





**HEC MONTRÉAL**

**Rendements de l'éducation au Chili : une application du modèle  
endogène de scolarisation optimale**

**par**

**Andrés Díaz Hernández**

**Sciences de la gestion  
(Économie appliquée)**

*Mémoire présenté en vue de l'obtention  
du grade de maîtrise ès sciences en gestion  
(M. Sc.)*

Décembre 2019  
© Andrés Díaz Hernández, 2019



## Résumé

Cette recherche porte sur l'estimation des rendements de l'éducation pour les travailleurs et travailleuses chiliens. En utilisant les données de l'Enquête supplémentaire de revenus (ESR) pour le Chili, nous identifions des écarts en termes d'années de scolarité et des salaires entre les individus ayant une université à proximité et ceux qui ne l'ont pas. Contrairement aux résultats trouvés pour les pays développés, les bénéfices en termes d'années de scolarité et salaires se concentrent sur une grande partie de la population chilienne et pas seulement pour les individus appartenant à un milieu familial défavorisé. Ces résultats confirment que les taux de rendement marginaux sont supérieurs aux rendements estimés par les méthodes conventionnelles.

**Mots clés :** modèle endogène de scolarisation optimale, variables instrumentales, rendements de l'éducation au Chili, économétrie, enquête, recherche quantitative



## Table de matières

Résumé.....	iii
Liste des tableaux et des figures .....	vii
Remerciements.....	ix
Introduction.....	1
1. Revue de la littérature.....	3
1.1 Marché du travail et éducation au Chili .....	8
2. Construction des données et variables disponibles .....	11
3. Méthodologie.....	17
3.1 Modèle endogène de scolarisation optimale.....	17
3.2 Modèles de régression .....	21
4. Résultats .....	23
4.1 Caractéristiques des travailleurs et travailleuses chiliens.....	23
4.2 Résultats des modèles d'estimation.....	24
4.3 Validité et légitimité de l'instrument.....	29
4.4 Discussion des résultats et analyse du contexte institutionnel chilien .....	31
4.5 Interaction de proximité et milieu familial défavorisé comme instrument pour l'éducation complétée.....	32
4.6 Différences de rendement de l'éducation entre hommes et femmes.....	33
Conclusion .....	39
Bibliographie.....	41



# Liste des tableaux et des figures

## Liste des tableaux

Tableau 1.1. Sommaire d'études qui estiment le rendement de l'éducation .....	5
Tableau 2.1. Correspondance entre le niveau éducatif et le maximum d'années accordées à chaque niveau .....	12
Tableau 2.2. Variables disponibles .....	13
Tableau 4.1. Caractéristiques de l'échantillon – Enquête supplémentaire de revenus pour 2017 .....	23
Tableau 4.2. Variables explicatives des modèles de régression.....	24
Tableau 4.3. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation .....	25
Tableau 4.4. Variables indicatrices selon les sous-groupes de revenus familiaux.....	26
Tableau 4.5. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation - hommes.....	35
Tableau 4.6. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation - femmes .....	36

## Liste des figures

Figure 3.1. Taux d'escompte et bénéfices marginaux .....	18
Figure 3.2. Niveau optimal de scolarité .....	20



## **Remerciements**

Un gros merci à mon directeur du mémoire, Daniel Parent. Je serai toujours reconnaissant de votre aide et conseils.

Merci à Tatyana Koreshkova, Frederick Davis, Annette Bowen et Linda Campione, qui m'ont encouragé à continuer mes études supérieures.

Merci à l'Institut national de statistiques au Chili pour clarifier toutes mes questions méthodologiques concernant la base de données utilisée dans cette recherche.

Finalement, un gros merci à Jean-François pour son accompagnement et son soutien pendant toute la période de mes études.



## Introduction

Les rendements sur l'investissement en éducation ont été estimés depuis les années 50. Dans le cadre théorique du capital humain, ce rendement est défini comme l'arbitrage entre coûts et bénéfices encourus dans le temps pour atteindre un certain niveau de scolarité. Récemment, le débat se centre sur le lien de causalité entre éducation et salaires en se basant sur l'équation du capital humain de Jacob Mincer (1974). Ce lien de causalité est particulièrement important pour le Chili, un pays qui a subi dans les dernières décennies des changements importants en termes d'accès à l'éducation. Effectivement, l'augmentation de la réussite scolaire aux niveaux secondaire et post-secondaire amène une préoccupation sur la structure des coûts relatifs pour ceux qui ont potentiellement des contraintes d'accès (Card, 2001). En outre, les dernières études remarquent l'importance de tenir compte que l'éducation est aussi une variable endogène et par conséquent, un modèle économétrique convaincant doit inclure des variations exogènes dans les décisions de scolarité.

Cette recherche vise à déterminer les rendements de l'éducation des travailleurs et travailleuses chiliens. Pour ce faire, nous utilisons la proximité d'une université comme une source de variation exogène quant aux décisions de scolarité pour estimer le rendement de l'éducation par variables instrumentales (VI). Ensuite, nous utilisons le cadre conceptuel du modèle endogène de scolarisation optimale de Gary Becker (1967) pour interpréter les résultats de nos estimations. Les données de notre recherche proviennent d'un échantillon de travailleurs et travailleuses chiliens de l'Enquête supplémentaire de revenus (ESR) sondés en 2017.

Notre recherche est organisée comme suit : le chapitre 1 développe un survol sur les études scientifiques qui estiment le rendement de l'éducation ainsi qu'une vue globale sur le contexte du Chili; le chapitre 2 présente le traitement effectué aux bases de données; le

chapitre 3 présente le modèle endogène de scolarité qui nous sert comme cadre théorique pour l'analyse de nos données. Finalement, le chapitre 4 présente les résultats des différentes spécifications du modèle économétrique.

# 1. Revue de la littérature

Il existe un consensus parmi les académiciens que l'éducation de l'individu est un déterminant important de plusieurs variables (Blau, 1996), tel que les salaires. Plusieurs études citées dans la littérature scientifique ont mis en évidence cette relation, en estimant des modèles économétriques basés sur le modèle de capital humain de Jacob Mincer<sup>1</sup>. Toutes ces études montrent que les individus ayant un niveau de scolarité élevé bénéficient des salaires élevés. Malgré cette conclusion, les économistes ont souvent remis en question une interprétation causale de cette relation statistique, étant donné que les individus choisissent leurs niveaux d'éducation. Au cours des dernières années, certains auteurs essaient de répondre à cet enjeu, en mettant l'accent sur les façons de tenir compte de l'endogénéité inhérente aux décisions de scolarité et tenir compte des variables non observées. Si l'on ne considère cette variable comme endogène, il peut exister des erreurs d'inférence et de spécification. Tel que remarqué par David Card (2001), ces études utilisent des sources de variation exogènes quant aux décisions de scolarité. Plusieurs méthodologies sont identifiées dans la littérature économique pour corriger le problème d'endogénéité et ainsi estimer le taux de rendement de l'éducation, notamment la méthode des variables instrumentales (des facteurs observables qui sont indépendants des aptitudes individuelles) et les estimateurs à effets fixes. Les études les plus importantes visant à expliquer l'effet de l'éducation sur les salaires sont résumées dans le tableau 1.1.

Dans leur étude, Angrist et Krueger (1991) utilisent le trimestre de naissance comme instrument pour la scolarité complétée dans leur modèle de scolarité et salaires pour des hommes américains. Selon leurs résultats, les hommes qui sont nés dans les premiers mois de l'année ont des salaires moins élevés par rapport aux hommes qui sont nés dans les derniers mois de l'année. Cela est la conséquence des lois qui stipulent un âge obligatoire de fréquentation scolaire. Les individus qui sont nés à la fin de l'année, quant à eux, sont forcés à rester plus longtemps à l'école en raison de ces lois.

---

<sup>1</sup> Voir l'étude de Jacob Mincer (1967) pour en approfondir sur le modèle du capital humain.

Plusieurs auteurs mettent l'accent sur les mesures de qualité et leurs effets sur la performance scolaire et le marché d'emploi. La contribution la plus connue est celle de David Card et Alan B. Krueger (1992). Ces auteurs arrivent à la conclusion générale qu'il existe une relation positive entre les ressources et les résultats scolaires. Ils s'intéressent notamment aux salaires futurs et à la réussite scolaire. Pour examiner cette relation, ils spécifient un cadre théorique basé sur quatre propositions : a) il existe une relation positive entre les années de scolarité et les salaires, b) les bénéfices marginaux d'une année de scolarité est supérieur pour les individus qui profitent des écoles de meilleure qualité, c) plus la qualité scolaire augmente, plus certains élèves assisteront à l'école plus longtemps et d) ceux qui optent pour un niveau d'études supérieur auront tendance à avoir une plus grande fonction de salaire, indépendamment de leur niveau d'éducation.

Ensuite, ils construisent un modèle de régression de forme réduite pour estimer les effets des ressources scolaires sur le profil salaire-éducation des individus. En analysant une base de données de salaires pour trois cohortes de travailleurs recensés aux États-Unis en 1980, les deux auteurs estiment le taux de rendement de la scolarité par état de résidence et l'année de naissance. Cette procédure génère un taux de rendement de l'éducation d'une année additionnelle de scolarité pour chaque cohorte définissant un intervalle d'années de naissance (mesure de réussite scolaire). Dans la deuxième étape, les auteurs estiment, pour chaque état américain, la relation entre le taux de rendement de l'éducation et différentes mesures pour la qualité des écoles telles que les salaires moyens de professeurs, la taille de la classe et la durée de la session scolaire, et ce, tout en tenant compte des différentes variables démographiques. En utilisant des données historiques de la *Biennial Survey of Education and Digest of Educational Statistics*, les auteurs ont calculé les valeurs moyennes de ces variables pour les années où les individus poursuivaient des études préuniversitaires.

