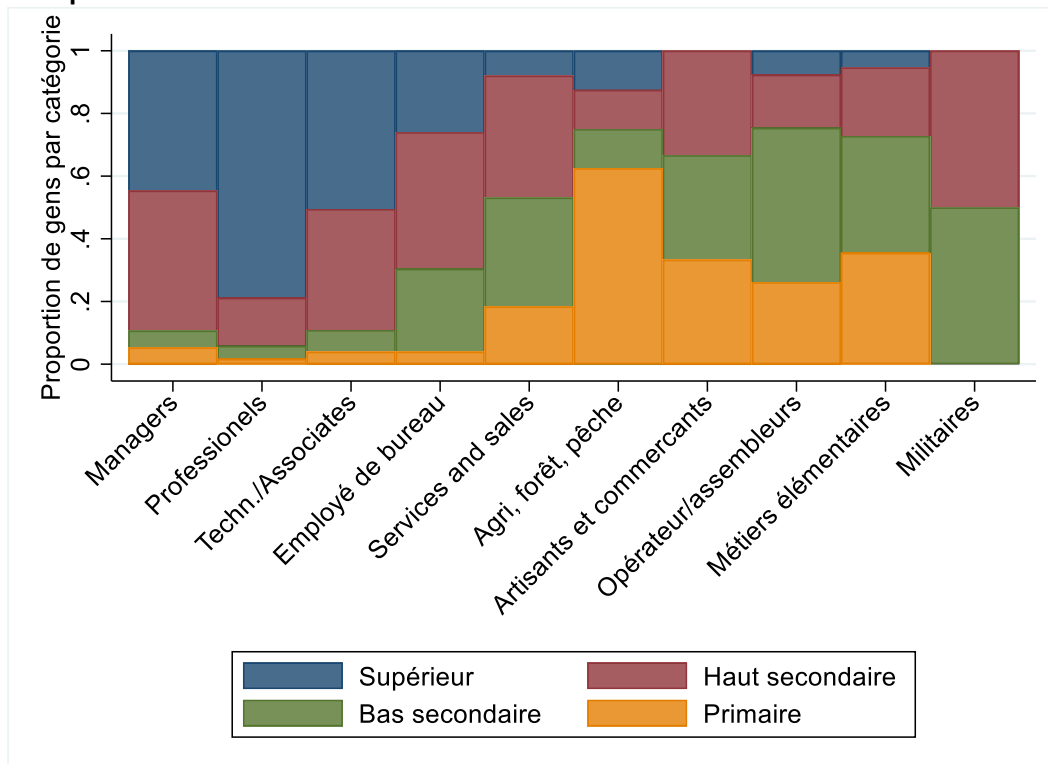
L'étude du premier graphique nous apprend que les catégories des managers et professionnels sont composées en majorité de personnes possédant un niveau d'enseignement supérieur tandis que les catégories de l'agriculture, forêt et pêche, des artisans et commerçants, des opérateurs et assembleurs et des métiers élémentaires sont composées en majorité par des individus n'ayant pas terminé l'enseignement secondaire.

Ce constat s'applique aux femmes de manière plus ou moins similaire ; les catégories des managers et professionnels sont composées, en grande partie, de personnes possédant un niveau d'enseignement supérieur auxquelles on peut ajouter la catégorie des techniciens et professionnels associés. Tandis que les catégories cumulant une majorité de personnes ne possédant pas de diplôme de secondaire sont, comme pour les hommes, celles de l'agriculture, forêt et pêche, des artisans et commerçants, des opérateurs et assembleurs et des métiers élémentaires auxquelles vient s'ajouter la catégorie des services et ventes. Aussi, on peut remarquer le nombre inquiétant de femmes ne possédant qu'un certificat d'étude de base (Au plus primaire) dans la catégorie de l'agriculture, forêt et pêche.

Graphique 4.6.3.b : Proportion d'hommes de 50-54 ans de chaque niveau d'éducation par catégorie d'emploi



Graphique 4.6.3.c : Proportion de femmes de 50-54 ans de chaque niveau d'éducation par catégorie d'emploi



Intéressons-nous désormais à la relation emploi/santé par type d'emploi en contrôlant pour l'éducation. Pour les hommes, la santé semble avoir un effet statistiquement significatif à plus de 95% sur l'emploi pour les 4 mêmes catégories d'emploi qu'auparavant ; donc il semblerait que le type d'emploi ait son importance dans la relation emploi/santé des individus indépendamment de l'éducation. Tout comme avant, les hommes travaillant dans les secteurs des services et des ventes sont ceux dont la santé a le plus d'effet sur leur emploi bien qu'il soit légèrement plus faible ; il y aura, en moyenne, une différence de taux d'emploi de 60% entre un individu de santé parfait et un individu de la pire santé. La catégorie d'emploi la moins dépendante de sa santé pour travailler est celle des professionnels.

Tableau 4.6.3.d : Régression de l'emploi des hommes de 50-54 ans sur la santé par catégorie d'emploi

	Managers	Professionnels	Techniciens et professionnel associé	Employés de bureau	Service et vente	Ouvriers qualifiés de l'agriculture sylviculture et pêche	Artisans et ouvriers de métier	Opérateurs d'installations et de machines et assembleurs	Métiers élémentaires
PVWindex-IVa	0.0016 (0.0012)	0.0012 (0.0008)	0.0014 (0.0011)	0.0034*** (0.0010)	0.0060*** (0.0011)	0.0037 (0.0025)	0.0023** (0.0011)	0.0047*** (0.0012)	0.0030* (0.0017)
Primaire									
Bas second.	0.2153 (0.3115)	-0.4478*** (0.1715)	0.4307 (0.3182)	-0.1141 (0.1298)	0.1664 (0.1093)	0.3115 (0.2147)	0.1871** (0.0883)	0.1018 (0.0888)	0.2254* (0.1342)
Haut second.	0.2656 (0.3056)	-0.1614** (0.0718)	0.4651 (0.3033)	0.0500 (0.1125)	0.1423 (0.1050)	0.1581 (0.2537)	0.3050*** (0.0820)	-0.0379 (0.1151)	0.2200 (0.1509)
Supérieur	0.2194 (0.3050)	-0.0621** (0.0294)	0.5761* (0.2994)	0.1564 (0.1086)	0.3164*** (0.0986)	-0.1010 (0.3694)	0.1934 (0.1432)	0.2302** (0.1093)	0.0486 (0.1942)
_cons	0.5630* (0.3011)	0.9376*** (0.0454)	0.2739 (0.3058)	0.5799*** (0.1302)	0.2184* (0.1212)	0.3972 (0.2789)	0.3223*** (0.0883)	0.3781*** (0.1062)	0.2971* (0.1547)
Obs.	143	221	88	177	157	31	198	132	80
R-squared	0.0435	0.1322	0.1146	0.1702	0.2673	0.2395	0.0943	0.1388	0.0962

