Capital humaine

Sam Gyetvay

ECO8000

September 17, 2025

Pourquoi différentes personnes gagnent-elles des salaires différents?

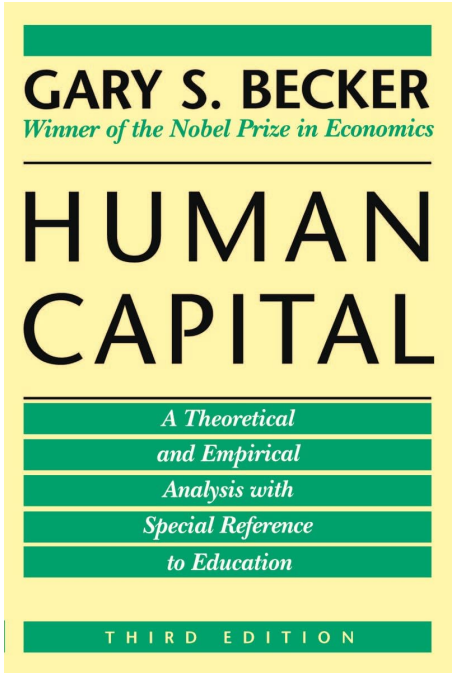
Dans le modèle de Roy, les travailleurs avaient différents niveaux de productivité dans la chasse, la pêche, etc.

Cependant, nous n'avons pas parlé de *pourquoi* les travailleurs sont différents. Nous avons simplement supposé que les personnes sont différentes *a priori*

Certaines différences sont **exogènes**, comme dans le modèle de Roy. Par exemple, certaines personnes sont plus grandes et donc meilleures au basketball. Comme on ne peut pas choisir ses gènes, on ne peut pas influencer sa taille : c'est en dehors ("exo") de notre contrôle

D'autres différences sont **endogènes** : elles sont le résultat de décisions que nous avons prises concernant des variables sous notre contrôle

En économie du travail, nous appelons la capacité productive qui résulte de décisions intentionnelles **capital humain**


GARY S. BECKER
Winner of the Nobel Prize in Economics
HUMAN CAPITAL
A Theoretical and Empirical Analysis with Special Reference to Education
THIRD EDITION

Capital humain

Il existe une énorme littérature sur le capital humain en économie. Becker (1962) a $>$ 17000 citations selon Google scholar

Le capital humain inclut l'éducation formelle, les apprentissages, la formation, les compétences acquises sur le lieu de travail par l'expérience, etc

Cependant, aujourd'hui nous allons essentiellement nous concentrer sur la question:

Quel est l'effet causal de l'éducation sur les revenus?

et sur les méthodes économétriques que les économistes du travail ont utilisées pour répondre à cette question

Nous concluons par une discussion de **Arteaga (2018)**, qui demande : l'éducation augmente-t-elle les revenus par le capital humain, ou par le **signalement** (ang. **signalling**)?

Quel est l'effet causal de l'éducation sur les revenus?

Pour répondre à cette question, nous devons d'abord définir ce que nous entendons par **effet causal**

Nous le faisons en utilisant le cadre des **résultats potentiels** (angl. **potential outcomes**)

Nous allons ensuite étudier un modèle simple où les travailleurs choisissent combien de temps passer à l'école, en équilibrant l'augmentation des revenus futurs avec le **coût d'opportunité** des revenus perdus dans le présent

Nous étudierons ensuite un article classique sur la façon d'estimer les rendements causaux de la scolarité (Angrist et Krueger, 1991), qui utilise une **variable instrumentale** pour estimer

Ensuite, nous examinerons des travaux plus récents. Finalement, je vais vous montrer un modèle simple de **signalement** pour motiver notre discussion d'Arteaga (2018)

Résultats potentiels (Potential outcomes)

Considérons une personne i décidant si elle doit aller à l'université

La variable indicatrice $D_i \in \{0, 1\}$ vaut 1 si i va à l'université, et 0 sinon

$Y_i(1)$ désigne les revenus potentiels de i si elle va à l'université $Y_i(0)$ désigne les revenus potentiels de i si elle ne va pas à l'université

Les résultats potentiels sont définis par une **manipulation hypothétique**: que se passerait-il pour une personne particulière qui n'est pas allée à l'université si elle y était allée?

L'effet causal de l'université sur les revenus de la personne i est défini comme

$$\delta_i = Y_i(1) - Y_i(0)$$

Le problème fondamental de l'inférence causale

Dans le monde réel, une personne va à l'université ou elle n'y va pas

Cela signifie qu'un seul résultat potentiel sera jamais observé – l'autre est **contrefactuel** (**counterfactual**)

Le résultat observé, Y_i, est égal à $Y_i(0)$ si $D_i = 0$ et à $Y_i(1)$ si $D_i = 1$

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))D_i$$

Comme nous ne pouvons jamais observer à la fois $Y_i(0)$ et $Y_i(1)$, nous ne pouvons pas voir δ_i pour un individu. Ceci est connu comme le **problème fondamental de l'inférence causale**

Nous ne pouvons jamais espérer retrouver δ_i pour une personne individuelle, mais parfois nous pouvons retrouver certaines moyennes

Effets moyens du traitement (Average treatment effects)

L'**effet moyen du traitement** pour une population est défini comme

$$ATE = E[Y_i(1) - Y_i(0)]$$

- Le langage des "effets de traitement" est adopté des essais médicaux
- $Y_i(1)$ est le résultat de i si elle reçoit le "traitement" (université)
- $Y_i(0)$ est le résultat de i si elle reçoit le "contrôle" (pas d'université)
- $\delta_i = Y_i(1) - Y_i(0)$ est l'effet de traitement de i
- Autres paramètres d'intérêt :
  - **TOT** $= E[Y_i(1) - Y_i(0)|D_i = 1]$ (effet du traitement sur les traités)
  - **TNT** $= E[Y_i(1) - Y_i(0)|D_i = 0]$ (effet du traitement sur les non-traités)

Effets du traitement et biais de sélection

Considérons la comparaison des revenus moyens observés pour université vs. pas d'université :

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]$$

Ajoutons et soustrayons $E[Y_i(0)|D_i = 1]$:

$$\begin{aligned} E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_i(1) - Y_i(0)|D_i = 1] \\ + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \end{aligned}$$

Effets du traitement et biais de sélection

$$E[Y_i|D_i=1] - E[Y_i|D_i=0] = \underbrace{E[Y_i(1) - Y_i(0)|D_i=1]}_{\text{TOT}} + \underbrace{E[Y_i(0)|D_i=1] - E[Y_i(0)|D_i=0]}_{\text{Biais de sélection}}$$

Différence observée = TOT + biais de sélection

Le biais de sélection apparaît si les résultats de contrôle ne correspondent pas au contrefactuel manquant pour le groupe traité

Expériences randomisées

Supposons que le traitement soit attribué indépendamment des résultats potentiels : $(Y_i(1), Y_i(0)) \perp D_i$, comme dans un **essai contrôlé randomisé** (**randomized control trial**) (RCT)

Alors :

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_i(1)] - E[Y_i(0)] = ATE$$

L'attribution aléatoire garantit l'indépendance, éliminant le biais de sélection

Implique différence traitement/contrôle = ATE = TOT = TNT

Souvent le traitement n'est pas randomisé $\rightarrow$ besoin de conceptions **quasi-expérimentales**

Modèle du choix de durée de scolarité (Card, 1999)

Modèle simple pour répondre aux questions suivantes :

Combien d'années devriez-vous aller à l'école?

Quand devriez-vous arrêter d'aller à l'école et commencer à travailler?