**Tableau 1.1. Sommaire d'études qui estiment le rendement de l'éducation**

Étude	Échantillon	Méthodologie	Commentaires
Angrist et Krueger (1991)	Recensement aux É.-U. en 1970 et 1980. Cohortes des hommes nés avant 1950	Estimation des moindres carrés ordinaires à deux étapes. Instruments utilisés : année * trimestre de naissance, état * trimestre de naissance, en incluant des variables de contrôle	L'estimateur par variables instrumentales est 28% plus élevé que ceux obtenu par moindres carrés ordinaires.
Card et Krueger (1992)	Recensement aux É.-U. en 1980. Hommes blancs âgés 31-60 ans.	Données agrégées par état pour la période 1926-1966 appariées aux individus, par état de naissance et années de scolarité. Variable dépendante : logarithme des salaires en 1979. Mesure de qualité de l'éducation : ratio élève-professeur. Autres variables explicatives : état de naissance, état du domicile, expérience (au carré), état civil.	La qualité de l'éducation a des effets significatifs sur la réussite scolaire.
Card (1993)	<i>National Longitudinal Survey of Youth</i> pour 1966 et 1976.	Estimation des moindres carrés ordinaires à deux étapes. Instruments : proximité d'une université et interaction de proximité d'une université et milieu familial	Les rendements scolaires sont plus élevés pour les individus qui proviennent d'un milieu familial défavorisé.
Ashenfelter et Krueger (1994)	<i>Enquête Twinburg pour 1991</i>	Estimation d'un modèle de moindres carrés généralisés (estimateur à effets fixes) pour comparer les salaires des jumeaux identiques ayant des niveaux de scolarité différents. Instrument : années de scolarité de l'individu obtenues de son jumeau.	Les estimés du rendement scolaire obtenus à l'aide de la méthode de variables instrumentales sont plus élevés que ceux obtenus par la méthode de moindres carrés ordinaires.

**Tableau 1.1 (Suite)**

Lemieux et Card (1998)	Recensement canadien de 1971. Hommes Ontariens et Québécois âgés 25-65 en 1971.	Estimation des moindres carrés ordinaires à deux étapes. Instrument : variable dichotomique pour les hommes de l'Ontario âgés de 19 à 22 ans en 1946.	Les estimés du taux de rendement scolaire obtenus à l'aide de la méthode de variables instrumentales sont plus élevés que ceux obtenus par la méthode de moindres carrés ordinaires.
Urzúa (2004)	<i>National Longitudinal Survey of Youth</i> pour 1979.	Estimation d'un modèle de forme réduite pour estimer des écarts salariaux pour des hommes noirs et blancs, en incluant des variables pour mesurer les habiletés non cognitives et observées et non observées.	Les habiletés non observées sont aussi importantes que l'environnement familial pour expliquer les différences des décisions d'éducation et des salaires observés.

Le modèle de Card et Kruger arrive à des conclusions significatives quant à la relation entre la qualité de l'éducation et les augmentations salariales attribuées à la réussite scolaire. Les augmentations salariales des professeurs et la diminution du ratio professeur-élève sont positivement reliées aux gains du rendement de l'éducation.

Card (1993) utilise la proximité d'une université comme un déterminant exogène du niveau de scolarité. L'auteur souligne que les individus ayant une université à proximité de leur domicile ont des salaires plus élevés. De plus, l'auteur démontre que les individus appartenant aux ménages défavorisés ont des taux marginaux de rendement de l'éducation plus élevés, ce qui rationalise le fait que les estimés du taux de rendement scolaire obtenus à l'aide de la méthode de variables instrumentales (VI) soient plus élevés que ceux obtenus par la méthode de MCO. Notons que cette méthode est utilisée pour cette recherche<sup>2</sup>.

Ashenfelter et Krueger (1994) utilisent la méthode de moindres carrés généralisés (MCG) pour comparer les salaires des jumeaux identiques qui ont des différents niveaux de scolarité. Après avoir tenu compte des erreurs de mesure et les variables omises dans leur modèle, les auteurs soutiennent qu'une année de scolarité additionnelle fait augmenter les salaires d'un individu de 12-16%. Cela confirme les biais négatifs trouvés dans les coefficients obtenus par les estimateurs de MCO.

Lemieux et Card (1998), estiment les effets des bénéfices offerts aux vétérans sur des hommes francophones au Québec et sur des hommes anglophones en Ontario. La variable instrumentale dans ce cas-ci est une variable dichotomique pour les hommes de l'Ontario âgés de 19 à 22 ans en 1946. Cela permet de mettre en valeur les effets de cohorte sur la réussite scolaire et les salaires observés, en se reposant sur les différences entre Ontariens et Québécois dans une cohorte par rapport aux autres pour identifier le taux de rendement de l'éducation. La variation exogène exploitée dans cette étude est le fait que, d'une part, peu de Québécois francophones ont servi durant le Seconde Guerre Mondiale et, d'autre

---

<sup>2</sup> Nous présenterons dans le chapitre 4 une description plus détaillée du modèle endogène de scolarisation optimale.

part, même ceux qui ont servi n'étaient pas vraiment touchés par ce programme de subvention des études universitaires en raison de leur faible niveau de scolarité complétée.

Un autre groupe d'études constatent le rôle des habiletés cognitives et non cognitives sur les déterminants des salaires futurs. Prenons par exemple l'étude de Sergio Urzúa (2004). En traitant les décisions scolaires comme étant endogènes, l'auteur estime les différences de salaires entre hommes blancs et noirs. En utilisant le *National Longitudinal Survey of Youth* pour 1979, ses résultats indiquent que les habiletés non observées sont aussi importantes que l'environnement familial pour expliquer les différences des décisions d'éducation et des salaires observés.

Finalement, on observe un vif intérêt pour expliquer les déterminants de la croissance économique et le rôle du capital humain sur le développement économique (Topel, 1999). Ces études analysent la relation dans un cadre de croissance économique; quelques-unes parmi ces études incluent des pays en voie de développement dans leurs analyses (Hanushek et Kimko, 2000; Hanushek et Woessmann, 2008). La conclusion générale qui en découle est une forte association entre la performance économique des nations et leur performance sur des tests standardisés internationaux. Par exemple, Eric Hanushek et Ludger Woessmann (2008) utilisent des données agrégées pour 95 pays et concluent que les habiletés cognitives ont des effets significatifs sur les salaires individuels, la distribution de la richesse et la croissance économique.

## **1.1 Marché du travail et éducation au Chili**

Le Chili a démontré un progrès notable en termes d'accès à l'éducation primaire et secondaire. À la suite d'une importante réforme éducative mise en place par le gouvernement chilien en 2003, l'éducation préuniversitaire est devenue obligatoire. Selon les données mondiales de l'éducation (UNESCO, 2010), un élève qui commence l'école primaire à 6 ans doit compléter son éducation secondaire à 18 ans, pour atteindre douze

années de scolarité. Grâce à cette réforme, le taux d'inscription au niveau secondaire est passé de 79 % en 1990 à 99 % en 2015 (MINEDUC, 2017).

De plus, Patrinos (1995) constate que le taux moyen du rendement de l'éducation au Chili se classe parmi les plus élevés par rapport au reste de l'Amérique latine, et même aux pays de l'Organisation de coopération et de développement économique (OCDE). Malgré ceci, il existe plusieurs enjeux quant à la qualité de l'éducation et aux compétences nécessaires pour l'intégration des individus sur le marché du travail. Selon les résultats du test standardisé PISA<sup>3</sup> obtenus pour 2018, les élèves chiliens démontrent des déficiences au niveau de la capacité de lecture et de l'analyse mathématique (OCDE, 2019c). En moyenne, près de 32 % des élèves chiliens se classent au niveau de base en lecture, tandis que la moyenne pour les pays de l'OCDE est de 23 %. Les résultats en mathématiques sont plus alarmants, avec 62 % des élèves chiliens placés au niveau le plus bas versus 24 % pour ceux de l'OCDE. Selon Carrillo et Ramos (2016), ce problème de qualité a un impact sur la capacité productive du pays et sur les perspectives de croissance économique. Également, l'importance de la qualité du capital humain délimite les salaires sur le marché du travail.

Une analyse dynamique de l'offre et la demande du travail au Chili (Carrillo, Espinoza et Valenzuela, 2018) précise une tendance vers une forte demande quant au nombre de travailleurs qualifiés, ce qui indique une valorisation d'un niveau avancé de scolarité et des compétences. Le vieillissement de la population active a lieu en même temps que les plus jeunes travailleurs deviennent plus éduqués.

À mesure que les bases de données deviennent disponibles, plus il est possible d'appliquer la méthodologie utilisée pour les pays développés aux pays en voie de développement. Les bases de données utilisées dans ce mémoire permettront de comparer les divers résultats à ceux obtenus par les études mentionnées précédemment.