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

Aussi, l'effet positif de l'éducation sur l'emploi n'est plus aussi évident qu'auparavant ; par exemple, on peut observer un coefficient négatif (mais non significatif) au plus haut niveau d'éducation dans la régression n'incluant que les ouvriers qualifiés de l'agriculture, de la sylviculture et de la pêche. Ce résultat est à prendre avec précaution au vu du nombre restreint d'observations lors de la régression. Un souci similaire apparaît dans la régression du secteur des professionnels ; très peu de personnes ne possèdent qu'une éducation « au plus primaire » ce qui explique les valeurs négatives aux estimateurs des niveaux plus élevés.

Chez les femmes c'est également l'effet de la santé sur l'emploi qui n'est plus aussi évident puisque celui-ci n'est largement significatif (99%) que pour 2 des 9 catégories d'emploi présentées ici. La catégorie d'emploi qui semble être la plus dépendante de sa santé pour être capable de travailler est celle de l'agriculture, de la pêche et de la sylviculture puisque l'augmentation (diminution) d'un centile de santé entraîne une augmentation (diminution) de 0.83 points de pourcentage du taux d'emploi. Or, les individus de cette catégorie possédant une santé parfaite et se trouvant donc dans le 100^{ième} centile devraient alors avoir un taux

d'emploi de 116% en moyenne ($0.33+0.83*100$) ce qui est incohérent. Ce phénomène révèle une des limites de la régression par moindres carrés ordinaires (MCO) qui apparaît lorsque la variable dépendante est binaire ; elle ne contraint pas ses prédictions de probabilité entre 0 et 1. On aurait pu régler ce problème en utilisant une régression logistique qui limite bel et bien les probabilités prédites entre 0 et 1 mais on privilégie les MCO dans ce travail notamment parce que l'interprétation des coefficients est bien plus simple. Aussi, dans cette régression particulière, nous sommes face à un cas spécial ; un manque d'observation flagrant puisque nous n'en disposons que de 8. Aussi, 7 de ces 8 femmes sont en emploi (87,5%) ce qui pousse nos estimateurs à surestimer l'effet de la santé sur l'emploi (on suppose qu'avec plus d'observations on aurait un taux d'emploi moyen plus faible pour cette catégorie).

Tableau 4.6.3.e : Régression de l'emploi des femmes de 50-54 ans sur la santé par catégorie d'emploi

	Managers	Professionels	Techniciens et professionnels associés	Employés de bureau	Services et vente	Ouvriers qualifiés de l'agriculture sylviculture et pêche	Artisans et ouvriers de métier	Opérateurs d'installations et de machines et assembleurs	Métiers élémentaires
PVWindex-IVa	0.0055*** (0.0018)	0.0005 (0.0008)	0.0015 (0.0017)	0.0021** (0.0010)	0.0018* (0.0010)	0.0083 (0.0112)	0.0019 (0.0017)	0.0008 (0.0016)	0.0035*** (0.0011)
Primaire									
Bas second.	0.6408** (0.2712)	-0.0037 (0.2970)	0.2821 (0.2936)	-0.0908 (0.1594)	0.0404 (0.0952)	-0.1509 (0.2555)	0.0410 (0.1307)	0.2043 (0.1468)	-0.0812 (0.0889)
Haut second.	0.4105 (0.2711)	0.2797 (0.2614)	0.3169 (0.2397)	0.0105 (0.1550)	0.0645 (0.0941)	0.6271 (0.8157)	0.1774 (0.1509)	-0.1185 (0.1657)	0.0487 (0.1007)
Supérieur	0.6177** (0.2654)	0.3725 (0.2535)	0.3968* (0.2372)	0.1305 (0.1576)	0.2478* (0.1260)	-0.0019 (0.0967)		0.1261 (0.2552)	0.0126 (0.1891)
_cons	-0.0434 (0.2777)	0.4765* (0.2551)	0.1860 (0.2353)	0.5284*** (0.1530)	0.3364*** (0.0966)	0.3316 (0.8713)	0.1089 (0.1213)	0.2491* (0.1369)	0.3646*** (0.0864)
Obs.	56	241	100	250	238	8	60	65	169
R-squared	0.3074	0.0647	0.0412	0.0581	0.0286	0.3279	0.0547	0.0696	0.0704