L'individu i choisit la durée de scolarité S pour maximiser la valeur actualisée des revenus :

$$\int_S^\infty e^{-r_i t} Y_i(S) dt$$

$Y_i(S)$ est le revenu potentiel au niveau de scolarité S

Zéro revenu jusqu'à S, puis $Y_i(S)$ ensuite

Le taux d'actualisation r_i détermine comment les revenus futurs sont valorisés

- Représente à quel point les personnes sont patientes (ou impatientes)
- Pourrait aussi représenter le coût d'emprunt/l'accès au crédit

Choix de scolarité optimal

Scolarité optimale :

$$S^* = \arg\max \int_S^\infty e^{-r_i t} Y_i(S) dt$$

Condition du premier ordre :

$$\frac{Y_i'(S^*)}{Y_i(S^*)} = r_i$$

Bénéfice marginal vs. coût marginal :

- Investir les revenus $Y_i(S)$ et obtenir un rendement r_i
- Ou différer les revenus et obtenir un rendement proportionnel $Y_i'(S)/Y_i(S)$

La scolarité optimale égalise les deux

Biais de capacité

Les empiriques visent à estimer les caractéristiques de $Y_i(S)$

Problème : on observe seulement un $Y_i(S)$ par individu

Les choix de scolarité diffèrent en raison de r_i ou de $Y_i(S)$

Biais de capacité : les individus choisissant plus de scolarité peuvent avoir des revenus potentiels plus élevés

Rendements observés $\neq$ rendements causaux

Rendements observés de la scolarité

Régression MCO :

$$Y_i = \hat{a} + \hat{b} S_i + \hat{e}_i$$

Rendement observé de la scolarité = pente MCO :

$$\hat{b} = \frac{Cov(Y_i, S_i)}{Var(S_i)}$$

Question : est-ce que S_i est corrélé avec les résidus ajustés $\hat{e}_i$?

MCO approxime la fonction d'espérance conditionnelle

Question : est-ce que S_i est corrélé avec les résidus ajustés $\hat{e}_i$?

Non. Par construction, $Cov(e_i, S_i) = 0$

MCO donne l'approximation à erreur quadratique minimale de la fonction d'espérance conditionnelle (**conditional expectation function**) (CEF):

$$(a, b) = \arg\min_{a_0, b_0} E[(E[Y_i|S_i] - a_0 - b_0 S_i)^2]$$

MCO ajuste la CEF quel que soit le modèle

Meilleure question : est-ce que la CEF a un sens économique ?

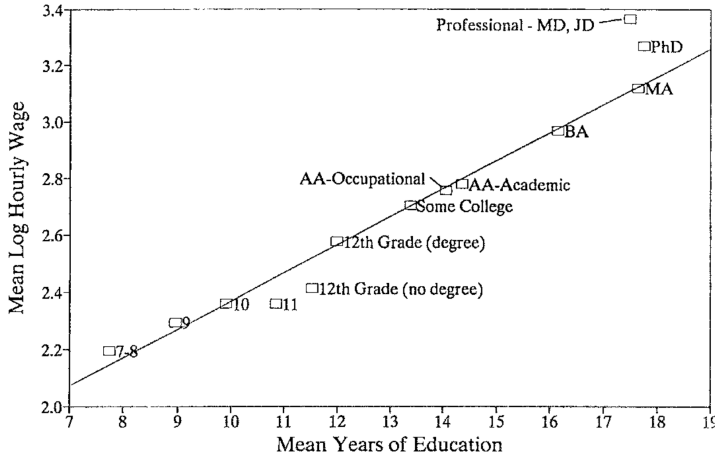


Fig. 2. Relationship between mean log hourly wages and completed education, men aged 40–45 in 1994–1996 Current Population Survey. Mean education by degree category estimated from February 1990 CPS.

Biais de capacité (Ability bias)

Considérons une fonction de revenus potentiels à **effets constants** :

$$Y_i(S) = \alpha_i + \beta S$$

Le rendement causal $\beta > 0$ est le même pour toutes les personnes et tous les niveaux de scolarité

Ce modèle implique

$$\frac{Y_i'(S^*)}{Y_i(S_i^*)} = r_i \quad \Rightarrow \quad S^* = \frac{1}{r_i} - \frac{\alpha_i}{\beta}$$

Supposons que le taux d'intérêt $r_i = r$ soit le même pour tout le monde. Le rendement observé de la scolarité est-il trop grand ou trop petit par rapport au rendement causal ?

Biais de capacité négatif

Le rendement observé est trop petit

Lorsque $r_i = r$ pour tout i, tout le monde gagne le même montant :

$$Y_i(S_i^*) = \alpha_i + \beta\left(\frac{1}{r} - \frac{\alpha_i}{\beta}\right) = \frac{\beta}{r}$$

Le rendement observé est donc nul, ce qui est inférieur au rendement causal β

Intuition : Le coût principal de la scolarité est le coût d'opportunité des revenus sacrifiés. Les personnes à plus grande capacité font face à des coûts d'opportunité plus élevés et abandonnent donc plus tôt

Dans ce cas le biais de capacité est négatif – le rendement causal dépasse le rendement observé. Est-ce réaliste ?


CAVALIERS
23

Biais de capacité général

Plus généralement, le rendement observé est

$$b = \frac{Cov(Y_i(S_i^*), S_i^*)}{Var(S_i)}$$
$$= \frac{Cov\left(\frac{\beta_i}{r_i}, \frac{1}{r_i} - \frac{\alpha_i}{\beta}\right)}{Var\left(\frac{1}{r_i} - \frac{\alpha}{\beta}\right)} = \beta \times \left(\frac{\sigma_{1/r}^2 - \sigma_{\alpha,1/r}\beta}{\sigma_{1/r}^2 - 2\sigma_{\alpha,1/r}\beta + \sigma_\alpha^2/\beta^2}\right)$$

Le biais de capacité dépend des variances et covariances des taux d'actualisation et de la capacité entre personnes. La direction est incertaine *a priori*

Pour obtenir un **biais de capacité positif**, il faut une force qui contrebalance l'histoire de coût d'opportunité de base

Estimer les rendements causaux

Le rendement observé de la scolarité peut être contaminé par un biais de capacité de signe et d'ampleur incertains. Comment peut-on estimer le rendement causal ?

Maintenons le modèle simple à effets constants : $Y_i(S) = \alpha_i + \beta S$

Revenus observés :

$$Y_i = \bar{\alpha} + \beta S_i + \varepsilon_i$$

où $\bar{\alpha} = E[\alpha_i]$ et $\varepsilon_i = \alpha_i - \bar{\alpha}$

Question : Dois-je m'inquiéter de savoir si S_i est corrélé avec ε_i ?

Rendements observés et causaux

Question : Dois-je m'inquiéter de savoir si S_i est corrélé avec ε_i ?

$$Y_i = \bar{\alpha} + \beta S_i + \varepsilon_i$$

Réponse : **Oui**. β est maintenant défini causalement, donc aucune garantie que $Cov(S_i, \varepsilon_i) = 0$

La scolarité n'est pas attribuée aléatoirement $\Rightarrow$ peut ne pas être indépendante des résultats potentiels, résumés par ε_i

Ainsi la pente MCO b peut ne pas être égale à l'effet causal β

Variables instrumentales (instrumental variables) (IV)

Les **variables instrumentales** (IV) sont une conception courante pour éliminer le biais de sélection

$Y_i = \bar{\alpha} + \beta S_i + \varepsilon_i$

Supposons que nous ayons un instrument Z_i qui satisfait deux conditions :

- **Première étape** (**First stage**) : $Cov(S_i, Z_i) \neq 0$
- **Restriction d'exclusion** (**Exclusion restriction**): $Cov(\varepsilon_i, Z_i) = 0$

La première étape exige que Z_i soit corrélé avec S_i

L'exclusion exige que Z_i ne soit pas corrélé avec les résultats potentiels (ε_i)

Z_i doit être **aussi bon qu'aléatoirement assigné** (**as good as randomly assigned**), et ne peut pas affecter Y_i sauf via S_i

Le coefficient IV de population

Covariance entre le résultat et l'instrument :

$$Cov(Y_i, Z_i) = Cov(\bar{\alpha} + \beta S_i + \varepsilon_i, Z_i)$$

$$= \beta Cov(S_i, Z_i) + Cov(\varepsilon_i, Z_i)$$

Restriction d'exclusion $\Rightarrow Cov(\varepsilon_i, Z_i) = 0$

Donc,

$$\beta = \frac{Cov(Y_i, Z_i)}{Cov(S_i, Z_i)}$$

Ce rapport de covariances est le **coefficient IV de population**

Interprétation IV

Divisons numérateur et dénominateur par $Var(Z_i)$:

$$\beta^{IV} = \frac{Cov(Y_i, Z_i)/Var(Z_i)}{Cov(S_i, Z_i)/Var(Z_i)}$$

Le coefficient IV est un rapport de deux régressions :