---

<sup>3</sup> *Programme for International Student Assessment.*



## 2. Construction des données et variables disponibles

La base de données utilisée pour notre analyse est l'*Encuesta Suplementaria de Ingresos* (traduite en français comme l'Enquête supplémentaire des revenus) pour l'année 2017. Elle est composée d'un module de l'Enquête nationale d'emploi, qui vise à caractériser les revenus individuels et des ménages de la population active. Les données sont collectées sur une base annuelle à travers les quinze<sup>4</sup> régions du territoire chilien. Afin de rendre l'échantillon représentatif de la population, chaque observation est pondérée par le poids échantillonnal du trimestre de la collecte des données.

Nous décrivons les modifications apportées à l'Enquête supplémentaire des revenus (ESR), ainsi que les traitements réalisés afin de construire l'échantillon final, qui servira à analyser notre modèle.

### **Statut d'emploi**

Notre analyse se limite aux individus disposant d'une expérience de travail potentielle positive. Les travailleurs autonomes, excluant le personnel du service domestique, sont inclus dans notre échantillon. De plus, nous avons conservé les travailleurs à temps partiel. Puisque l'enquête contient des personnes qui étaient à la recherche d'emploi au moment de la collecte des données, nous avons conservé seulement les personnes qui détenaient un emploi depuis au moins un mois. Nous avons créé une variable indicatrice afin d'indiquer les conditions précédentes.

### **Expérience**

Nous avons généré la variable « expérience » à partir de l'âge de la personne, selon l'équation suivante<sup>5</sup> :

---

<sup>4</sup> À partir de 2019, une région additionnelle a été créée, en augmentant le nombre de régions à seize.

<sup>5</sup> Le chiffre 6 représente l'âge à laquelle la première année d'éducation est entamée.

$$\text{expérience} = \text{âge} - \text{éducation} - 6 \quad (1)$$

Nous utilisons l'expérience potentielle lorsque l'expérience actuelle des individus n'est pas disponible dans les données utilisées. Selon Bédard (2014), cette variable proxy peut cependant surestimer l'impact de l'expérience de travail sur le salaire. En effet, l'auteur constate que cette construction assume qu'il n'y a aucune interruption entre la dernière année de scolarité et l'entrée sur le marché du travail ainsi que durant la carrière de l'individu.

### Éducation

La base de données ne contient pas les années de scolarité pour chaque personne. La base rapporte uniquement l'éducation de l'individu dans deux variables : « niveau », pour chaque niveau éducatif (de la maternelle au doctorat) et « cours », qui est associé à chaque niveau. La dernière variable prend les valeurs de 1 à 8, en fonction de chaque niveau. La procédure suivante vise donc à trouver le niveau maximum d'éducation et le dernier cours complété associé à chaque niveau. Le tableau 2.1 spécifie le maximum d'années accordées à chaque combinaison des niveaux et des cours.

**Tableau 2.1. Correspondance entre le niveau éducatif et le maximum d'années accordées à chaque niveau**

Niveau éducatif	Maximum d'années accordées
Maternelle	1
Primaire	8
Secondaire	4
Formation technique	3

**Tableau 2.1. (suite)**

Institut professionnel	4
Université	7
Études supérieures spécialisées	2
Maîtrise	2
Doctorat	5

Pour accorder les années de scolarité, une base de données contenant la correspondance entre chaque combinaison de niveau et de cours a été converti en années de scolarité. Après, chaque observation de l'ESR se voit assigner une valeur qui servira pour l'analyse de la régression.

Le tableau 2.2 ci-dessous présente une liste des variables disponibles pour l'année 2017.

**Tableau 2.2. Variables disponibles**

<b>Variable</b>	<b>Description</b>	<b>Type de variable (valeur minimale- valeur maximale)</b>
<i>Région</i>	Code de la région de domicile	Discrète (1-15)
<i>Ville</i>	Code de la ville de domicile	Discrète (1-346)
<i>Âge</i>	Âge de l'individu	Discrète (18-77)
<i>Niveau</i>	Dernier niveau d'études atteint	Discrète (0-9)

**Tableau 2.2. (suite)**

<i>Cours</i>	Dernier cours approuvé correspondant au niveau	Discrète (0-14)
<i>Marié</i>	Statu matrimonial de l'individu	Dichotomique = 1 si l'individu est marié
<i>Employé</i>	Statu d'emploi de l'individu	Dichotomique = 1 si l'individu est travailleur ou travailleuse
<i>Salaires</i>	Salaire mensuel principal	Discrète (125 000 - 20 000 000)
<i>Décile</i>	Groupe de revenu familial	Discrète (1-10)
<i>Proximité</i>	Université à proximité	Dichotomique = 1 s'il y a une université dans la ville de domicile
<i>Éducation</i>	Années de scolarité	Discrète (0-23)

### **Milieu familial**

La base de données attribue à chaque individu la variable *décile*, correspondant à un groupe de revenu familial. La procédure consiste à diviser l'échantillon total de ménages en dix groupes de la même taille, où le premier décile correspond aux ménages ayant le revenu le plus faible, et le 10ème décile correspond aux ménages ayant le revenu le plus élevé.

### **Proximité d'une université**

Dans le cadre de notre recherche, nous utiliserons la proximité d'une université comme un déterminant exogène de scolarité. Vu que cette variable n'est pas incluse dans l'ESR, nous avons créé une variable indicatrice égale à 1 si une université est située dans le marché du travail local. Celle-ci est construite à partir des données du système éducatif du Ministère de l'éducation chilien. En effet, en 2017, il y a 55 établissements universitaires accrédités à travers le territoire chilien qui offrent un programme de premier cycle de 4 ans ou plus. La base de données utilisée contient des codifications uniques pour la région et la ville où chaque université est située. En effet, dans l'ESR chaque individu se voit donner une de ces valeurs uniques. La procédure consiste alors à appairer les codes en question.

### **Salaires mensuels**

Pour simplifier l'analyse, nous avons choisi de travailler avec les salaires mensuels principaux, en excluant les transferts ainsi que les revenus de pensions, de location et de l'extérieur. La base contient des salaires imputés lorsqu'un individu ne déclare pas de revenu, même s'il fait partie de la population active.

En somme, après le traitement de données, l'échantillon final contient 42 830 observations correspondant à 33 859 ménages.



### 3. Méthodologie

#### 3.1 Modèle de scolarisation optimale

Cette section développe le modèle de capital humain conçu par Gary S. Becker (1967) et résumé par David Card (1995b), lequel nous servira comme cadre théorique pour expliquer l'effet causal de la scolarité.

Supposons un individu qui maximise une fonction d'utilité  $U$ . Les opportunités de l'individu sont regroupées dans une fonction  $y = g(S)$ , qui représente les niveaux de salaires associés à chaque année de scolarité. Soit une fonction  $\emptyset(S)$  qui représente les coûts d'opportunité pour un nombre d'années  $S$  de scolarité complétée.

Le problème de l'individu est le suivant :

$$\max U(y, S) = \log(y) - \emptyset(S) \quad (2)$$

où  $y$  est le salaire mensuel et  $S$  représente les années de scolarité. Afin de simplifier l'analyse, l'hypothèse posée ici assume que  $\emptyset$  est une fonction convexe croissante.

La condition de premier ordre est obtenue en égalisant à zéro la dérivée de (2) par rapport à  $S$  :

$$\frac{g'(S)}{g(S)} = \emptyset'(S) \quad (3)$$

Cette équation indique le niveau optimal de scolarité atteint lorsque le rendement marginal de la scolarité est égal au coût marginal.

Pour des fins d'interprétation, il est utile de définir l'utilité marginale et le coût marginal comme étant des fonctions homogènes linéaires pour un individu  $i$  :

$$\frac{g'(S)}{g(S)} \equiv \beta_i(S) = b_i - k_1 S \quad (4)$$

$$\phi'(S) \equiv \delta_i(S) = r_i + k_2 S \quad (5)$$

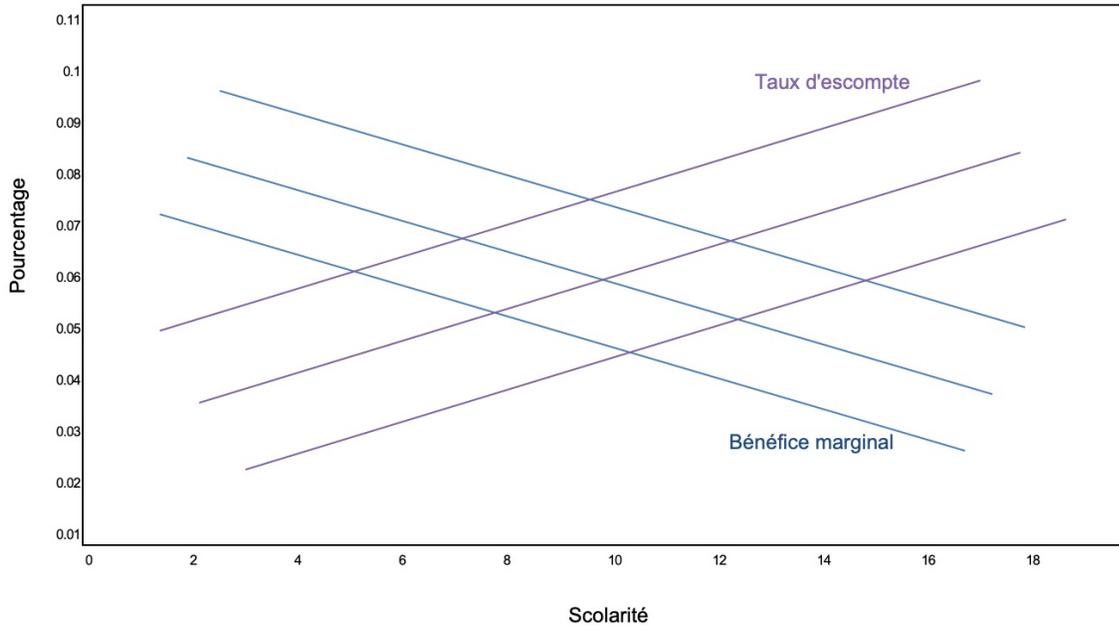
En égalisant les fonctions (4) et (5), nous obtenons l'expression pour le niveau optimal de scolarité  $S$  :

$$S = (b_i - r_i)/k \quad (6)$$

où  $k = k_1 + k_2$  et  $k_1, k_2 \geq 0$ .