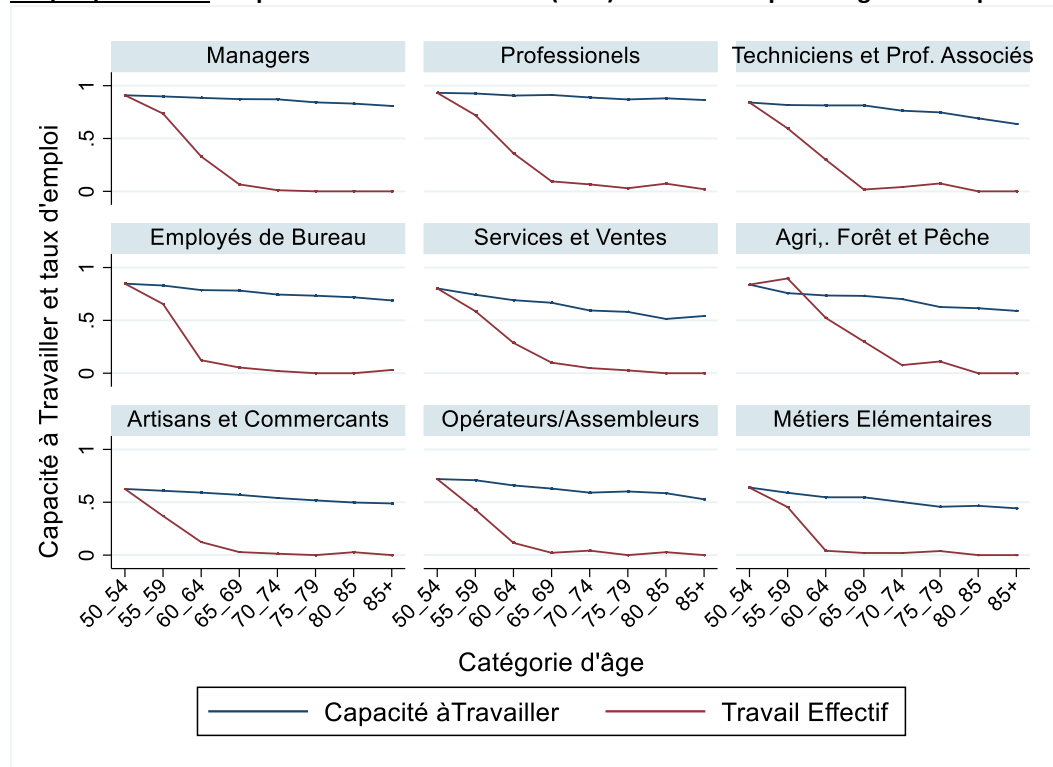
*** p<0.01, ** p<0.05, * p<0.1

En réalité, estimer et présenter la capacité à travailler des femmes pour cette catégorie d'emploi n'est pas vraiment pertinent mais on les présente quand même car elles nous permettent d'introduire la limite des MCO lorsque la variable dépendante est binaire. Enfin, d'une manière générale, les résultats obtenus par Moindres Carrés Ordinaires sont très proches de ceux obtenus par des régressions logistiques. D'autant plus qu'ici, nous présentons la capacité à travailler prédite par moyenne de tranche d'âge (on fait une moyenne des prédictions d'un groupe d'observation) et non pas des probabilités individuelles d'être en emploi.

Pour les hommes, il existe une capacité à travailler supplémentaire pour toutes les catégories d'emploi et pour tous les âges sauf pour ceux de 55-59 ans travaillant dans les domaines de l'agriculture, forestier et pêche. Pour ces individus en particulier, la méthode de Cutler prédit une capacité à travailler en dessous de leur taux d'emploi effectif. Ce résultat s'explique par un taux d'emploi effectif à 55-59 ans étrangement supérieur aux taux d'emploi de cette catégorie à 50-54 ans.

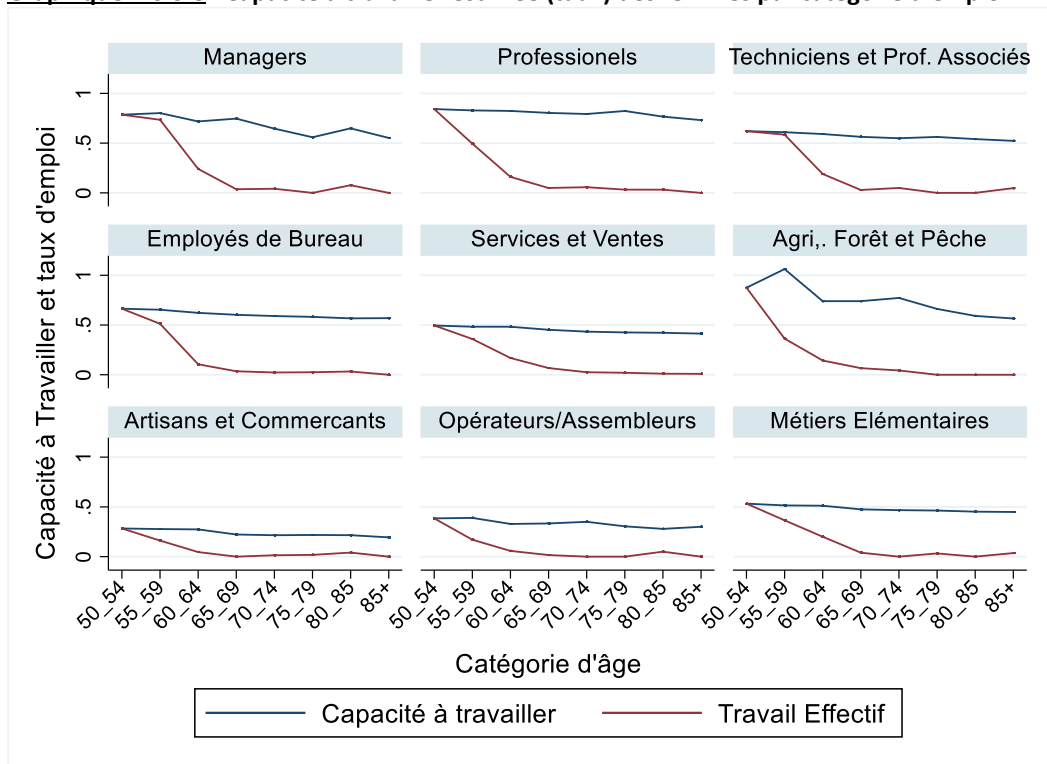
Enfin, la capacité à travailler estimée des femmes en régressant par catégorie d'âge suit toujours les résultats obtenus jusqu'ici ; il existe une capacité à travailler supplémentaire importante parmi les femmes âgées bien que celle-ci dispose de moins d'opportunités d'emploi à la base.

Graphique 4.6.3.d : Capacité à travailler estimée (taux) des hommes par catégorie d'emploi



On peut remarquer le résultat étrange évoqué précédemment ; une capacité à travailler dépassant les 100% pour les femmes de 55-59 ans de la catégorie agriculture, sylviculture et pêche. On en présente les prédictions de capacité à travailler ici pour montrer la limite des MCO lorsque la variable dépendante est dichotomique bien qu'on aurait pu (et peut être dû) ne pas les présenter du tout sous critère d'un manque d'observation trop important conduisant à des estimations trop imprécises.

Graphique 4.6.3.e : Capacité à travailler estimée (taux) des femmes par catégorie d'emploi



4.6.4. Les régions

Nous disposons de données sur la région d'habitation de nos individus ; est-ce que la capacité à travailler estimée ici diffère entre les régions ? Et est-ce que les relations d'emploi/santé changent lorsqu'on contrôle pour la région de résidence ? C'est le but de cette dernière partie sur la mise en pratique de la méthode de Cutler.

On présente d'abord les résultats de la régression lorsqu'on contrôle pour la région (tableau 4.6.4.a) pour définir si oui ou non il existe des différences significatives entre les régions. On calculera ensuite chaque relation emploi/santé par genre, éducation et région pour observer si la relation emploi/santé par éducation diffère entre les différentes régions.

La régression nous informe que pour les femmes, il n'y a statistiquement pas de différences dans la relation estimée entre la Flandre et la Wallonie et entre la Flandre et Bruxelles lorsqu'on observe uniquement les femmes bien qu'il semble y avoir des différences entre les régions lorsqu'on observe la relation pour les hommes. Les coefficients associés aux différents niveaux d'éducation vont toujours dans le sens évoqué précédemment (sauf pour les femmes du bas secondaire) ; l'éducation offre de meilleurs résultats en termes d'emploi à 50-54 ans et entraîne donc des estimations de la capacité à travailler à des âges avancés plus élevés de base.