- Forme réduite : régression de Y_i sur Z_i
- Première étape : régression de S_i sur Z_i

Si Z_i est binaire

$$\beta^{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[S_i|Z_i = 1] - E[S_i|Z_i = 0]}$$

connu sous le nom d'**estimateur de Wald**

Estimations IV du rendement de la scolarité: Angrist et Krueger (1991)

Angrist et Krueger (QJE 1991): étude IV classique sur les rendements de l'éducation

Stratégie: interaction entre **lois sur la scolarité obligatoire** et **lois sur l'âge d'entrée**

- Les étudiants peuvent quitter l'école le jour de leurs 16 ans
- La date de rentrée est liée à l'année civile des six ans
- Crée des différences de scolarité selon la date de naissance

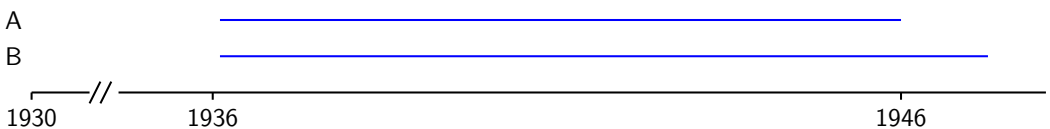
Exemple

Étudiant A: Né le 1 janvier 1930

- Commence l'école le 1 septembre 1936 (âge 6)
- A 16 ans le 1 janvier 1946, pendant la 10e année
- Peut quitter le 2 janv 1946 $\rightarrow$ complète environ 9.5 années de scolarité

Étudiant B: Né le 31 décembre 1930

- Commence l'école le 1 septembre 1936 (âge 5.75)
- A 16 ans le 31 décembre 1946, pendant la 11e année
- Peut quitter le 31 déc 1946 $\rightarrow$ complète environ 10.5 années de scolarité



Instruments par trimestre de naissance

L'instrument AK91: **trimestre de naissance (quarter of birth) (QOB)** d'après les données du recensement

$Z_i = 1\{\text{né au premier trimestre}\}$

Que signifient la première étape et la restriction d'exclusion pour le QOB?

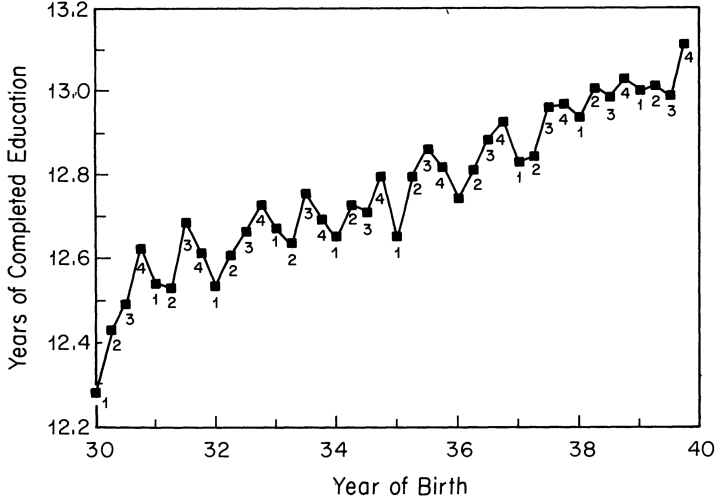


FIGURE I
Years of Education and Season of Birth
1980 Census
Note. Quarter of birth is listed below each observation.

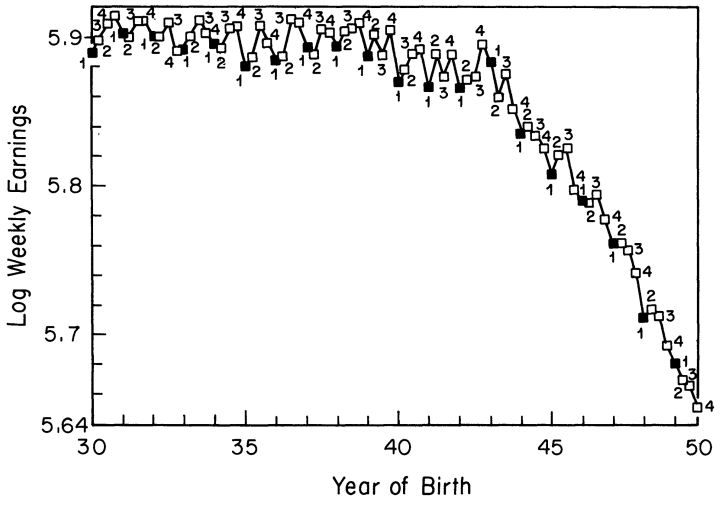


FIGURE V
Mean Log Weekly Wage, by Quarter of Birth
All Men Born 1930–1949; 1980 Census

TABLE III

PANEL A: WALD ESTIMATES FOR 1970 CENSUS—MEN BORN 1920–1929[a]

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year	(3) Difference (std. error) (1) − (2)
ln (wkly. wage)	5.1484	5.1574	−0.00898
			(0.00301)
Education	11.3996	11.5252	−0.1256
			(0.0155)
Wald est. of return to education			0.0715
			(0.0219)
OLS return to education[b]			0.0801
			(0.0004)

Panel B: Wald Estimates for 1980 Census—Men Born 1930–1939

	(1) Born in 1st quarter of year	(2) Born in 2nd, 3rd, or 4th quarter of year	(3) Difference (std. error) (1) − (2)
ln (wkly. wage)	5.8916	5.9027	−0.01110
			(0.00274)
Education	12.6881	12.7969	−0.1088
			(0.0132)
Wald est. of return to education			0.1020
			(0.0239)
OLS return to education			0.0709
			(0.0003)

Interprétation du QOB

Les estimations IV utilisant le QOB suggèrent un rendement de la scolarité de 7–10% par an

Comparable ou supérieur aux estimations MCO

Dans notre modèle cela suggère un **biais d'aptitude négatif**: les personnes moins capables vont plus longtemps à l'école

Autres interprétations?

Effets de traitement hétérogènes (heterogeneous treatment effects)

Notre modèle simple supposait des effets constants de la scolarité entre les personnes

Retour au modèle général des résultats potentiels avec traitement binaire D_i et résultats potentiels $Y_i(1)$ et $Y_i(0)$

Supposons que nous ayons un instrument binaire Z_i, et considérons deux nouveaux résultats potentiels définis par une manipulation hypothétique de Z_i:

$D_i(1)$: statut de traitement de i si $Z_i = 1$
$D_i(0)$: statut de traitement de i si $Z_i = 0$

Le traitement observé est $D_i = D_i(0) + (D_i(1) - D_i(0))Z_i$

Hypothèses IV

Hypothèses IV dans un monde à effets de traitement hétérogènes:

- Exclusion: $(Y_i(1), Y_i(0), D_i(1), D_i(0)) \perp Z_i$
- Première étape: $\Pr[D_i = 1|Z_i = 1] > \Pr[D_i = 1|Z_i = 0]$
- **Monotonicité:** $D_i(1) \geq D_i(0)\ \forall i$

Par rapport à notre configuration IV à effets constants, la monotonicité est l'hypothèse nouvelle

La monotonicité exige que l'instrument affecte le statut de traitement de chacun dans la même direction

Groupes de conformité

Sous monotonicité, nous pouvons partitionner la population en trois groupes définis par leurs réponses comportementales à l'instrument (Angrist, Imbens, et Rubin 1996):

1. **Toujours preneurs** (**always-takers**): $D_i(1) = D_i(0) = 1$
2. **Jamais preneurs** (**never-takers**): $D_i(1) = D_i(0) = 0$
3. **Compliers:** $D_i(1) = 1,\ D_i(0) = 0$

Les compliers ont $D_i(1) > D_i(0)$: leur statut de traitement augmente avec l'instrument

La monotonicité exclut les **défiants** avec $D_i(1) = 0,\ D_i(0) = 1$

Effet de traitement moyen local (LATE)

Sous ces hypothèses, IV identifie un **effet de traitement moyen local (LATE)**:

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0)]$$

C'est le **théorème LATE d'Imbens et Angrist (1994)**

Le LATE est l'**effet de traitement moyen pour les compliers** – individus dont le statut de traitement est déterminé par l'instrument

Preuve du théorème LATE

Notons que $Y_i = Y_i(D_i) = Y_i(D_i(Z_i))$, donc par indépendance

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = E[Y_i(D_i(1))|Z_i = 1] - E[Y_i(D_i(0))|Z_i = 0]$$
$$= E[Y_i(D_i(1)) - Y_i(D_i(0))]$$

Par monotonicité nous avons soit $D_i(1) = D_i(0)$ soit $D_i(1) > D_i(0)$, donc

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0)]\,\Pr[D_i(1) > D_i(0)]$$

La même logique implique

$$E[D_i|Z_i = 1] - E[D_i|Z_i = 0] = \Pr[D_i(1) > D_i(0)]$$

Donc

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0)]$$

□

Interprétation des estimations IV

Comment interpréter les effets identifiés par l'instrument QOB dans un cadre de traitements hétérogènes ?