Les bénéfices et coûts marginaux pour un individu  $i$ , ainsi que le niveau optimal de scolarité respective, sont illustrés dans la figure 3.1.

**Figure 3.1. Taux d'escompte et bénéfices marginaux**



Selon Card (1995b), si  $\delta_i(S)$  est une fonction constante, celle-ci peut être interprétée comme le taux marginal d'escompte pour l'individu  $i$ . Tel que soulevé par Becker (1967), ceci est le taux de rendement des fonds requis pour financer la  $S^{\text{e}}$  année de scolarité. Nous reprenons cette constatation pour notre analyse.

Dans ce modèle, les choix des individus diffèrent lorsque : a) ils ont différents rendements d'éducation (variation en  $b_i$ ) et b) ils ont différents taux d'escompte (variation en  $r_i$ ). Afin d'interpréter les concepts d'une façon intuitive, la variation de  $b_i$  correspond aux

différences en habileté, tandis que la variation de  $r_i$  correspond aux différences en opportunité ou liquidité.

Par hypothèse,  $b_i$  et  $r_i$  sont corrélés négativement, vu que l'habileté est un trait hérité ou influencé par le milieu familial. À titre d'exemple, les parents ayant un haut niveau d'habileté ( $b_i$ ) auront tendance à avoir des niveaux d'éducation et des salaires plus élevés. Sous cette logique, les enfants des familles ayant un  $r_i$  bas auront aussi un  $r_i$  bas et un  $b_i$  haut.

En intégrant la fonction (4), nous obtenons l'équation pour le logarithme des salaires de l'individu  $i$  :

$$\log y_i = a + b_i S_i - \frac{1}{2} k_1 S_i^2 \quad (7)$$

À partir des équations (6) et (7), il est possible de déterminer la relation entre les années de scolarité et les salaires. Cette relation est illustrée dans la figure 3.2.

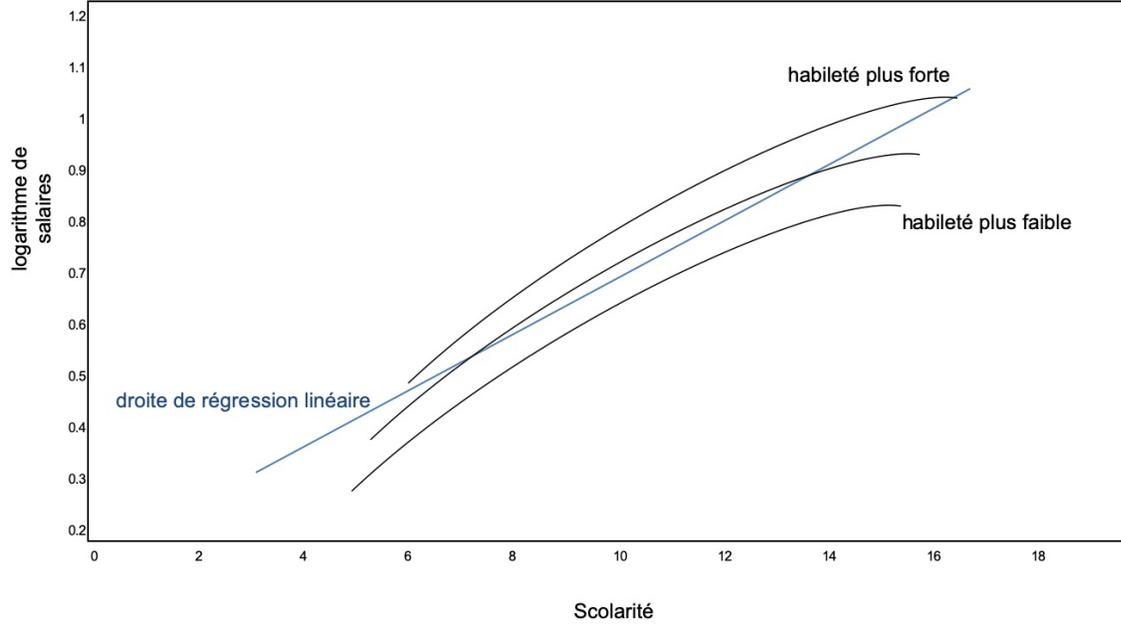
Afin de simplifier la dérivation du coefficient du rendement scolaire par moindres carrés ordinaires,  $\rho$ , nous établissons l'hypothèse que les individus maximisent la valeur actuelle des salaires à des taux d'escompte fixes. Donc, nous avons  $k_2 = 0$ . Par définition, le coefficient de régression du modèle de population est égal à l'expression suivante :

$$\rho = \text{cov}(\log y_i, S_i) \div \text{var}(S_i) \quad (8)$$

En utilisant l'opérateur d'espérance, le terme pour la covariance dans (8) peut se réécrire comme suit :

$$\rho = E(\log y_i) (S_i - \bar{S}) \div \text{var}(S_i) \quad (9)$$

**Figure 3.2. Niveau optimal de scolarité**



Où  $\bar{S}$  représente la moyenne de  $S_i$ .

En utilisant les équations (6) et (7), nous pouvons réécrire l'équation (9) comme suit:

$$\rho = E(\log y_i) (S_i - \bar{S}) = E \left[ b_i S_i (S_i - \bar{S}) - \frac{1}{2} k_1 S_i^2 (S_i - \bar{S}) \right] \quad (10)$$

$$= E \left[ b_i \frac{b_i - r_i (b_i - \bar{b}) - (r_i - \bar{r})}{k} - \frac{1}{2} k_1 S_i^2 (S_i - \bar{S}) \right] \quad (11)$$

Où  $\bar{b}$  et  $\bar{r}$  sont les moyennes de  $b_i$  et  $r_i$ , respectivement. Par hypothèse, nous supposons que les distributions de  $b_i$  et  $r_i$  sont symétriques. Sous cette hypothèse, le coefficient de régression du modèle de population est une moyenne pondérée de  $\bar{b}$  et  $\bar{r}$  et il peut s'écrire comme suit:

$$\rho = \lambda \bar{b} + (1 - \lambda) \bar{r} \quad (12)$$

où

$$\lambda = \frac{\sigma_b^2 - \sigma_{br}}{(\sigma_b^2 - \sigma_{br}) + (\sigma_r^2 - \sigma_{br})} \quad (13)$$

Dans l'équation (13),  $\sigma_b^2$  et  $\sigma_r^2$  sont les variances de  $b_i$  et  $r_i$ , respectivement, et  $\sigma_{br}$  est leur covariance. Dans le cas où  $\sigma_{br} = 0$ , l'équation (13) peut être interprétée comme la proportion de la variance du niveau de scolarité attribuée à la variation en habileté.

Finalement, nous procédons à comparer l'estimateur obtenu en (12) avec le bénéfice marginal moyen de l'éducation. À partir de l'équation (4) nous avons l'expression suivante :

$$\bar{\beta} = \bar{b} - k_1 \bar{S} \quad (14)$$

Avec l'équation (14), nous pouvons réécrire l'équation (12) comme suit :

$$\rho = \bar{\beta} + \lambda(\bar{b} - \bar{r}) \quad (15)$$

Selon l'équation (15) le coefficient estimé par MCO est égal au bénéfice marginal moyen de l'éducation plus un terme positif. Ce terme peut être interprété comme un biais d'endogénéité, dû au fait que les individus ayant un taux de rendement d'éducation plus haut choisissent des niveaux de scolarité plus élevés.

Comment l'a démontré Card (1995b), le cadre théorique présenté ci-dessus peut servir à expliquer pourquoi les estimateurs du rendement scolaire par VI sont supérieurs aux ceux par MCO lorsque les instruments utilisés pour l'estimation par VI visent à réduire les coûts d'accès à l'éducation. Puisque les individus choisissent d'investir en éducation quand le taux d'escompte est égal à leur bénéfice marginal, ceux ayant des taux d'escompte très élevés nécessitent un bénéfice marginal très élevé. En conséquence, les coefficients par VI auront tendance à capter cet effet.

### 3.2 Modèles de régression

Jacob Mincer est le premier académicien à avoir établi une relation statistique reliant le nombre d'années de scolarité au salaire d'un travailleur (Bédard, 2014). Après lui, de nombreux chercheurs ont contribué à cette relation en y intégrant plusieurs nouvelles variables explicatives susceptibles d'améliorer la prédiction du modèle (Lemieux, 2006).