Tableau 4.6.4.a : Régression de l'emploi de 50-54 ans par genre avec toutes les variables de contrôles

	Hommes		Femmes	
	Coef.	St.Err.	Coef.	St.Err.
PVWindex-Iva	0.0027***	0.0004	0.0020***	0.0004
Primaire	0.0000	.	0.0000	.
Bas secondaire	0.1268***	0.0441	0.0175	0.0483
Haut secondaire	0.1644***	0.0433	0.0839*	0.0490
Supérieure	0.2268***	0.0441	0.2097***	0.0539
Managers	0.0000	.	0.0000	.
Professionnels	0.0104	0.0315	0.0057	0.0563
Techn./Associates	-0.0351	0.0450	-0.176**5	0.0697
Clerical support	-0.0238	0.0351	-0.0824	0.0593
Services and sales	-0.0507	0.0378	-0.2027***	0.0632
Agri, forêt, pêche	-0.0179	0.0705	0.1878	0.1242
Artisans et commerçants	-0.1773***	0.0445	-0.3909***	0.0815
Opérateur/assembleurs	-0.0751	0.0485	-0.3012***	0.0836
Métiers élémentaires	-0.1755***	0.0597	-0.1438**	0.0683
Militaires	0.1133**	0.0483	-0.1950	0.2807
Brabant flammand	0.0000	.	0.0000	.
Brabant Wallon	-0.0675***	0.0238	-0.0188	0.0309
Bruxelles	-0.0930***	0.0289	0.0087	0.0341
Constant	0.5749***	0.0573	0.5413***	0.0773
Mean dependent var	0.8015		Mean dependent var	0.6165
R-squared	0.1660		R-squared	0.1440
Prob > F	0.0000		Prob > F	0.0000
Number of obs	1234.0000		Number of obs	1189.0000

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

Intéressons-nous désormais à la régression par région (tableau 4.6.4.b) ce qui nous permettra de déterminer si les relations estimées diffèrent entre les régions. La santé a un effet significatif sur l'emploi dans toutes les régions bien que cet effet semble plus fort à Bruxelles et plus faible en région flamande par rapport à la région wallonne.

Les coefficients associés aux différentes catégories d'emploi sont presque tous non différenciables de 0 (non-significatifs) et ceux qui le sont (différenciables de 0) sont tous négatifs sauf pour les femmes travaillant dans les secteurs de l'agriculture, forestier et pêche pour les raisons déjà évoquées précédemment. Le fait que ceux-ci soient presque tous négatifs révèle que la catégorie « managers » offre, en moyenne, des taux d'emploi plus élevés.

Tableau 4.6.4.b : Régression de l'emploi de 50-54 ans par genre et par région avec contrôle de l'éducation et des catégories d'emploi

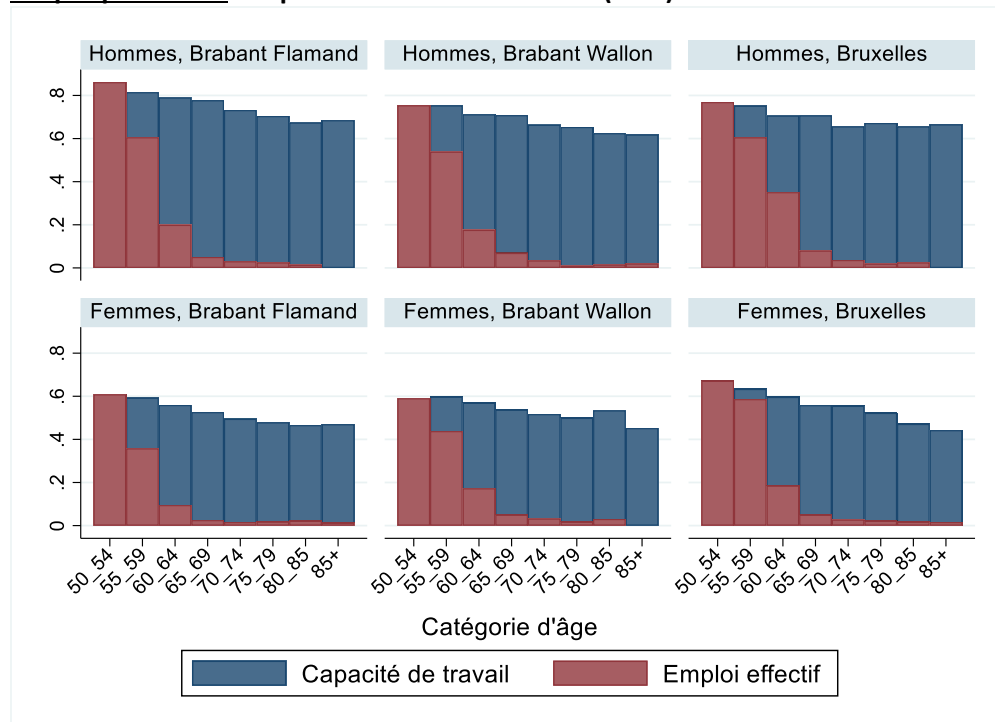
	Hommes			Femmes		
	Région flamande	Région wallonne	Bruxelles	Région flamande	Région wallonne	Bruxelles
PVWindex-IVa	0.0023*** (0.0006)	0.0026*** (0.0007)	0.0035*** (0.0009)	0.0014* (0.0007)	0.0021*** (0.0007)	0.0031*** (0.0009)
b. Primaire						
Bas secondaire	0.1905*** (0.0658)	0.1168 (0.0713)	-0.0001 (0.1171)	0.0375 (0.0773)	0.0179 (0.0746)	0.0134 (0.1238)
Haut secondaire	0.2037*** (0.0668)	0.1702** (0.0712)	0.0768 (0.0990)	0.0791 (0.0757)	0.0912 (0.0775)	0.1023 (0.1262)
Supérieure	0.2452*** (0.0654)	0.2850*** (0.0730)	0.1095 (0.0989)	0.2756*** (0.0878)	0.1162 (0.0858)	0.2522* (0.1307)
b. Managers						
Professionels	0.0414 (0.0424)	-0.0246 (0.0608)	-0.0135 (0.0649)	-0.0619 (0.0871)	0.0312 (0.0940)	0.0761 (0.1115)
Techn./Associates	-0.0313 (0.0650)	-0.0991 (0.0753)	0.1244 (0.1133)	-0.2195* (0.1137)	-0.1675 (0.1156)	-0.0952 (0.1337)
Clerical support	0.0099 (0.0442)	-0.0613 (0.0720)	-0.0131 (0.0833)	-0.0948 (0.0968)	-0.1191 (0.0972)	0.0068 (0.1169)
Services and vente	0.0190 (0.0490)	-0.0638 (0.0702)	-0.1496 (0.0912)	-0.1979* (0.1007)	-0.2552** (0.1030)	-0.1302 (0.1286)
Agri, forêt, pêche	0.0630 (0.0771)	-0.0386 (0.1220)	-0.2520 (0.3346)	0.0819 (0.2284)	0.2522** (0.1018)	
Artisans et commerçants	-0.1211** (0.0580)	-0.2026*** (0.0783)	-0.2725** (0.1212)	-0.3967*** (0.1227)	-0.4089*** (0.1346)	-0.4391** (0.1776)
Opérateur/assembleurs	-0.0065 (0.0692)	-0.1133 (0.0825)	-0.1472 (0.1211)	-0.3638*** (0.1245)	-0.3047** (0.1412)	-0.1181 (0.1875)
Métiers élémentaires	-0.1084 (0.0832)	-0.2363** (0.1144)	-0.2275* (0.1210)	-0.1342 (0.1055)	-0.2711** (0.1152)	0.0513 (0.1504)
Militaires	0.1050** (0.0494)	0.1574* (0.0832)			-0.2556 (0.2872)	
_cons	0.5278*** (0.0854)	0.5298*** (0.0868)	0.5577*** (0.1244)	0.5834*** (0.1199)	0.5753*** (0.1168)	0.3833** (0.1658)
Obs.	506	468	260	468	447	274
R-squared	0.1468	0.1711	0.1795	0.1367	0.1412	0.1895