Interprétation des estimations IV

Comment interpréter les effets identifiés par l'instrument QOB dans un cadre de traitements hétérogènes ?

Qui sont les complieurs pour l'instrument QOB ?

Interprétation des estimations IV

Comment interpréter les effets identifiés par l'instrument QOB dans un cadre de traitements hétérogènes ?

Qui sont les complieurs pour l'instrument QOB ?

L'interprétation LATE suggère que l'instrument QOB identifie l'effet causal de la scolarité supplémentaire pour les individus à la marge d'abandonner tôt au milieu du siècle

Ensuite, nous considérerons des preuves plus récentes portant sur d'autres marges de scolarité

Rendements de l'université pour les étudiants marginaux : Zimmerman (2014)

Le rendement observé de l'université a augmenté de façon spectaculaire ces dernières décennies

La prime salariale universitaire est passée de 50% à 97% entre 1980 et 2008 (Acemoglu et Autor, 2011)

Peut refléter une forte croissance de la demande de compétences couplée à une croissance lente de l'offre de compétences (Goldin et Katz, 2008)

En même temps, de nombreux étudiants aux États-Unis commencent l'université mais ne terminent pas

Question : La fréquentation de l'université améliore-t-elle les revenus des étudiants académiquement marginaux ?

Zimmerman (JOLE 2014) exploite un **design de discontinuité de régression** (RDD)

Régression sur discontinuité (Regression discontinuity)

Considérons un cadre avec un traitement binaire $D_i \in \{0, 1\}$, et des résultats potentiels $Y_i(1)$ et $Y_i(0)$

Supposons que le traitement soit une fonction déterministe et discontinue d'une covariable observée R_i, telle que

$$D_i = 1\{R_i > c\}$$

R_i est appelée la **variable de contrôle** ou **variable de forçage**

Ceci est une "**sharp RD**" parce que la probabilité de traitement passe de zéro à un au seuil

Zimmerman (2014) : seuil de GPA pour l'admission dans les universités publiques en Floride

Régression sur discontinuité

Nous observons $Y_i(1)$ lorsque $R_i > c$ et $Y_i(0)$ lorsque $R_i \leq c$

Idée de base du design RD : comparer les observations juste au-dessus et juste en dessous du seuil pour inférer l'effet du traitement

Intuitivement, le traitement peut être aussi bon que tiré au hasard pour les individus dans le "voisinage" de $R_i = c$, donc comparer traités et non traités près de c révèle un effet de traitement

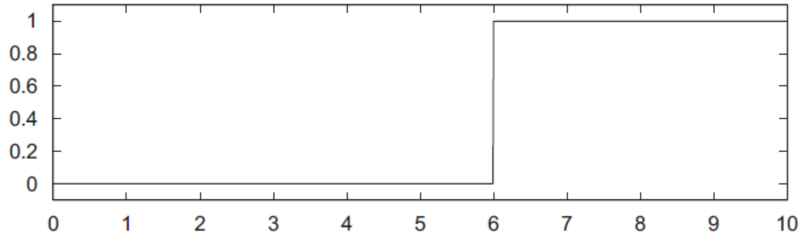


Fig. 1. Assignment probabilities (SRD).

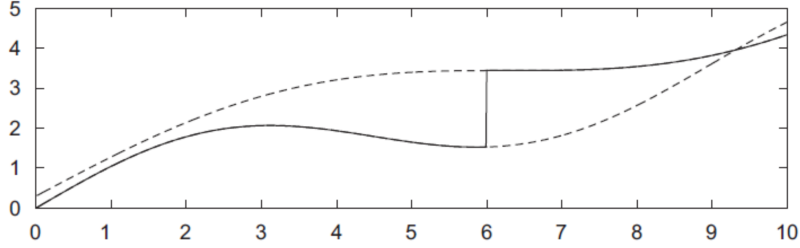


Fig. 2. Potential and observed outcome regression functions.

Source: Imbens and Lemieux (2008)

Identification RD

Hypothèse clé : les résultats potentiels sont lisses au seuil

Formellement :

$$\lim_{r \to c^+} E[Y_i(d)|R_i = r] = \lim_{r \to c^-} E[Y_i(d)|R_i = r], \quad d \in \{0, 1\}$$

Les CEF des résultats potentiels doivent être continues au seuil

La population juste en dessous ne doit pas être discrètement différente de la population juste au-dessus

Identification RD

Si cette hypothèse tient, nous avons

$$\lim_{r \to c^+} E[Y_i|R_i = r] - \lim_{r \to c^-} E[Y_i|R_i = r] = E[Y_i(1) - Y_i(0)|R_i = c]$$

Lorsque les résultats potentiels sont lisses autour du seuil, une comparaison des individus juste au-dessus et juste en dessous donne l'effet moyen du traitement pour ceux au seuil

L'argument d'identification est **non paramétrique** : nous n'avons pas besoin de supposer quoi que ce soit sur la distribution des résultats potentiels, à part la continuité des CEF

Interprétation RD

Intuition centrale de RD : pour ceux proches du seuil, les choses auraient pu aller dans un sens ou dans l'autre

Interprétez RD comme un **essai randomisé local** parmi ceux proches de $R_i = c$

Explique pourquoi les preuves RD peuvent être particulièrement convaincantes par rapport à d'autres stratégie empique – proches de l'idéal RCT

La vision de "randomisation locale" motive des diagnostics RD communs :
- Vérifier l'équilibre des caractéristiques prédéterminées pour les observations au-dessus et en dessous du seuil - Rechercher des anomalies dans la distribution de la variable de contrôle autour du seuil (McCrary, 2008)

"Fuzzy" RD

Parfois le traitement est généré par une règle d'attribution discontinue qui n'est pas déterministe

Supposons que

$$\lim_{r \to c^+} \Pr[D_i = 1 | R_i = r] > \lim_{r \to c^-} \Pr[D_i = 1 | R_i = r]$$

La probabilité de traitement saute à $R_i = c$, mais pas nécessairement de zéro à un

Ceci est un scénario de **fuzzy RD** car le traitement est seulement partiellement déterminé par le seuil

Zimmerman (2014) : Les étudiants au-dessus du seuil de GPA sont éligibles à l'admission, mais non garantis

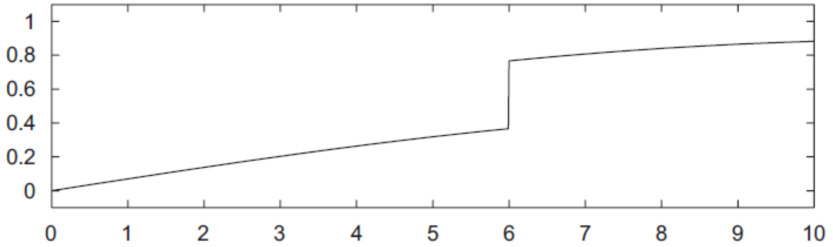


Fig. 3. Assignment probabilities (FRD).