Pour l'estimation du modèle structurel, nous avons repris le modèle standard utilisé par la plupart des études montrés dans le tableau sommaire du survol de la littérature, afin d'identifier le rendement de l'éducation. Soit l'équation suivante :

$$S_i = X_{1i}\gamma_i + C_i\delta_0 + C_i * P_i\delta_1 + v_i \quad (16)$$

Où  $i$  dénote un travailleur et chaque variable est définie comme suit :

$S_i$  : années de scolarité;

$C_i$  : variable indicatrice égale à 1 s'il y a une université dans la ville du travailleur  $i$ ;

$P_i$  : variable indicatrice égale à 1 si le travailleur appartient à un groupe familial à faible revenu;

$X_{1i}$  : vecteur de variables de contrôle telles que l'expérience potentielle, le statut matrimonial, la résidence dans une région métropolitaine et le sexe.

L'interaction  $C_i * P_i\delta_1$  trouvée dans l'équation (16) est utilisée comme un des instruments dans nos estimations.

L'équation du salaire mensuel  $S$  estimé est la suivante :

$$y_i = X_i\alpha + S_i\beta + u_i \quad (17)$$

## 4. Résultats

### 4.1 Caractéristiques des travailleurs et travailleuses chiliens

Le tableau 4.1 met en évidence certaines caractéristiques à partir de notre échantillon, afin de mieux comprendre les résultats de nos modèles économétriques. Nous constatons qu'un quart de l'échantillon habite dans la région métropolitaine chilienne. Le pourcentage des personnes ayant une université à proximité varie à travers les régions (plus bas pour les régions plus pauvres) et par sexe (plus haut pour les femmes). Le tableau nous montre aussi que les femmes sont plus éduquées que les hommes, en termes de moyenne d'années de scolarité et de pourcentage de personnes ayant des études post-secondaires. Finalement, nous constatons un écart salarial assez important entre les hommes et les femmes, en ajustant les montants pour la parité du pouvoir d'achat<sup>6</sup> en 2017.

**Tableau 4.1. Caractéristiques de l'échantillon – Enquête supplémentaire de revenus pour 2017**

	Tous	Hommes	Femmes
Région métropolitaine (%)	25,3	24,4	26,5
Domicile ayant une université à proximité (%)	52	50,4	54
Moyenne des années de scolarité	11,9	11,6	12,2
Salaire mensuel moyen (en dollars USD, 2017)	1 314	1 492	1 084
Études post-secondaires (%)	21,7	20,1	23,4

<sup>6</sup> L'indicateur est obtenu des données de l'OCDE.

## 4.2 Résultats des modèles d'estimation

La présente section porte sur l'analyse des coefficients estimés des régressions par MCO et par moindres carrés à deux étapes (2SLS).

Le tableau 4.3 présente une série d'estimateurs du rendement de l'éducation pour les travailleurs et travailleuses chiliens en utilisant la base de données décrite dans le chapitre 2. La variable dépendante pour le modèle d'éducation est « années de scolarité », tandis que pour le modèle de salaires la variable est le logarithme du salaire mensuel du travail principal (incluant le salaire des travailleurs indépendants). Le tableau 4.2 présente les variables de contrôle incluses dans toutes les spécifications :

**Tableau 4.2. Variables explicatives des modèles de régression**

<b>Variable</b>	<b>Description des variables (type de variable)</b>
<i>Proximité</i>	Université à proximité dans la ville du domicile du travailleur (binaire)
<i>Exp</i>	Expérience potentielle du travail (discrète)
<i>Exp2</i>	Expérience potentielle du travail au carré divisé par 100 (discrète)
<i>Marié</i>	Statut matrimonial de la personne (dichotomique)
<i>Metro</i>	Résidence dans une région métropolitaine (dichotomique)

Tableau 4.3. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation

	(1)	(2)	(3)	(4)
<b>Estimation par MCO</b>				
a. Éducation	0,1130 (0,0012)	-- --	-- --	-- --
<b>Estimateur de VI (instrument: proximité d'une Université)</b>				
a. Éducation	0,1724 (0,0067)	-- --	-- --	-- --
<b>Modèle d'éducation de forme réduite</b>				
c. Proximité	--	1,1390 (0,0304)	1,1071 (0,0314)	0,9817 (0,0352)
d. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,2754 (0,1193)	-0,4039 (0,0799)	-0,2469 (0,0585)
<b>Modèle de salaires de forme réduite</b>				
e. Proximité	--	0,1803 (0,0083)	0,1604 (0,0086)	0,1209 (0,0095)
f. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,0277 (0,0328)	-0,0532 (0,2190)	-0,0443 (0,0160)
<b>Estimateurs de VI (instrument: interaction proximité d'une Université * sous-groupe de revenu familial défavorisé)</b>				
g. Éducation	--	0,1008 (0,1112)	0,1317 (0,0517)	0,1795 (0,0664)
h. Proximité	--	0,0654 (0,1249)	0,0146 (0,0545)	-0,0552 (0,0598)
Variables de contrôle pour le milieu familial	oui	oui	oui	oui

Écarts types en parenthèses.  
Nombre d'observations: 42 830

Concernant notre choix de l'expérience potentielle comme variable explicative, nous reprenons les spécifications de Mincer (1974) et Murphy et Welch (1990), lesquelles constatent que les salaires sont mieux décrits comme fonction de l'éducation et l'expérience que comme fonction de l'éducation et de l'âge.

Chaque colonne du tableau 4.3 correspond à une spécification différente de nos estimations. La colonne 1) présente le modèle d'éducation estimé par MCO et par VI, en utilisant la proximité d'une université comme instrument. Les colonnes 2-4 présentent une série d'estimations des modèles réduits et leurs modèles structurels correspondants, en utilisant l'interaction de la proximité d'une université et l'appartenance à un milieu familial défavorisé comme instrument. La colonne 2) utilise le premier décile de revenu familial comme variable indicatrice d'un environnement familial défavorisé. La colonne 3) utilise une variable indicatrice égale à 1 si l'individu appartient au décile 1 et 2, tandis que la colonne 4) utilise une variable indicatrice égale à 1 si le ménage de l'individu se place en dessous de la médiane des revenus familiaux. Celles-ci sont résumées dans le tableau 4.4.

**Tableau 4.4. Variables indicatrices selon les sous-groupes de revenus familiaux**

<b>Variable</b>	<b>Définition (type de variable)</b>
Pauvre1	Personne dans le premier décile de la distribution de revenu familial (dichotomique)
Pauvre2	Personne dans le premier et deuxième décile la distribution de revenu familial (dichotomique)
Pauvre5	Personne au-dessous de la médiane de la distribution de revenu familial (dichotomique)

La première colonne du tableau 1 montre le coefficient du modèle standard estimé par moindres carrés ordinaires. Toutes choses étant égales par ailleurs, le rendement d'une année additionnelle de scolarité est estimé à 11,3 %, en utilisant les variables de contrôle spécifiées dans le tableau 4.2. On constate que notre coefficient se rapproche de la moyenne du rendement de l'éducation pour les pays à revenus moyens estimée par Sakellariou (2007).

Malgré les avantages en termes de comparabilité et stabilité de l'estimateur par MCO, il est possible que l'estimateur du vrai rendement de l'éducation soit biaisé. Pour justifier cette observation, considérons un système de deux équations décrivant le logarithme de salaires mensuels  $y_i$  et les années de scolarité  $S_i$ :

$$S_i = X_i\gamma + v_i \quad (18)$$

$$y_i = X_i\alpha + S_i\beta + u_i \quad (19)$$

où  $X_i$  est un vecteur des variables observées par l'économètre,  $\beta$  représente le rendement du vrai modèle d'éducation,  $u_i$  et  $v_i$  sont des termes d'erreur, avec  $E(X_i u_i) = E(X_i v_i) = 0$ .

Pour que le coefficient estimé de  $\beta$  soit convergent, la condition requise est que les termes d'erreur  $u_i$  et  $v_i$  ne soit pas corrélés, c'est-à-dire, que  $S_i$  soit une variable exogène dans le modèle de salaires.

Selon Card (1995b), il existe au moins deux raisons pour lesquelles le niveau de scolarité est corrélé avec les habiletés non observées par l'économètre. La première raison est reliée aux problèmes de variable omise, particulièrement le « biais d'habileté ». Supposons qu'il existe un individu ayant des habiletés non observées et que celles-ci lui permettent d'augmenter son salaire indépendamment de son niveau d'éducation. Si cet individu possède un niveau de scolarité supérieur à la moyenne, le coefficient de  $\beta$  estimé par MCO sera biaisé à la hausse. Tel que remarqué par Card (1995b), le fait que les individus

ayant des bons résultats aux tests standardisés ont tendance à avoir des salaires plus élevés est souvent établi comme évidence du biais d'habileté.

La deuxième raison de la corrélation entre les termes d'erreur  $u_i$  et  $v_i$  est l'erreur de mesure dans  $S_i$ . Cette erreur provoque un biais d'atténuation vers zéro si les erreurs sont indépendantes d'une observation à l'autre et indépendantes de la vraie valeur de  $S_i$ . Tel que constaté par Parent (2018), il est tout simplement difficile de mesurer des associations lorsque les données contiennent du bruit. On observe un biais similaire si les individus ayant des niveaux de scolarité bas ont des rendements élevés, tel qu'observé par Card (1995b) et Lang (1993).