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

On peut également remarquer que l'effet de l'éducation sur l'emploi des personnes de 50-54 ans change avec la région observée. Pour les hommes, il existe toujours un effet fort et significatif de l'éducation en région flamande. En région wallonne aussi, bien que nos estimateurs sont moins précis à mesure qu'on se rapproche de l'éducation de base (primaire). A Bruxelles par contre, l'éducation ne semble pas avoir d'effet significatif bien que nos coefficients montrent qu'en moyenne être plus éduqué permet d'obtenir plus d'opportunités d'emploi. Pour les femmes, l'effet de l'éducation sur l'emploi n'est jamais statistiquement significatif sauf pour celles possédant une éducation supérieure et résidant en Flandre. Aussi, remarquons qu'à Bruxelles, l'effet de l'éducation sur l'emploi est plus important pour les

femmes que pour les hommes.

Graphique 4.6.4.a : Capacité à travailler estimée (taux) des hommes et des femmes par région



L'analyse de la régression précédente nous apprend que les coefficients estimés pour l'éducation partagent en réalité des interactions parfois importantes avec les régions. On décide donc de présenter ici les derniers résultats de ce travail ; les estimations des relations entre l'emploi et la santé en séparant les hommes et les femmes, les régions et les niveaux d'éducation sans contrôler pour les types d'emploi (pour éviter de surcharger ce travail et mes régressions et parce qu'ils ne semblent pas jouer de rôle fort dans la relation emploi/santé par niveau d'éducation comme on l'a vu précédemment dans le tableau 4.6.3.a, bien qu'il pourrait être intéressant néanmoins de les inclure tout de même).

Tableau 4.6.4.c : Régression de l'emploi des 50-54 ans sur la santé par genre et par niveau d'éducation pour la région flamande

	Hommes				Femmes			
	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0035*	0.0037***	0.0027**	0.0011	0.0012	0.0044***	-0.0002	-0.0004
	(0.0020)	(0.0013)	(0.0012)	(0.0009)	(0.0020)	(0.0013)	(0.0013)	(0.0012)
_cons	0.3769**	0.5788***	0.6695***	0.8710***	0.3879***	0.2382***	0.5956***	0.8759***
	(0.1459)	(0.1098)	(0.1004)	(0.0752)	(0.1337)	(0.0834)	(0.0935)	(0.0845)
Obs.	64	125	155	162	65	122	163	118
R-squared	0.0456	0.0900	0.0465	0.0173	0.0059	0.0825	0.0002	0.0012

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

Tableau 4.6.4.d : Régression de l'emploi des 50-54 ans sur la santé par genre et par niveau d'éducation pour la région Wallonne

	Hommes				Femmes			
	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0028 (0.0017)	0.0033** (0.0014)	0.0034*** (0.0013)	0.0016* (0.0009)	0.0009 (0.0017)	0.0051*** (0.0014)	0.0008 (0.0013)	0.0024* (0.0014)
_cons	0.3798*** (0.1081)	0.4824*** (0.0942)	0.5506*** (0.0936)	0.8404*** (0.0774)	0.3903*** (0.0969)	0.2223*** (0.0779)	0.5647*** (0.0796)	0.6195*** (0.0962)
Obs.	79	112	151	126	76	106	134	131
R-squared	0.0321	0.0524	0.0533	0.0504	0.0040	0.1055	0.0026	0.0253

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

Tableau 4.6.4.e : Régression de l'emploi des 50-54 ans sur la santé par genre et par niveau d'éducation pour la région Bruxelloise

	Hommes				Femmes			
	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure	Au plus primaire	Bas secondaire	Haut secondaire	Supérieure
PVWindex-IVa	0.0019 (0.0021)	0.0026 (0.0022)	0.0046*** (0.0016)	0.0035** (0.0014)	0.0038 (0.0028)	0.0024 (0.0018)	0.0029* (0.0015)	0.0032** (0.0015)
_cons	0.4692*** (0.1554)	0.4883*** (0.1608)	0.4783*** (0.1268)	0.6267*** (0.1121)	0.2893* (0.1622)	0.3624*** (0.1097)	0.4588*** (0.1061)	0.6446*** (0.1046)
Obs.	39	39	66	116	25	58	76	115
R-squared	0.0221	0.0361	0.1450	0.0936	0.0741	0.0294	0.0443	0.0632