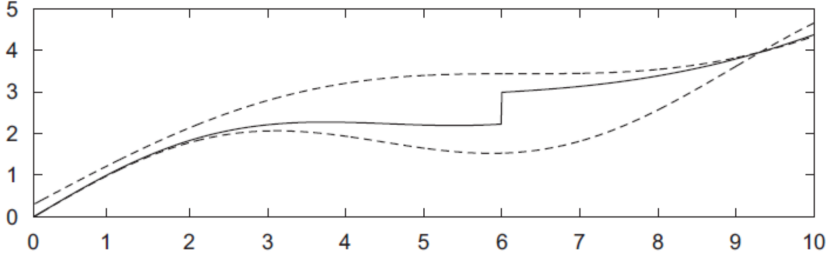


Fig. 4. Potential and observed outcome regression (FRD).

Hypothèses de fuzzy RD

Comme auparavant, supposons que les distributions de $Y_i(1)$ et $Y_i(0)$ soient lisses autour du seuil

Soient $D_i(1)$ et $D_i(0)$ les statuts de traitement potentiels pour l'individu i s'il est au-dessus ou en dessous du seuil. Supposons qu'ils soient également lisses, et que $D_i(1) \geq D_i(0)$ pour tout i

Franchir le seuil augmente faiblement la probabilité de traitement pour tout le monde

Fuzzy RD

Sous ces hypothèses :

$$\frac{\lim_{r \to c^+} E[Y_i|R_i = r] - \lim_{r \to c^-} E[Y_i|R_i = r]}{\lim_{r \to c^+} E[D_i|R_i = r] - \lim_{r \to c^-} E[D_i|R_i = r]} = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0), R_i = c]$$

- Numérateur : saut des résultats au seuil (comme dans RD franche)
- Dénominateur : variation de la probabilité de traitement au seuil
- Le ratio identifie un effet moyen du traitement pour les **switchers au seuil**

Ça vous semble familier ?

Fuzzy RD est IV

Fuzzy RD est IV avec l'indicateur de seuil $Z_i = 1\{R_i > c\}$ comme instrument

Pensez à FRD comme à un essai randomisé local avec **non-conformité**

Implique que les estimations RD floues sont locales en deux sens :

- Locales au seuil, $R_i = c$ (vrai aussi pour RD franche)
- Locales aux complieurs au seuil (le "L" dans LATE)

Zimmerman (2014) : rendements de l'université pour les étudiants marginaux

Seuil de GPA $\rightarrow$ admission au système universitaire public de Floride

Concerné par Florida International University (FIU)

Population proche du seuil : SAT faible (21e percentile) et faibles taux de diplomation

Les estimations capturent les rendements pour les étudiants marginaux

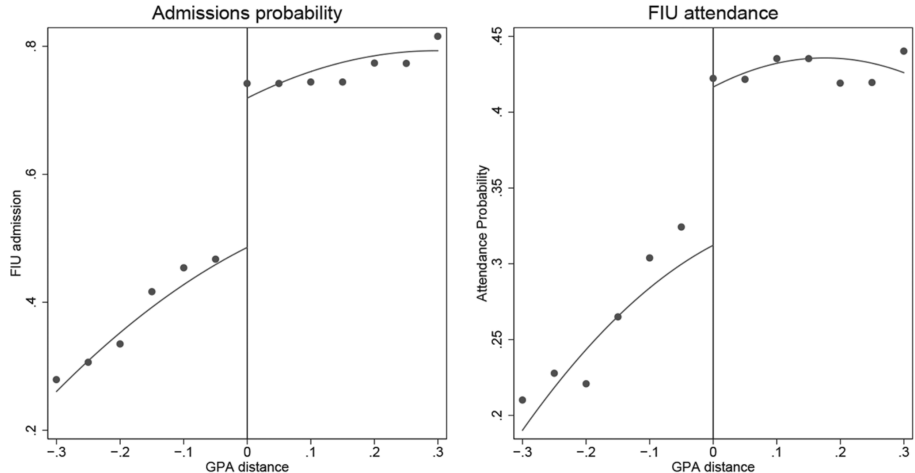


FIG. 4.—Admissions and FIU attendance. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

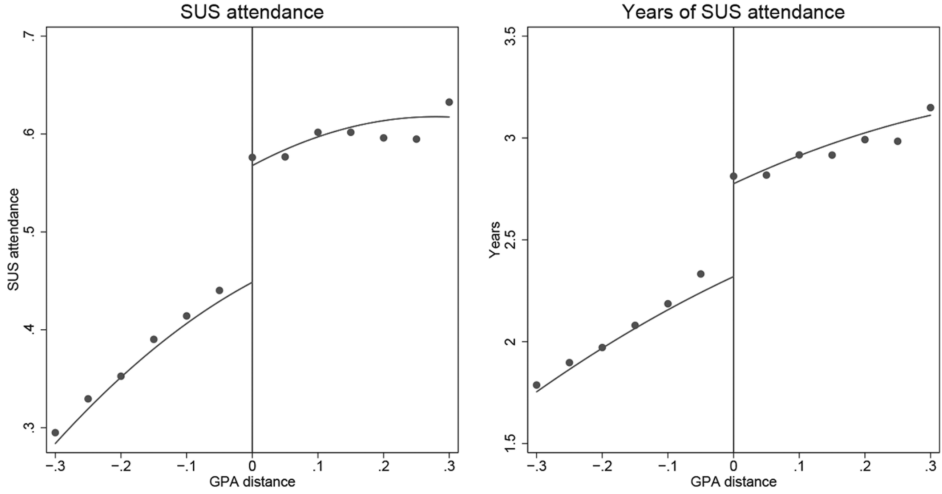


FIG. 5.—SUS attendance and persistence. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

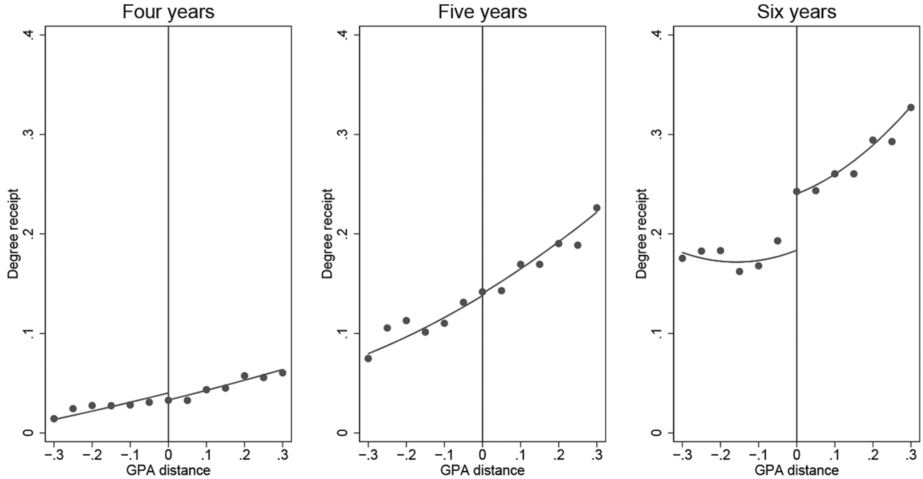


FIG. 6.—SUS BA receipt by years elapsed since high school. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

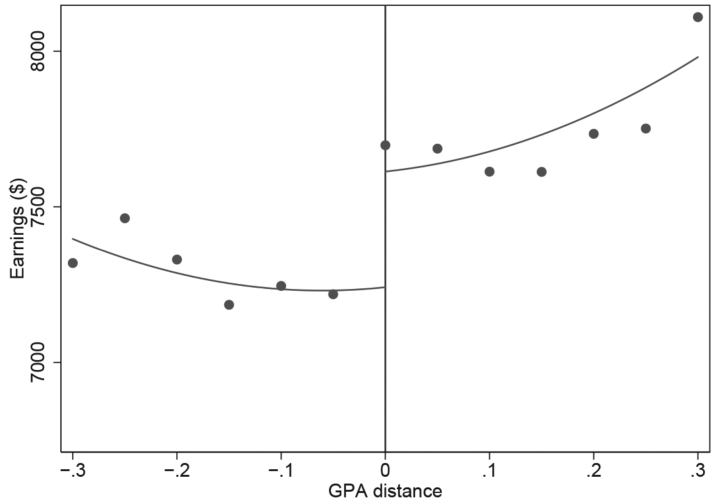


FIG. 8.—Quarterly earnings by distance from GPA cutoff. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