Dû aux deux raisons précédentes, il est requis de briser la corrélation entre le terme d'erreur et  $S_i$ , c'est-à-dire d'obtenir une variable  $z_i$  contenue dans  $X_i$  qui n'ait aucun effet sur  $y_i$  au-delà de son association avec  $S_i$  et qui soit corrélée le plus fortement possible avec  $S_i$ . Sous une assignation aléatoire des niveaux de scolarité parmi la population, la réalisation du processus aléatoire pourrait être utilisée pour estimer l'équation des salaires par la méthode des VI. Cependant, en l'absence de cette situation, nous avons besoin d'identifier un déterminant de la scolarité et l'exclure du modèle de salaires.

Il est possible que la présence d'une université à proximité dans la ville du domicile de l'individu soit cette variable recherchée. Il est clair que les élèves ayant grandi dans une ville sans université sont plus sensibles aux coûts éducatifs. En ce sens, ce coût réduit les possibilités d'investir dans l'éducation supérieure, et ce particulièrement parmi les élèves appartenant aux ménages à faible revenu, tel que soulevé par Card (1995b).

Pour reprendre les constatations précédentes, nous avons estimé le modèle structurel des rendements de l'éducation, en utilisant la proximité d'une université comme instrument pour les années de scolarité complétées. Les résultats sont présentés dans la colonne (1) du tableau 4.3. L'effet est très grand dans les salaires (17,2%) pour un individu ayant une université à proximité de sa ville de domicile. L'estimateur par VI est 50 % plus élevé que l'estimateur par MCO. Notre estimation se rapproche de celle trouvée par Card (1995a, 1995b), Butcher et Case (1993), Kane et Rouse (1993) et Angrist et Krueger (1993).

Les sections a) et b) des colonnes 2-4 présentent les résultats de l'estimation des modèles d'éducation et salaires de forme réduite. Tel qu'indiqué dans les paragraphes précédents, la proximité d'une université est définie comme un déterminant exogène de la scolarité. Prenons d'abord la colonne 2, qui utilise le premier décile de revenus comme définition d'un ménage défavorisé. Le coefficient positif de proximité confirme que l'effet d'avoir une université à proximité est très significatif. Toute chose étant égale par ailleurs, l'effet pour les individus ayant une université à proximité dans leur ville de domicile est de 1,13 années de scolarité additionnelles. Concernant le modèle de salaires de forme réduite, le coefficient estimé est aussi positif et significatif. Toute chose étant égale par ailleurs, l'effet pour les individus ayant une université à proximité dans leur ville de domicile est une augmentation de 18 % dans leur salaire mensuel.

Il est également intéressant de constater à quel point l'impact diffère selon la position de l'individu sur la distribution des revenus familiaux. Ces effets sont présentés dans les sections c) et e) des colonnes (3) et (4), qui utilisent des sous-groupes distincts, selon les spécifications 2 et 3 du tableau 4.4. Les effets pour les individus ayant une université à proximité dans leur ville de domicile est de 1,11 et 0,98 années de scolarité additionnelles respectivement. Concernant le modèle de salaires, le fait d'avoir une université à proximité représente une augmentation de 16 % et 12 % pour chaque sous-groupe.

### **4.3 Validité et légitimité de l'instrument**

Il est important de remarquer que la variable créée pour tenir compte de la proximité d'une université a été construite à partir de la ville de domicile de l'individu. Tel que remarqué par Card (1995b), cette variable est plus utile quand elle indique la ville de domicile lorsque l'individu avait 18 ans, puisque la décision de poursuivre des études supérieures se fait d'habitude à ce moment. Cette situation affecte particulièrement les individus plus âgés inclus dans la base de données, qui pourraient avoir déménagé dans une ville ayant une université à proximité. Par conséquent, il est possible qu'une relation de causalité inverse se produise entre la proximité d'une université et la réussite scolaire.

La crédibilité de notre instrument repose sur le fait que la proximité d'une université n'ait aucun effet direct sur les salaires. Card (1995b) considère trois facteurs pour lesquels la présence d'une université dans une ville affecte différemment les salaires des hommes nord-américains, en contrôlant par des variables démographiques :

1. Il est possible que les familles qui mettent une forte valeur à l'éducation de leurs enfants décident de s'établir proche d'une université.
2. Il est possible que la présence d'une université ait une influence sur la qualité de l'éducation primaire et secondaire.
3. Si c'est le cas, les hommes qui ont grandi dans une ville ayant une université ont une tendance à choisir ces villes pour s'établir, il est possible que la proximité soit corrélée avec les primes de salaires géographiques non observées.

Il est suggéré par Card (1995b) que le fait que la proximité d'une université représente une réduction dans les coûts de l'éducation, l'effet est plus important pour les individus appartenant aux ménages plus défavorisés. Le tableau 4.3 présente aussi l'estimation des modèles d'éducation et salaires par moindres carrés à deux étapes, en utilisant l'interaction entre la proximité d'une université et l'appartenance à un milieu familial défavorisé. La section c) des colonnes 2-4 présentent les estimations en utilisant les trois définitions de ménages défavorisés, tel que spécifié dans le tableau 4.4.

Contrairement aux résultats constatés par Card (1995a, 1995b), les effets sur l'éducation et les salaires sont négatifs pour les individus provenant d'un milieu familial défavorisé, et ce pour chacun des trois sous-groupes. Prenons d'abord les résultats de la colonne 2. Pour les individus les plus défavorisés, l'effet d'avoir une université à proximité représente une réduction de 0,277 années de scolarité. Dans notre modèle de salaires, cette réduction représente 2,7 %.

L'effet négatif est plus grand pour les déciles 1 et 2 de la distribution de revenus familiaux. L'effet d'avoir une université à proximité représente une réduction de 0,4 années de scolarité. Dans notre modèle de salaires, cette réduction représente 5,32 %.

À partir des résultats de la colonne 4 pour le modèle d'éducation, on observe que l'effet est encore négatif mais décroissant à mesure que l'on relâche la définition de ménage défavorisé. En prenant les ménages qui se placent au-dessous de la médiane de revenus comme sous-groupe, l'effet d'avoir une université à proximité représente une réduction de 0,25 années de scolarité. Dans notre modèle de salaires, cette réduction est la plus marquée à 5,32%.

#### **4.4 Discussion des résultats et analyse du contexte institutionnel chilien**

Bien que les résultats ci-dessus ne soient pas les mêmes que ceux trouvés par Card, il est possible que les différences soient attribuables au contexte institutionnel chilien. Un facteur qui pourrait expliquer les résultats est la définition des ménages défavorisés au Chili et les contraintes de liquidité subies par la population la plus vulnérable.

Selon l'Institut national de statistiques au Chili, 50 % des travailleurs et travailleuses chiliens ont un salaire mensuel égal ou inférieur à \$USD 943,60<sup>7</sup> en 2017. De plus, 71 % des travailleurs et travailleuses ont un salaire égal ou inférieur à la moyenne nationale, \$USD 1 378.

Afin d'avoir un portrait global des contraintes de liquidité reliées à l'accès à l'éducation supérieure, les frais de scolarité moyens mensuels sont de \$ USD 695 par étudiant (Meller, 2011), tandis que le coût du panier de consommation pour la taille moyenne d'une famille économique s'élève à \$USD 458 (Ministère du développement social et famille, 2018)<sup>8</sup>. Les montants précédents nous renseignent sur les contraintes de liquidité subies chez les familles au-dessous de la médiane de la distribution de revenus familiales, particulièrement pour les familles les plus pauvres. Il est possible que ces contraintes leur empêchent de faire des investissements en éducation supérieure, même si les coûts diminuent grâce à la proximité d'une université. Cependant, les individus qui semblent en profiter se trouvent en haut de la distribution de revenus familiaux.

---

<sup>7</sup> Montants rajustés en fonction de la parité du pouvoir d'achat (PPA).

<sup>8</sup> Montants rajustés en fonction de la parité du pouvoir d'achat (PPA).

Cette situation met en évidence l'inégalité de la distribution de revenus, qui pourrait donc expliquer les résultats de notre estimation.

Pour confirmer les constatations précédentes, nous avons estimé les modèles d'éducation et salaires de forme réduite, en définissant le groupe de ménages défavorisés au-dessous du 9<sup>ème</sup> décile. En utilisant l'interaction de proximité et ce sous-groupe, les effets additionnels du rendement de l'éducation sont 0,08, avec un écart type de 0,0699 et 0,03, avec un écart type de 0,0191, pour les modèles d'éducation et salaires respectivement.

En conclusion, même si les coefficients des modèles de forme réduite sont imprécis, ils suggèrent que c'est la classe moyenne qui profite d'avoir une université à proximité. Pour les individus appartenant aux ménages défavorisés, le taux d'escompte est très élevé et ils nécessitent un très haut taux de rendement pour pouvoir bénéficier des études universitaires, tel que vu par les résultats obtenus dans le cadre de variables instrumentales.

#### **4.5 Interaction de proximité et milieu familial défavorisé comme instrument pour l'éducation complétée**

Les lignes g) et h) et les colonnes 2-4 du tableau 4.3 présentent les résultats de l'estimation du modèle de salaires pour les variables instrumentales en utilisant comme instrument l'interaction entre la proximité d'une université et la provenance d'un milieu familial défavorisé, et ce, pour chacune des définitions des ménages défavorisés, tel que spécifié dans le tableau 4.4. L'estimateur par variables instrumentales et l'effet direct sont indiqués. En utilisant le premier décile de la distribution de revenus familiaux, le rendement de l'éducation est plus faible que celui estimé dans la colonne (1) est l'écart type est plus grand. Par contre, l'effet direct de la proximité d'une université sur les salaires est positif, mais non significatif.