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

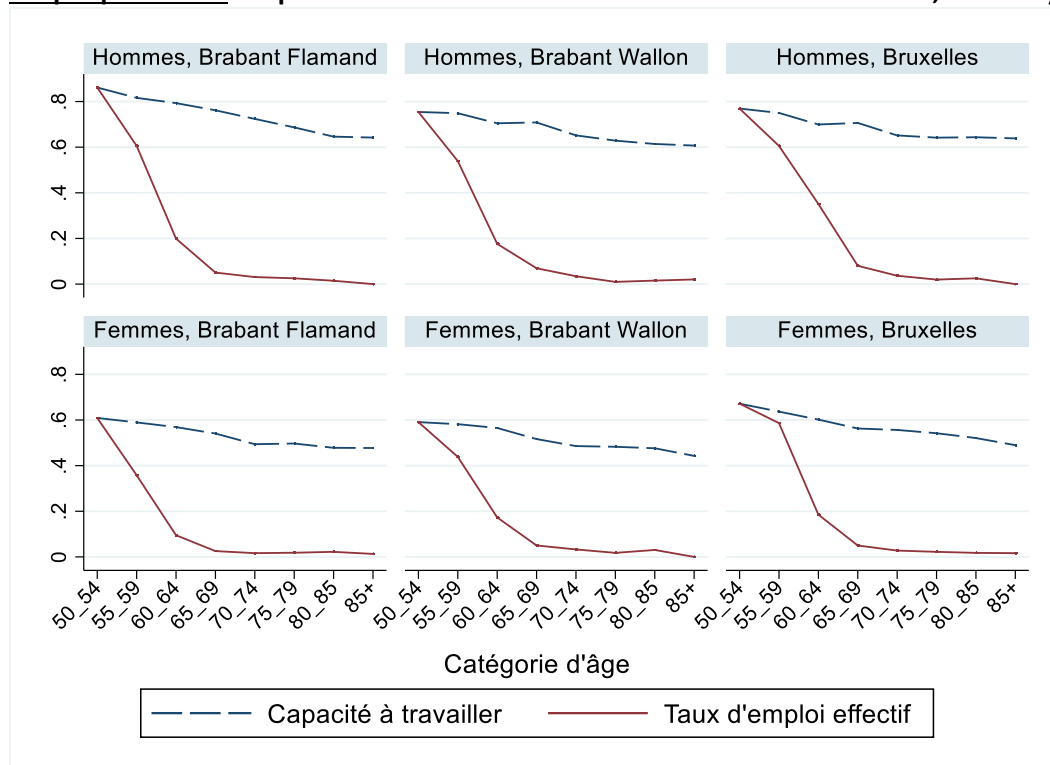
Ces différents tableaux de régressions nous en apprennent beaucoup sur les différentes relations emploi/santé par niveau d'éducation entre les régions. On voit que ces relations emploi/santé par niveau d'éducation et par région sont intimement liées à leur région respective. D'abord, l'effet relativement faible (par rapport aux autres niveaux d'éducation) de la santé pour l'enseignement supérieur découvert plus en amont dans ce travail serait typique de la région flamande ; dans cette région les femmes d'éducation supérieure ne semblent plus du tout dépendre de leur santé tout comme celles ayant terminé l'enseignement secondaire. Pour les hommes de la région flamande, ce résultat est similaire pour le dernier niveau d'éducation mais pas pour ceux ayant terminé l'enseignement secondaire. Aussi, les opportunités d'emploi pour les femmes ayant commencé sans terminer l'enseignement secondaire sont largement inférieures à celles des hommes comme le montre le coefficient associé à la constante (0.48 pour les hommes contre 0.22 pour les femmes)

Concernant la région Wallonne, ce sont les femmes du niveau d'éducation primaire et ayant terminé leur éducation secondaire qui dépendent le moins de leur santé et, dans une moindre mesure, les hommes et les femmes ayant une éducation supérieure. À contrario, on le voit, les femmes ayant commencé, sans terminer, l'enseignement secondaire sont celles qui semblent dépendre le plus de leur santé pour travailler. Aussi, les opportunités d'emploi pour les femmes possédant un diplôme d'enseignement supérieur de cette région sont largement inférieures à celles des hommes comme le montre le coefficient associé à la constante (0.84 pour les hommes et 0.62 pour les femmes).

Enfin, la région bruxelloise possède des relations emploi/santé assez forte pour chaque niveau d'éducation bien que tous nos coefficients de la santé ne soient pas significatifs. Cette région, comparée aux autres, voit également ses hommes ayant terminé l'enseignement secondaire ou supérieur posséder les plus fortes relations emploi/santé. Aussi, les opportunités d'emploi pour les femmes ne possédant pas de diplôme d'enseignement secondaire de cette région sont largement inférieures à celles des hommes ; 0.47 pour les hommes contre 0.29 pour les femmes de l'enseignement « au plus primaire » et 0.49 pour les hommes 0.36 pour les femmes ayant commencé sans terminer l'enseignement secondaire.

Enfin, on présente les capacités de travail moyennes par région obtenues avec l'aide de nos dernières régressions. Elles ne diffèrent pratiquement pas de celles obtenues lorsqu'on plaçait l'éducation et les catégories d'emploi en contrôle.

Graphique 4.6.4.b : Capacité à travailler estimée des hommes et des femmes, taux moyen par région



5. Conclusions

5.1. Résumer des résultats

Il existe un fort potentiel de travail inexploité au sein de la population âgée de Belgique. Tous nos résultats vont dans ce sens et concordent avec ceux des différents auteurs traitant sur le sujet. La capacité à travailler, qui est le travail potentiel pour une certaine santé, ne décline que très légèrement au cours de la vie des individus reflétant, non pas que la santé n'a pas d'effet sur l'emploi, mais que, dans nos sociétés modernes, le déclin de la santé est très lent avec l'âge. Par contre, nos résultats démontrent un second fait bien plus important ; tout le monde n'est pas dans une situation égale sur le marché du travail. A partir du moment où l'on

s'intéresse à certaines différences démographiques on voit apparaître des différences clefs en termes d'opportunités d'emploi et dans les relations emploi/santé estimées :

- Les femmes subissent encore une situation inégale concernant les opportunités d'emploi.
- L'éducation est un facteur essentiel puisqu'il offre aux individus des opportunités d'emploi bien plus importantes.
- L'éducation supérieure permet également de réduire largement ces inégalités d'opportunités d'emploi entre les hommes et les femmes, voire, de les anéantir totalement dans certains cas.
- En contrôlant pour les différents types d'emploi, on ne voit pas la relation emploi/santé par niveau d'éducation se modifier. Notre analyse par catégorie d'emploi nous a conduit à penser que les catégories d'emploi reflèteraient des différences d'éducation, mais le contrôle de l'éducation lors de la régression par catégorie d'emploi nous fait nuancer cette hypothèse ; tous les métiers ne sont pas égaux face aux effets de la santé sur l'emploi.
- Aux différences de relation emploi/santé par type d'emploi et par genre viennent s'ajouter celles provenant des différents niveaux d'éducation. Comme nous l'avons vu les femmes, ayant commencé mais n'ayant pas terminé l'enseignement secondaire, font face à une relation emploi/santé plus importante que les autres catégories. Ce qui semble indiquer que celles-ci dépendent plus de leur santé que les autres bien que ce résultat se nuance également selon les régions (tableau 4.6.4. (c, d, e)).
- En contrôlant pour les régions on voit certains coefficients associés à la relation emploi/santé par niveau d'éducation changer et/ou perdre leur significativité, d'autres en gagner ; ce résultat nous apprend que les interactions entre l'éducation et les régions sont plus complexes qu'il n'y paraît et méritent un plus large approfondissement.