Table 5
Earnings Effects 8–14 Years after High School Completion

	Main	Controls	BW=.5	BW=.15	Local Linear
Reduced-form estimates:					
Above cutoff	372*	366**	409**	479**	410**
	(141)	(130)	(154)	(198)	(147)
Instrumental variables estimates:					
FIU admission	1,593*	1,575**	1,665**	1,700**	2,001*
	(604)	(584)	(645)	(621)	(696)
Years of SUS attendance	815**	792**	833**	966***	977**
	(276)	(262)	(271)	(305)	(306)
BA degree	6,547*	6,442*	7,366*	10,769	5,958**
	(2,496)	(2,411)	(2,998)	(5,726)	(2,024)
N	6,542	6,542	9,659	3,294	6,542

NOTE.—FIU = Florida International University; SUS = State University System; BA = bachelor's degree. Standard errors are clustered within grade bins. The *p*-values are calculated using a clustered wild bootstrap-*t* procedure described in Sec. III and app. B. The dependent variable in each regression is average quarterly earnings in 2005 dollars. The "BW=.15" specification uses observations within .15 grade points above and below the cutoff and allows for a linear trend in distance from the cutoff. The "BW=.5" specification uses observations within the .5 grade points on either side of the cutoff and allows for a quartic polynomial in distance from the cutoff. The "Local Linear" specification is identical to the main specification, but it allows for linear slope terms in distance from the cutoff that differ above and below the threshold.

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Dale et Krueger

- Comparer des étudiants qui ont postulé/été admis aux mêmes écoles mais qui en ont fréquenté différentes
- Les choix de candidature révèlent des informations sur la capacité des étudiants
- Les choix d'admission révèlent des informations sur la capacité des collèges
- Données : enquête College and Beyond (34 écoles sélectives)
- Mise à jour de 2014 liée aux revenus de la SSA

TABLE I

ILLUSTRATION OF HOW MATCHED-APPLICANT GROUPS WERE CONSTRUCTED

		Student applications to college							
		Application 1		Application 2		Application 3		Application 4	
Student	Matched-applicant group	School average SAT	School admissions decision	School average SAT	School admissions decision	School average SAT	School admissions decision	School average SAT	School admissions decision
Student A	1	1280	Reject	1226	Accept*	1215	Accept	na	na
Student B	1	1280	Reject	1226	Accept	1215	Accept*	na	na
Student C	2	1360	Accept	1310	Reject	1270	Accept*	1155	Accept
Student D	2	1355	Accept	1316	Reject	1270	Accept*	1160	Accept
Student E	2	1370	Accept*	1316	Reject	1260	Accept	1150	Accept
Student F	Excluded	1180	Accept*	na	na	na	na	na	na
Student G	Excluded	1180	Accept*	na	na	na	na	na	na
Student H	3	1360	Accept	1308	Accept*	1260	Accept	1160	Accept
Student I	3	1370	Accept*	1311	Accept	1255	Accept	1155	Accept
Student J	3	1350	Accept	1316	Accept*	1265	Accept	1155	Accept
Student K	4	1245	Reject	1217	Reject	1180	Accept*	na	na
Student L	4	1235	Reject	1209	Reject	1180	Accept*	na	na
Student M	5	1140	Accept	1055	Accept*	na	na	na	na
Student N	5	1145	Accept*	1060	Accept	na	na	na	na
Student O	No match	1370	Reject	1038	Accept*	na	na	na	na

* Denotes school attended.

na = did not report submitting application.

The data shown on this table represent hypothetical students. Students F and G would be excluded from the matched-applicant subsample because they applied to only one school (the school they attended). Student O would be excluded because no other student applied to an equivalent set of institutions.

TABLE V

LINEAR REGRESSIONS PREDICTING WHETHER STUDENT ATTENDED MOST SELECTIVE COLLEGE FOR C&B SAMPLE OF STUDENTS ADMITTED TO MORE THAN ONE SCHOOL

	Parameter estimates	
	Matched-applicant model*	Self-revelation model
Predicted log (parental income)	−0.024	−0.037
	(0.026)	(0.030)
Own SAT score/100	0.020	0.021
	(0.005)	(0.007)
Female	0.034	0.033
	(0.014)	(0.028)
Black	0.056	−0.005
	(0.026)	(0.037)
Hispanic	−0.019	0.042
	(0.064)	(0.074)
Asian	0.019	0.074
	(0.026)	(0.050)
Other/missing race	−0.095	0.010
	(0.093)	(0.081)
High school top 10 percent	−0.014	−0.020
	(0.021)	(0.028)
High school rank missing	−0.035	−0.040
	(0.036)	(0.058)
Athlete	0.056	0.059
	(0.023)	(0.045)
Average SAT score/100 of schools applied to		−0.122
		(0.040)
One additional application		0.149
		(0.037)
Two additional applications		0.076
		(0.033)
Three additional applications		0.020
		(0.038)
N	5536	8257

TABLE III

LOG EARNINGS REGRESSIONS USING COLLEGE AND BEYOND SURVEY, SAMPLE OF MALE AND FEMALE FULL-TIME WORKERS

	Model					
	Basic model: no selection controls		Matched-applicant model	Alternative matched-applicant models		Self-revelation model
	Full sample	Restricted sample	Similar school-SAT matches*	Exact school-SAT matches**	*Barron's* matches***	
Variable	1	2	3	4	5	6
School-average SAT score/100	0.076	0.082	−0.016	−0.106	0.004	−0.001
	(0.016)	(0.014)	(0.022)	(0.036)	(0.016)	(0.018)
Predicted log(parental income)	0.187	0.190	0.163	0.232	0.154	0.161
	(0.024)	(0.033)	(0.033)	(0.079)	(0.028)	(0.025)
Own SAT score/100	0.018	0.006	−0.011	0.003	−0.005	0.009
	(0.006)	(0.007)	(0.007)	(0.014)	(0.005)	(0.006)
Female	−0.403	−0.410	−0.395	−0.476	−0.400	−0.396
	(0.015)	(0.018)	(0.024)	(0.049)	(0.017)	(0.014)
Black	−0.023	−0.026	−0.057	−0.028	−0.057	−0.034
	(0.035)	(0.053)	(0.053)	(0.049)	(0.039)	(0.035)
Hispanic	0.015	0.070	0.020	−0.248	0.036	0.007
	(0.052)	(0.076)	(0.099)	(0.206)	(0.066)	(0.053)
Asian	0.173	0.245	0.241	0.368	0.163	0.155
	(0.036)	(0.054)	(0.064)	(0.141)	(0.049)	(0.037)
Other/missing race	−0.188	−0.048	0.060	−0.072	−0.050	−0.192
	(0.119)	(0.143)	(0.180)	(0.083)	(0.134)	(0.116)
High school top 10 percent	0.061	0.091	0.079	0.091	0.079	0.063
	(0.018)	(0.022)	(0.026)	(0.032)	(0.024)	(0.019)
High school rank missing	0.001	0.040	0.016	0.029	0.025	−0.009
	(0.024)	(0.026)	(0.038)	(0.066)	(0.027)	(0.022)
Athlete	0.102	0.088	0.104	0.169	0.093	0.094
	(0.025)	(0.030)	(0.039)	(0.096)	(0.033)	(0.024)
Average SAT score/100 of schools applied to						0.090
						(0.013)
One additional application						0.064
						(0.011)
Two additional applications						0.074
						(0.022)
Three additional applications						0.112
						(0.028)
Four additional applications						0.085
						(0.027)
Adjusted R^2	0.107	0.110	0.112	0.142	0.106	0.113
N	14,238	6,335	6,335	2,330	9,202	14,238

Table 3

Comparing Parameter Estimates of the Effect of College Average SAT Score on Earnings Using C&B and SSA Data, 1976 Cohort