En utilisant les déciles 1 et 2, le coefficient du rendement de l'éducation est supérieur à ceux estimés dans la colonne (1). L'effet direct de la proximité d'une université sur les salaires est très petit, mais non significatif.

La section c) de la colonne 4 est plus intéressante. En utilisant le sous-groupe qui se place au-dessous de la médiane de revenus familiaux, l'estimateur du rendement de l'éducation se rapproche à celui qui utilise seulement la proximité d'une université. En plus de ces précédents résultats, les coefficients de l'effet direct de la proximité sur les salaires sont petits et non significatifs pour les trois spécifications.

En résumé, notre analyse confirme les résultats des études empiriques qui estiment le rendement de l'éducation par variables instrumentales. En tenant compte que l'effet de l'instrument utilisé est plus fort pour les individus qui obtiendrait des niveaux de scolarité faibles, le coefficient par variables instrumentales est plus élevé que le coefficient estimé par MCO.

Selon la théorie économique derrière le modèle de Becker, le niveau optimal de scolarité pour l'individu est atteint quand son bénéfice marginal est égal à son taux d'escompte. Dans ce cadre, deux scénarios expliquent un niveau de scolarité bas pour l'individu : son rendement de l'éducation est bas ou son taux d'escompte est élevé. Si une université à proximité incite les individus de ce sous-groupe à augmenter leur niveau de scolarité, le rendement associé sera plus élevé que le rendement moyen pour toute la population si le taux d'escompte est plus élevé que le rendement de l'éducation.

#### **4.6 Différences de rendement de l'éducation entre hommes et femmes**

Nous avons refait le même exercice que dans le tableau 4.3 en considérant les hommes et les femmes séparément. Les tableaux 4.5 et 4.6 montrent les résultats des modèles de régression par moindres carrés ordinaires, par variables instrumentales et pour les trois spécifications présentées dans le tableau 4.2.

La première colonne montre le coefficient du modèle standard estimé par moindres carrés ordinaires. Le coefficient pour les femmes est 30 % supérieur à celui pour les hommes. En effet, le rendement d'une année additionnelle de scolarité pour les femmes est estimé

à 13,5 %. Ce résultat est cohérent avec ceux obtenus dans plusieurs études dans la littérature (Psacharopoulos, 2004). Selon Dougherty (2005), ce phénomène pourrait s'expliquer par le fait que l'éducation semble avoir un effet double sur les salaires chez les femmes : a) les gains considérables en matière de scolarité contribuent à augmenter leur productivité (effet direct), comme chez les hommes; et b) l'accroissement du niveau de scolarité contribue à réduire l'écart salarial entre les femmes et les hommes attribué aux facteurs de discrimination faite sur le marché du travail et des caractéristiques et circonstances personnelles (effet indirect).

Pour les hommes, le coefficient estimé est de 10,38 %. Même si cet estimateur est supérieur à ceux des modèles qui utilisent des données pour les États-Unis ou même pour les pays développés, il est plus proche que celui estimé par Sakellariou (2007) dans son modèle de rendements de l'éducation pour un échantillon d'hommes chiliens inclus dans l'Enquête internationale d'alphabetisation des adultes (IALS en langue anglaise).

La deuxième ligne de la colonne (1) présente l'estimation du modèle des salaires en utilisant la proximité d'une université comme instrument pour les années de scolarité complétées. Sous cette spécification, les femmes semblent bénéficier plus d'avoir une université dans leur ville de domicile. En effet, l'estimateur par variables instrumentales pour les femmes est 55 % plus élevé que pour les hommes.

Tableau 4.5. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation - hommes

	(1)	(2)	(3)	(4)
<b>Estimation par MCO</b>				
a. Éducation	0,1038 (0,0014)	-- --	-- --	-- --
<b>Estimateur de VI (instrument: proximité d'une Université)</b>				
a. Éducation	0,1532 (0,0073)	-- --	-- --	-- --
<b>Modèle d'éducation de forme réduite</b>				
c. Proximité	--	1,2732 (0,0414)	1,2516 (0,0428)	1,1408 (0,0478)
d. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,3558 (0,1573)	-0,5018 (0,0700)	-0,3504 (0,0793)
<b>Modèle de salaires de forme réduite</b>				
e. Proximité	--	0,1799 (0,0100)	0,1622 (0,0103)	0,1336 (0,0114)
f. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,0338 (0,0383)	-0,0580 (0,0257)	-0,0731 (0,0190)
<b>Estimateurs de VI (instrument: interaction proximité d'une Université * sous-groupe de revenu familial défavorisé)</b>				
g. Éducation	--	0,0950 (0,0997)	0,1156 (0,0486)	0,2086 (0,0606)
h. Proximité	--	0,0589 (0,1249)	0,0175 (0,0576)	-0,1044 (0,0623)
Variables de contrôle pour le milieu familial	oui	oui	oui	oui

Nombre d'observations: 24 144

Écarts types en parenthèses.

Tableau 4.6. Coefficients estimés par MCO et VI du modèle de rendement de l'éducation - femmes

	(1)	(2)	(3)	(4)
<b>Estimation par MCO</b>				
a. Éducation	0,1350 (0,0020)	-- --	-- --	-- --
<b>Estimateur de VI (instrument: proximité d'une Université)</b>				
a. Éducation	0,2380 (0,0131)	-- --	-- --	-- --
<b>Modèle d'éducation de forme réduite</b>				
c. Proximité	--	0,9391	0,8946	0,7426
	--	0,0443	(0,0458)	(0,0513)
d. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,1825	-0,2808	-0,1255
	--	(0,1817)	(0,1198)	0,0854
<b>Modèle de salaires de forme réduite</b>				
e. Proximité	--	0,2087	0,1844	0,1266
	--	(0,0133)	(0,0136)	(0,0152)
f. Proximité d'une université * sous-groupe de revenu familial défavorisé	--	-0,0509	-0,0600	-0,0067
	--	(0,0546)	(0,0356)	(0,0253)
<b>Estimateurs de VI (instrument: interaction proximité d'une Université * sous-groupe de revenu familial défavorisé)</b>				
g. Éducation	--	0,2788	0,2137	0,0534
	--	(0,3159)	(0,1254)	(0,1929)
h. Proximité	--	-0,0530	-0,0067	0,0869
	--	0,2936	(0,1078)	(0,1350)
Variables de contrôle pour le milieu familial	oui	oui	oui	oui

Nombre d'observations: 18 686

Écarts types en parenthèses.

Nous avons procédé par la suite à estimer les modèles d'éducation et salaires de forme réduite, en séparant les hommes et les femmes. Les sections a) et b) des colonnes 2-4 présentent les résultats de l'estimation. Chaque spécification inclut une définition différente des sous-groupes de revenu familial, tel qu'indiqué dans le tableau 3.4. Les résultats du modèle d'éducation de forme réduite des colonnes (2) et (4) mettent en évidence que les femmes ont des rendements inférieurs à ceux des hommes, peu importe la définition des ménages défavorisés. Toute chose étant égale par ailleurs, l'effet d'avoir une université à proximité pour les femmes est de 0,94 années de scolarité additionnelles. Pour les hommes, le coefficient estimé s'élève à 1,27 années de scolarité.

L'effet sur le sous-groupe de femmes appartenant aux ménages à revenu faible est négatif. Cela est particulièrement le cas pour la spécification de la colonne (1), qui utilise le sous-groupe du revenu familial le plus faible. Pour ce sous-groupe de femmes, l'effet d'avoir une université à proximité est de 0,74 années de scolarité additionnelles. Les effets additionnels négatifs sur les hommes et femmes sont respectivement -0,50 et -0,28 pour ceux appartenant aux déciles 1 et 2 de revenu familial et -0,35 et -0,12 pour ceux appartenant aux ménages au-dessous de la médiane de revenu familial.

Sur la base des résultats des sections e) et f), les effets sont plus importants chez les femmes appartenant aux groupes à faible revenu. Outre l'estimation pour le sous-groupe de la colonne (4), les résultats des tableaux nous montrent que les effets d'avoir une université à proximité sur les salaires sont plus grands chez les femmes que chez les hommes.

Finalement, nous avons également estimé les rendements de l'éducation par variables instrumentales pour chacun des groupes spécifiés dans le tableau 4.4. Outre que les régressions estimées des colonnes (3) et (4) chez les hommes, les coefficients sont imprécis mais supérieurs à l'estimateur par moindres carrés ordinaires. Cela est conforme avec les études mentionnées dans notre survol de la littérature scientifique sur le sujet.

En somme, les derniers paragraphes nous donnent un portrait global sur les différences de rendement de l'éducation et que les femmes ont des rendements de l'éducation plus élevés que les hommes.

## Conclusion

Cette recherche traite de l'estimation des rendements de l'éducation pour les travailleurs et travailleuses chiliens. En utilisant les données de l'Enquête supplémentaire de revenus (ESR), nous constatons que les personnes ayant une université à proximité ont des niveaux de scolarité et de salaires plus élevés. En tenant compte de différents sous-groupes de revenu familial, nous estimons 0,98-1,13 années de scolarité additionnelles et une augmentation de 12%-18% dans les salaires mensuels.