5.2. Résumé des limites de ce travail

Commençons par le commencement, la base de données provient d'une méthode d'échantillonnage aléatoire par niveau pour laquelle des randomisations sont faites à chaque niveau. Il n'est normalement pas approprié de réaliser de l'inférence statistique telle qu'utilisée ici ; on aurait dû utiliser une commande spéciale pour signaler à notre programme la méthode d'échantillonnage. Néanmoins, comme expliqué dans notre partie sur les données, nos erreurs-types et nos coefficients ne devraient pas être trop impactés par la

méthode d'échantillonnage et sont censés nous fournir des estimateurs très proches de ceux qu'on aurait eus en randomisant directement chaque individu de 50 ans et plus de Belgique.

Aussi, plus inquiétant, nos données originelles présentaient un nombre important d'individus avec des données manquantes ; en les nettoyant, on perd près de la moitié de ceux-ci. Si les individus supprimés de notre base de données sont différents de ceux utilisés dans notre analyse, alors les résultats obtenus peuvent être totalement remis en question. Là encore, la question mérite d'être approfondie.

Encore pour la base de données ; des travaux plus poussés et réalisés avec l'aide de données de panels permettent de quantifier les effets fixes au temps. Selon Blundell et al. (2017), nos estimateurs devraient surestimer l'effet de la santé par deux fois environ. Le travail de mon promoteur (Vandenberghe, 2019) semble le confirmer. Dans son travail mon promoteur étudie également une question restée en suspens dans la littérature sur la capacité à travailler ; est-ce que le travail définit uniquement par sa forme extensive (nombre de personnes au travail) est suffisant ? Et donc ; est-ce qu'étudier également la forme intensive (nombre d'heures de travail par travailleur) apporte un éclairage nouveau à la question ? Il trouve un effet significatif de la santé sur le nombre d'heures travaillées bien que celui-ci soit faible par rapport à la forme extensive. Il en conclut que les individus préfèrent arrêter totalement de travailler plutôt que d'ajuster leur temps de travail à leur santé. Ce fait semble confirmer l'étude plus historique réalisée par Costa (1998).

Comme expliqué plus haut, la régression par Moindres Carrés Ordinaires n'est pas la méthode d'estimation la plus appropriée pour étudier une variable dépendante binaire. Si elle est privilégiée dans ce travail c'est pour sa plus grande facilitée d'interprétation des résultats par rapport à un modèle logit ou probit et parce que ses résultats, tant sur le plan des estimateurs que de leurs variances, sont très similaires à ceux de ces modèles pour variable dépendante binaire. Aussi, parce qu'une très grande partie de la littérature empirique utilisée ici privilégie également la régression par MCO. Ce qui nous inquiète ici c'est que la régression par MCO se base sur des hypothèses qui ne sont, pour la plupart, pas vérifiées. Par exemple, on montre que la santé ne semble pas avoir un effet linéaire sur l'emploi mais posséderait un effet plutôt concave (forte pour une santé faible et faible pour une santé forte). On peut également citer l'hypothèse de normalité des résidus qui n'est pas respectée non plus... Sur base de ce non-respect de ses hypothèses, on ne peut que conclure que nos estimateurs sont biaisés de même que leurs variances.

La méthode en composante principale telle que je l'utilise ici me laisse personnellement perplexe ; le test de Kaiser-Meyer-Olkin de l'analyse en composante principale pour mesurer en combien de variables on devrait synthétiser l'ensemble de nos indicateurs de santé (qui n'est pas présenté dans ce travail) me confirme que synthétiser la santé en un seul indicateur ne permet pas d'en inclure l'ensemble des variations ; qu'utiliser plusieurs composantes principales serait plus approprié. Des travaux plus poussés synthétisent la santé en plusieurs indicateurs ; séparant la santé objective, la santé subjective et la cognition (Vandenberghe 2019). Ne pas découvrir d'effet de la santé sur l'emploi pour certaines catégories démographiques particulières renforce ces craintes ; Jousten A. (2017) nous informe que parmi plusieurs explications possibles, en voilà deux : (1) ces individus possèdent peut-être

moins de flexibilité dans la possibilité d'un arrêt de travail, et donc, face à la santé et (2) que notre indice de santé ne contienne pas les variations nécessaires pour déterminer l'emploi (notre indice de santé n'est pas suffisant, pas assez précis).

Enfin, j'émettrais une dernière limite car, en y réfléchissant bien, je pourrais sans doute en énumérer beaucoup trop ; je ne suis qu'un étudiant, mon travail est loin d'être parfait et, arrivant sur sa fin, je sais qu'il mérite une révision complète et fastidieuse que je ne peux me permettre actuellement, « Dead-line oblige ». Un lecteur avisé prendra chacun de mes résultats spécifiques « avec des pincettes », ce qui ne change pas les conclusions finales de celui-ci ; la santé seule ne permet pas de justifier un arrêt de travail à 65 ans. Bien que la santé puisse justifier l'incapacité à travail d'une personne en très mauvaise santé (le graphique 4.5.d montrait que l'élasticité de l'emploi à la santé dépend aussi du niveau de santé), l'évolution générale de la santé dans la population décline très lentement avec l'âge ; Il existe donc un fort potentiel de travail inexploité au sein de la population âgée de Belgique.