	C&B sample[a]		Merged C&B and SSA sample[b]									
	Log 1995 C&B earnings		Log 1995 C&B earnings		Log 1995 SSA earnings (topcoded)		Log 1995 SSA earnings (not topcoded)		Log (median of 1993 to 1997 earnings), SSA data		Log (median of 1993 to 1997 earnings), SSA data	
	1	2	3	4	5	6	7	8	9	10	11	12
	Basic	Self – revelation	Basic	Self - revelation	Basic	Self – revelation	Basic	Self - revelation	Basic	Self – revelation	Basic	Self – revelation
Parameter estimate for school	0.076	-0.001	0.068	-0.007	0.048	-0.021	0.058	-0.015	0.059	-0.025	0.061	-0.023
	(.008)	(.012)	(.007)	(.012)	(.009)	(.014)	(.009)	(.015)	(.008)	(.012)	(.007)	(.012)
SAT/100	{.016}	{.018}	{.014}	{.018}	{.016}	{.018}	{.017}	{.016}	{.012}	{.013}	{.013}	{.014}
N	14,238		10,886		10,886		10,886		11,932		12,075	
Sample restriction	Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Median earnings greater than zero (SSA data)		Median earnings greater than $13,822 in 2007 dollars (SSA data)	

Table 8

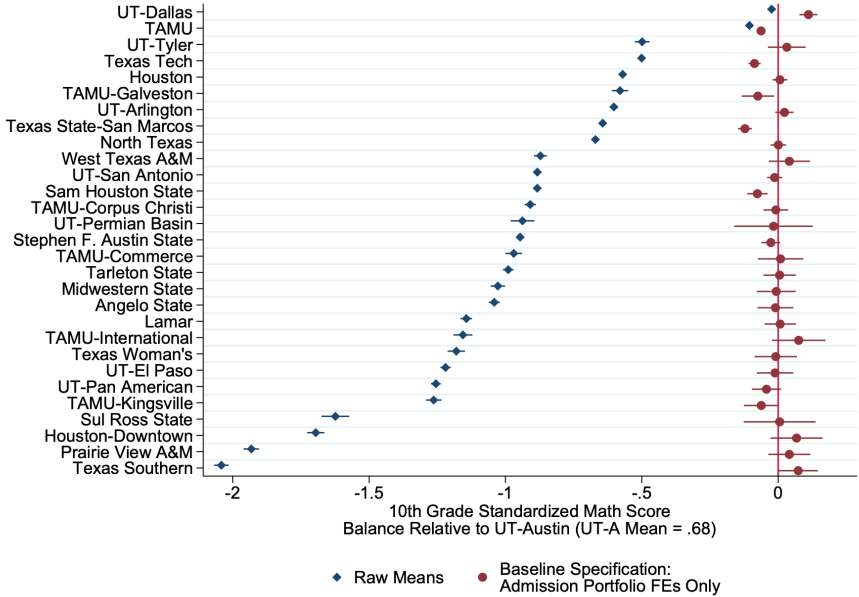
Effect of School Characteristics on 2007 Earnings (Black and Hispanic Students Only, 1989 Cohort)

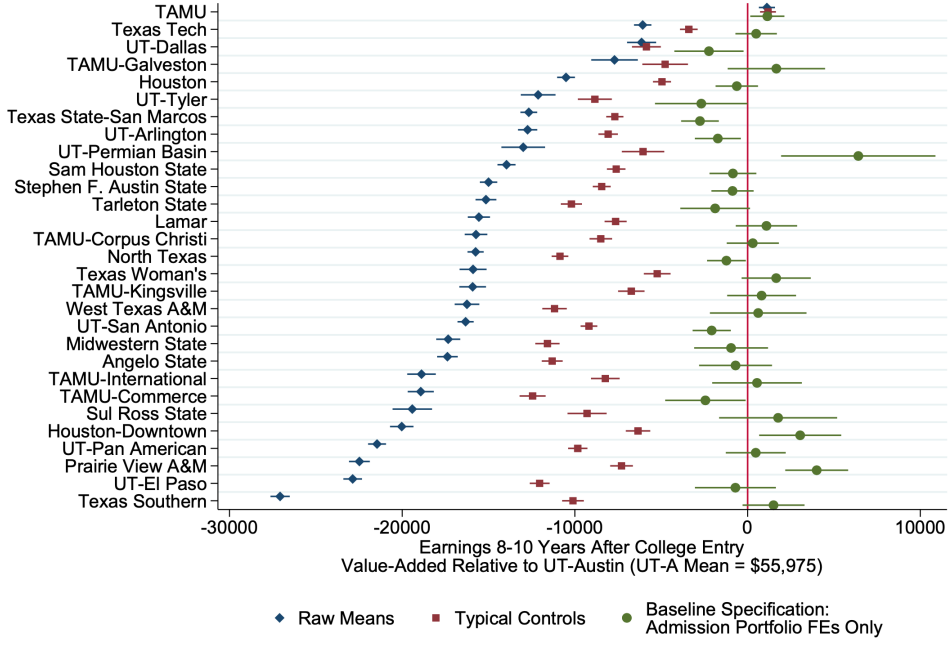
	School SAT score/100		Log net tuition		Barron's index	
Dependent variable	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation
All black and Hispanic students						
Parameter estimate for	0.067	0.076	0.173	0.138	0.063	0.049
effect of quality measure	(.019)	(.032)	(.056)	(.071)	(.022)	(.036)
on log 2007 earnings	{.028}	{.042}	{.076}	{ .092}	{.033}	{.046}
Sample size	1,508		1,508		1,508	
All black and Hispanic students, excluding historically black colleges and universities						
Parameter estimate for	0.122	0.120	0.187	0.116	0.158	0.143
effect of quality measure	(.030)	(.042)	(.064)	(.079)	(.040)	(.053)
on log 2007 earnings	{.035}	{.056}	{.081}	{.101}	{.038}	{.051}
Sample size	995		995		995	

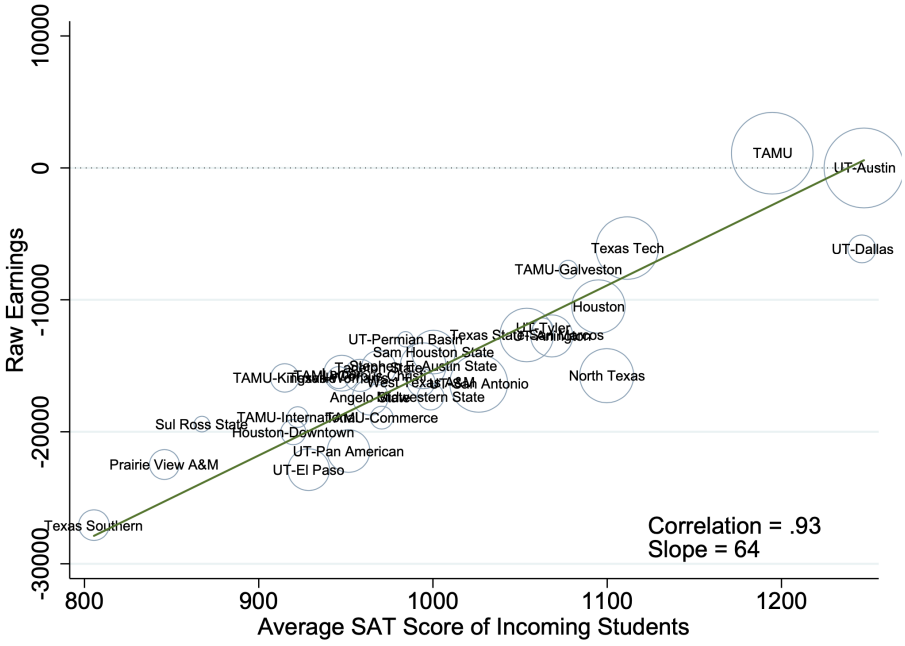
Mountjoy et Hickman (2020)

- Mettre à jour la stratégie DK en utilisant des données administratives du Texas
- Estimer un modèle de valeur ajoutée : effet pour chaque collège
- Conditionner sur les contrôles de candidature/admission DK
- Trouver des rendements limités à la sélectivité
- La valeur ajoutée est corrélée avec les intrants (dépenses, ratio enseignants)

Figure 3: Validating the Matched Applicant Approach: Ability Balance across College Treatments




TAMU
Texas Tech
UT-Dallas
TAMU-Galveston
Houston
UT-Tyler
Texas State-San Marcos
UT-Arlington
UT-Permian Basin
Sam Houston State
Stephen F. Austin State
Tarleton State
Lamar
TAMU-Corpus Christi
North Texas
Texas Woman's
TAMU-Kingsville
West Texas A&M
UT-San Antonio
Midwestern State
Angelo State
TAMU-International
TAMU-Commerce
Sul Ross State
Houston-Downtown
UT-Pan American
Prairie View A&M
UT-El Paso
Texas Southern
-30000
-20000
-10000
0
10000
Earnings 8-10 Years After College Entry
Value-Added Relative to UT-Austin (UT-A Mean = $55,975)
Raw Means
Typical Controls
Baseline Specification:
Admission Portfolio FEs Only