Nous avons procédé par la suite à estimer les effets additionnels de la scolarité et des salaires. Nous constatons des effets négatifs pour les individus en bas de la distribution des revenus familiaux. À mesure qu'on relâche la définition de ménage défavorisé en incluant les groupes en haut de la distribution de revenus familiaux, l'effet additionnel est encore négatif mais décroissant. Ceci confirme le résultat que les bénéfices se concentrent sur une masse très grande de travailleurs et travailleuses et pas seulement sur les individus appartenant à un milieu familial défavorisé. Quant à ceux-ci, ils nécessitent un très haut taux de rendement pour pouvoir bénéficier des études post-secondaires. Ceci nous renseigne que les estimations basées sur des facteurs exogènes qui affectent les décisions de scolarité sont différentes pour les pays en voie de développement.

Le contexte institutionnel chilien peut nous aider à comprendre les résultats de notre recherche. Même si le salaire mensuel moyen des Chiliens est parmi les plus élevés des pays latino-américains, des écarts en termes de répartition égalitaire de revenus persistent (OCDE, 2019b). Il est possible que cette situation amène à des crises politiques. Nous pensons par exemple aux séries de manifestations violentes qui ont pris place en octobre 2019, où presque un million de personnes ont participé à une des plus grandes protestations de l'histoire du pays, en réaction à des mesures d'augmentation des prix des services publics.



## Bibliographie

Angrist, Joshua D. et Alan B. Krueger (1991). « Does compulsory school attendance affect schooling and earnings? », *The Quarterly Journal of Economics*, vol. 106, no 4, p. 979-1014.

Angrist, Joshua D. et Alan B. Krueger (1993). *Split sample instrumental variables* [document inédit], New Jersey, Princeton University Department of Economics.

Ashenfelter, Orley et Alan Krueger (1994). « Estimates of the economic return to schooling from a new sample of twins », *The American Economic Review*, vol. 84, no 5, p. 1157-1173.

Becker, Gary S. (1967). *Human capital and the personal distribution of income*, Ann Arbor, University of Michigan Press.

Bédard, Alexandre (2014). *Écarts salariaux entre les femmes et les hommes dans l'industrie américaine : Une application des régressions quantiles non conditionnelles et de la décomposition d'Oaxaca-Blinder* [mémoire de maîtrise, version électronique], Montréal, HEC Montréal, 156 p.

Blau, Francine D. (1996). « Symposium on primary and secondary education », *Journal of Economic Perspectives*, vol. 10, no 4, p. 3-8.

Butcher, Kristin F. et Anne Case (1994). « The effect of sibling sex composition on women's education and earnings », *Quarterly Journal of Economics*, vol. 109, no 3, p. 531-564.

Card, David (1995a). « Earnings, schooling and ability revisited », dans Solomon W. Polacheck (dir.), *Research in labor economics*, vol 14, Greenwich, Connecticut et Londres, Jai Press Inc., p. 23-48.

Card, David (1995b). *Using geographic variation in college proximity to estimate the return to schooling*, Toronto, Buffalo et Londres, University of Toronto Press, coll. Aspects of labour market behaviour: Essays in honour of John vanderkamp.

Card, David (2001). « Estimating the return to schooling: Progress on some persistent econometric problems », *Econometrica*, vol. 69, no 5, p. 1127-1160.

Carrillo, F. et J. Ramos (2016). « Formación a lo largo de la vida: Motor de productividad », dans F. Bertranou et R. Astorga (dir.), *Chile: Desafíos de la productividad y el mundo laboral*, 1<sup>e</sup> éd, Santiago, Organisation internationale du travail.

Carrilo, Francisco, Sebastián Espinoza et Andrea Valenzuela (2018). « Mercado laboral y educación en Chile: Principales tendencias y resultados ». Récupéré le 17 juin de

[https://www.comisiondeproductividad.cl/wp-content/uploads/2018/04/Nota\\_Tecnica\\_Mercado\\_laboral.pdf](https://www.comisiondeproductividad.cl/wp-content/uploads/2018/04/Nota_Tecnica_Mercado_laboral.pdf)

Centro de Estudios MINEDUC, División de planificación et presupuesto (2017). « Informe nacional: Revisión de las políticas educativas ». Récupéré le 4 mai de [https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/06/CBR\\_Mineduc-WEB.pdf](https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/06/CBR_Mineduc-WEB.pdf)

Hanushek, Eric A. et Dennis D. Kimko (2000). « Schooling, labor-force quality, and the growth of nations », *The American Economic Review*, vol. 90, no 5, p. 1184-1208.

Hanushek, Eric A. et Ludger Woessman (2008). « The role of cognitive skills in economic development », *Journal of Economic Literature*, vol. 46, no 3, p. 607-668.

Institut national de statistiques (2018). *Síntesis de resultados ESI, encuesta suplementaria de ingresos 2017*. Récupéré de [https://www.cooperativa.cl/noticias/site/artic/20180718/asocfile/20180718102428/s\\_ntesis\\_encuesta\\_suplementaria\\_de\\_ingresos\\_2017\\_nacional.pdf](https://www.cooperativa.cl/noticias/site/artic/20180718/asocfile/20180718102428/s_ntesis_encuesta_suplementaria_de_ingresos_2017_nacional.pdf)

Institut national de statistiques (2019). Encuesta suplementaria de ingresos. Récupéré le 7 juin 2019 de <https://datosabiertos.ine.cl/datasets/187196/esi-2017/>

Kane, Thomas J. et Cecilia E. Rouse (1993). *Labor market returns to two- and four-year colleges: Is a credit a credit and do degrees matter?* [Working Paper no. 311], Princeton University Industrial Relations Section.

Lang, Kevin (1993). *Ability bias, discount rate bias, and the return to education*, Boston, Boston University Department of Economics.

Lemieux, Thomas (2006). « The “Mincer equation”, thirty years after schooling, experience and earnings », dans Shoshana Grosbard (dir.), *Jacob Mincer, a pioneer of labor economics*, New York Springer Science+Business Media.

Lemieux, Thomas et David Card (2001). « Education, earnings, and the ‘Canadian G.I. Bill’ », *The Canadian Journal of Economics / Revue canadienne d’Économique*, vol. 34, no 2, p. 331-344.

Meller, Patricio (2011). *Universitarios, ¿el problema no es el lucro, es el mercado!*, Santiago de Chile, Uqbar Editores.

Mincer, Jacob (1974). *Schooling, experience and earnings*, New York, National Bureau of Economic Research.

Ministère du développement social et famille (2018). *Valor de la canasta básica de alimentos y líneas de pobreza*. Récupéré de [http://observatorio.ministeriodesarrollosocial.gob.cl/ipc\\_pob\\_informe3.php?ano=2019](http://observatorio.ministeriodesarrollosocial.gob.cl/ipc_pob_informe3.php?ano=2019)

Murphy, Kevin M. et Finis Welch (1990). « Empirical age-earnings profiles », *Journal of Labor Economics*, vol. 8, no 2, p. 202-229.

OCDE (2018). Map of Chilean schooling system. Récupéré le 3 mars 2019 de <http://gpseducation.oecd.org/CountryProfile?primaryCountry=CHL>

OCDE (2019a). Purchasing power parities (ppp) (indicator). Récupéré le 26 octobre 2019 de [https://www.oecd-ilibrary.org/finance-and-investment/purchasing-power-parities-ppp/indicator/english\\_1290ee5a-en](https://www.oecd-ilibrary.org/finance-and-investment/purchasing-power-parities-ppp/indicator/english_1290ee5a-en)

OCDE (2019b). Inégalité de revenu (indicateur). Récupéré le 26 octobre 2019 de <https://data.oecd.org/fr/inequality/inegalite-de-revenu.htm>

OCDE (2019c). Chile Country Note. Programme for International Student Assessment (PISA). Results from PISA 2018. Récupéré le 25 janvier 2020 de [http://www.oecd.org/pisa/publications/PISA2018\\_CN\\_CHL.pdf](http://www.oecd.org/pisa/publications/PISA2018_CN_CHL.pdf)

Organisation des nations unies pour l'éducation, la science et la culture (UNESCO) (2010). Datos mundiales de la educación, UNESCO-IBE. Récupéré le 5 mai 2019 de [http://www.ibe.unesco.org/fileadmin/user\\_upload/Publications/WDE/2010/pdf-versions/Chile.pdf](http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/pdf-versions/Chile.pdf)

Parent, Daniel (2018). *8081603 - économétrie appliquée*, Endogénéité - séances 3 et 4, session d'été [présentation PowerPoint], HEC Montréal.

Patrinos, H. A. (1995). *Education and earnings differentials*, Washington, DC, Banque mondiale.

Psacharopoulos, George et Harry Anthony Patrinos (2004). « Returns to investment in education: A further update », *Education Economics*, vol. 12, no 2, p. 111-134.

Psacharopoulos, George et Harry Anthony Patrinos (2018). « Returns to investment in education: A decennial review of the global literature », *Education Economics*, vol. 26, no 5, p. 445-458.

Sakellariou, Chris (2007). *Cognitive ability and returns to schooling in Chile* [document inédit], Singapore, Division of Economics, Humanities and Social Sciences, Nanyang Technological University, 31 p.

Topel, Robert H. (1999). « Labor markets and economic growth », dans Orley Ashenfelter et Richard Layard (dir.), *Handbook of labor economics*, vol 3C, Amsterdam et New York, North Holland.

Urzúa, Sergio (2004). « Racial labor market gaps: The role of abilities and schooling choices », *The Journal of Human Resources*, vol. 43, no 4.

Willis, Robert J. (1986). « Wage determinants: A survey and reinterpretation of human capital earnings functions », *Handbook of Labor Economics*, vol. 1, p. 525-602