5.3. Ce que ce travail m'a apporté

Une question de recherche étrange selon moi, sensible selon mon promoteur ; chercher à augmenter l'âge de la retraite serait « mal vu » et pourtant, désormais, j'en suis convaincu, c'est l'option la plus sûre. Avant, j'aurais pu être contre l'idée en vociférant qu' « augmenter l'âge de la retraite c'est violer nos droits sociaux et la mémoire des générations passées qui se sont battues pour les obtenir ». En ayant étudié la question, j'ai désormais le sentiment que ne pas augmenter cet âge d'éligibilité c'est postposer un problème fondamental de la sécurité sociale ; sa soutenabilité. Les générations retraitées d'aujourd'hui vivent trop richement grâce au travail des générations présentes et postposer le problème c'est le léguer aux générations futures, ce que je ne désire pas. Ce travail m'a donc permis de réfléchir et de penser autrement.

Un travail empirique, une base de données, une méthode à appliquer et d'innombrables heures passées derrière un écran pour essayer d'obtenir des résultats dont je suis loin de comprendre les subtilités ; c'est exactement ce dont je rêvais pour mon mémoire. Encore merci à mon promoteur pour ce sujet passionnant. Partir de rien pour arriver à un résultat qui ressemble à quelque chose. Ceci m'a permis d'approfondir mes connaissances et surtout de les enregistrer sur un disque dur et pas dans mon cerveau. Appliquer, au lieu d'étudier, m'est bien plus enrichissant.

Ne pas avoir peur de l'inconnu ; j'ai perdu un temps considérable pour me décider à faire appel à un professeur pour m'aider dans le choix de mon sujet de mémoire. Si Mr. Vandenberghe n'avait pas répondu à mon appel ; j'ai bien peur que je n'aurais tout simplement pas pu rédiger mon mémoire à temps. Je n'ai pas pris l'option sur l'économie du travail. Je n'avais jamais entendu parler de la capacité à travailler avant de démarrer ce mémoire. Et j'utilise et parle ici de tout un tas de chose dont je ne connaissais rien et qui m'échappe encore de loin. Je suis

dans une situation similaire face au monde du travail ; une méconnaissance absolue. Pourtant aujourd'hui, ce qui n'était pas le cas il y a six mois, je ne le crains plus et l'attends avec impatience ; si ne pas être sûr de moi devait me justifier de ne pas essayer alors je ne ferais jamais rien.

6. Bibliographie

- Blanks J., Emmerson C., Tetlow G., (2017) «Healthy, Happy, and Idle: Estimating the Health Capacity to Work at Older Ages in Germany», Dans *Social Security Programs and Retirement around the World: the Capacity to Work at Older Ages*, de David A. Wise, 149 - 180, Chicago, University of Chicago Press.
- Blundell R., Britton J., Costa Dias M., French E., (2017), *The impact of health on labour supply near retirement*, disponible sur <https://ssrn.com/abstract=310388a>, consulté le 5 Avril 2019.
- Catherine S., (2015), *Introduction a l'économétrie*, disponible sur <https://sylvaincatherine.files.wordpress.com/2015/01/instruments2.pdf>, consulté le 6 Avril 2019.
- Costa Dora L. (1998), *The evolution of retirement : an American economic history, 1880-1990*, (s.l.), University of Chicago Press.
- Cutler D., Meara E., Richards-Shubik S., (2012), «Health and Work Capacity of Older Adults: Estimates and Implications for Social Security Policy», *Social Science Research Network*, disponible sur <http://ssrn.com/abstract=2577858>.
- Wise David A., (2017), *Social Security Programs and Retirement around the World : The Capacity to Work at Older Ages*, Chicago, University of Chicago Press.
- Eurostat, (s.d.), *Votre clé d'accès à la Statistique européenne*, base de données, disponible sur <https://ec.europa.eu/eurostat/web/lfs/data/database>
- Gruber, Jonathan, Wise David A., (1999), *Social Security Programs and Retirement around the World*, Chicago, University of Chicaco Press.
- Hindriks Jean, (2018), «Ces chiffres qui prouvent l'ampleur du défi lié au financement des pensions», *L'Echo*, disponible sur <https://www.itinerainstitute.org/fr/article/ces-chiffres-qui-prouvent-lampleur-du-defi-lie-au-financement-des-pensions/>
- Johansson P., Laun L., Palme M., (2017), «Health, Work Capacity, and Retirement in Sweden», Dans *Social Security Programs and Retirement Around the World: the Capacity to Work at Older Ages*, de David A. Wise, 301 – 327, Chicago, University of Chicago Press.
- Jousten A., Lefebvre M., (2017), «Work Capacity and Longer Working Lives in Belgium», Dans *Social Security Programs and Retirement around the World: the Capacity to Work at Older Ages*, de David A. Wise, 35 – 58, Chicago, University of Chicago Press.

- Jürges H., Thiel L., Bôrsch-Supan A., (2017), «Healthy, Happy, and Idle: Estimating the Health Capacity to Work at Older Ages in Germany», Dans *Social Security Programs and Retirement around the World: The Capacity to Work at Older Ages*, de David A. Wise, 329 – 357, Chicago, University of Chicago Press.
- Milligan K., Wise David A., (2015), «Health and Work at Older Ages: Using Mortality to Assess the Capacity to Work across Countries», *Journal of Population Aging*, 8 (1-2) : 27-50.
- National Bureau of Economic Research, (s.d.), *NBER Aging Program*, disponible sur <https://www.nber.org/programs/ag/>, consulté le 30 Juillet 2019.
- Poterba J., Venti S., Wise D.A., (2013), "Health, Education, and the Post-Retirement Evolution of Household Assets", *NBER Working Paper*, no.18695, Cambridge, MA.
- Scientific Institute of Public Health, (s.d.), *Manual for the users of the Health Interview Survey 2013*.
- Service fédéral des Pensions, (2019), *Comment calculer une date de pension?*, disponible sur <https://www.sfpd.fgov.be/fr/age-de-la-pension/quand#legal> consulté le 1 février, 2019.
- Stefaan Demarest, Johan Van der Heyden, Rana Charafeddine, Sabine Drieskens, Lydia Gisle, and Jean Tafforeau, (2013), "Methodological basics and evolution of the Belgian health interview survey 1997–2008", *Archives of Public Health*, 71(1) : 24.
- University of California, Berkley (USA), and Max Planck Institute for Demographic Ressearch (Belgium), (2019), «Human Mortality Database», base de données, disponible sur <https://www.mortality.org/cgi-bin/hmd/country.php?cntr=BEL&level=1>
- Vandenbergh, V., (2019), "Health, Cognition and Work Capacity Beyond the Age of 50", *GLO Discussion Paper*, No. 295, disponible sur <https://www.econstor.eu/handle/10419/190974>.

7. Annexes (code Stata – version secrétariat)