Raw Earnings
10000
0
-10000
-20000
-30000
800
900
1000
1100
1200
Average SAT Score of Incoming Students
TAMU
UT-Austin
Texas Tech
UT-Dallas
TAMU-Galveston
Houston
UT-Permian Basin
Sam Houston State
Texas State
UT-Tyler
Stephen F. Austin State
TAMU-Kingsville
West Texas A&M
UT-San Antonio
North Texas
Angelo State
Midwestern State
TAMU-International
TAMU-Commerce
Sul Ross State
Houston-Downtown
UT-Pan American
Prairie View A&M
UT-El Paso
Texas Southern
Correlation = .93
Slope = 64

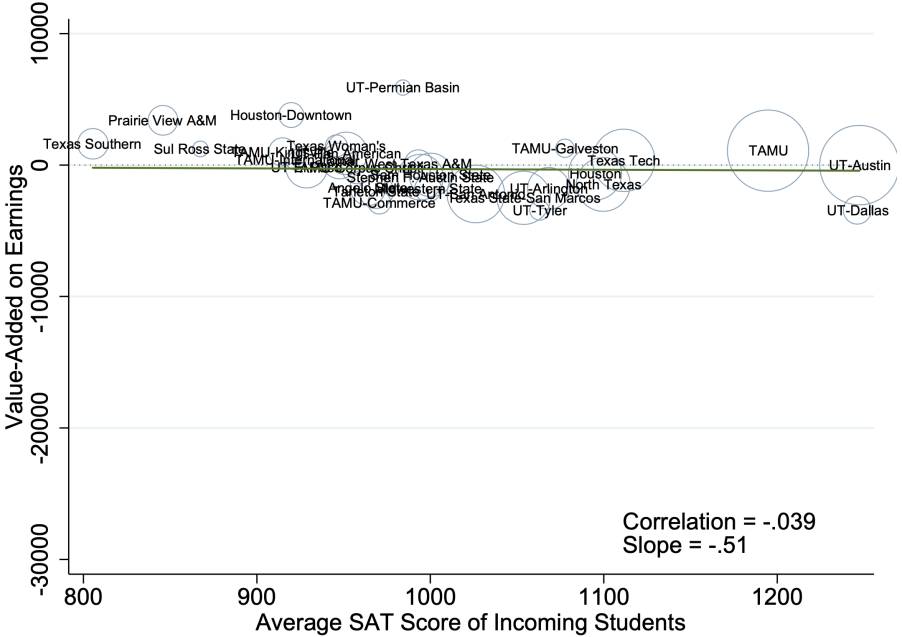
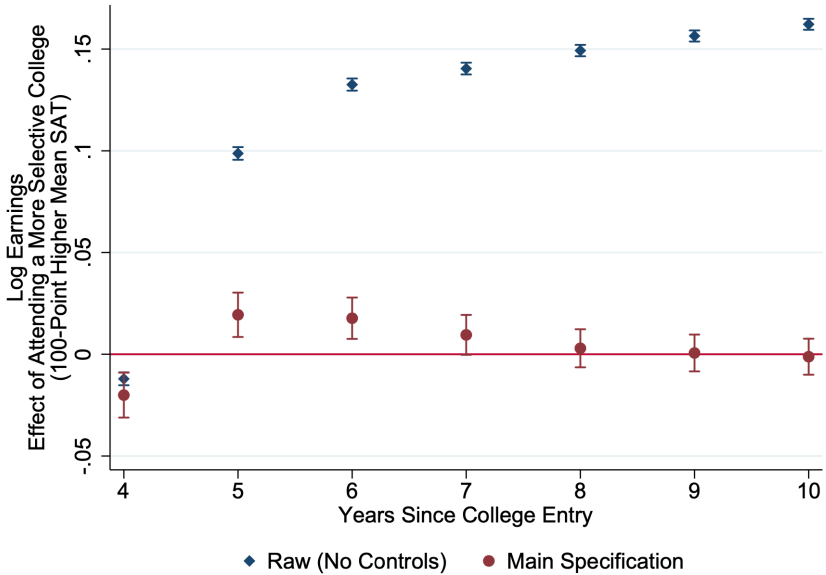

Value-Added on Earnings
10000
0
-10000
-20000
-30000
800
900
1000
1100
1200
Average SAT Score of Incoming Students
UT-Permian Basin
Prairie View A&M
Houston-Downtown
Texas Southern
Sul Ross State
Texas Woman's
TAMU-Galveston
TAMU
Texas Tech
UT-Austin
Houston
North Texas
UT-Arlington
UT-Tyler
UT-Dallas
TAMU-Commerce
Correlation = -.039
Slope = -.51

Figure 7: Early Career Dynamics of the Return to College Selectivity




Value-Added on Earnings
6000
4000
2000
0
-2000
-4000
-2
-1
0
1
2
Non-Peer Inputs Index,
Residualized on Selectivity (Average SAT Score)
Correlation = .67
Slope = 1,438

Spence (1973) signalisation

Les travailleurs diffèrent en capacité $a \in \{H, L\}$ avec $\Pr(H) = p$

L'éducation $e \in \{0, 1\}$ ne fournit **aucune valeur de productivité**, mais est plus coûteuse pour les travailleurs de faible capacité : $c_H < c_L$

Les entreprises observent le niveau d'éducation et paient un salaire égal à la capacité attendue conditionnellement à e

Si les travailleurs à haute productivité choisissent l'éducation et les travailleurs à faible productivité choisissent l'absence d'éducation

$$H - c_H \geq L \quad L \geq H - c_L$$

alors les entreprises infèrent parfaitement le type : $w(1) = H$, $w(0) = L$

Ceci est appelé un **équilibre séparateur**

L'éducation est purement un **signal** : elle n'augmente pas la productivité, mais elle distingue de manière crédible les types élevés des types faibles

Discussion : Arteaga (2018)

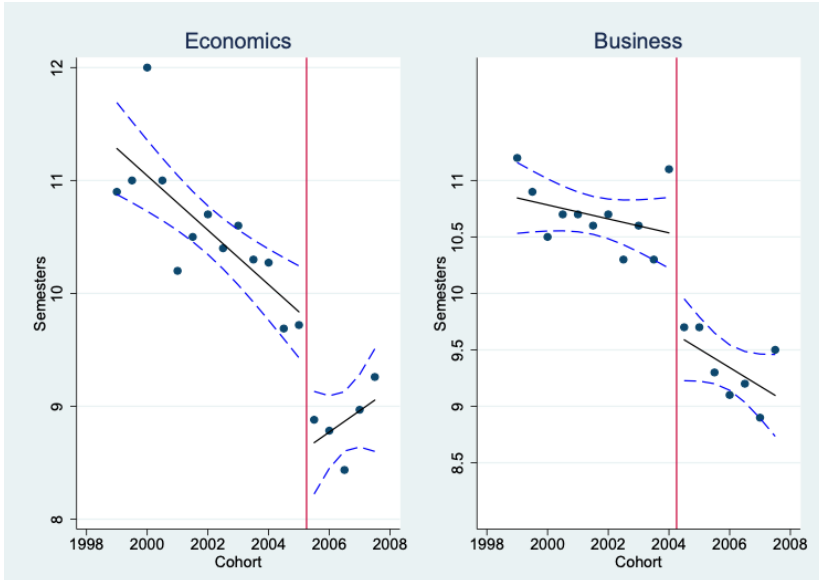
Quelle est la question de recherche ?

Quelles données utilise-t-elle ?

Quel changement de politique l'autrice exploite-t-elle ?

Quelles sont les hypothèses de régression par discontinuité dans ce contexte ?

Pourquoi cela teste-t-il la signalisation ?


Economics
Semesters
8
9
10
11
12
1998
2000
2002
2004
2006
2008
Cohort
Business
Semesters
8.5
9
9.5
10
10.5
11
1998
2000
2002
2004
2006
2008
Cohort

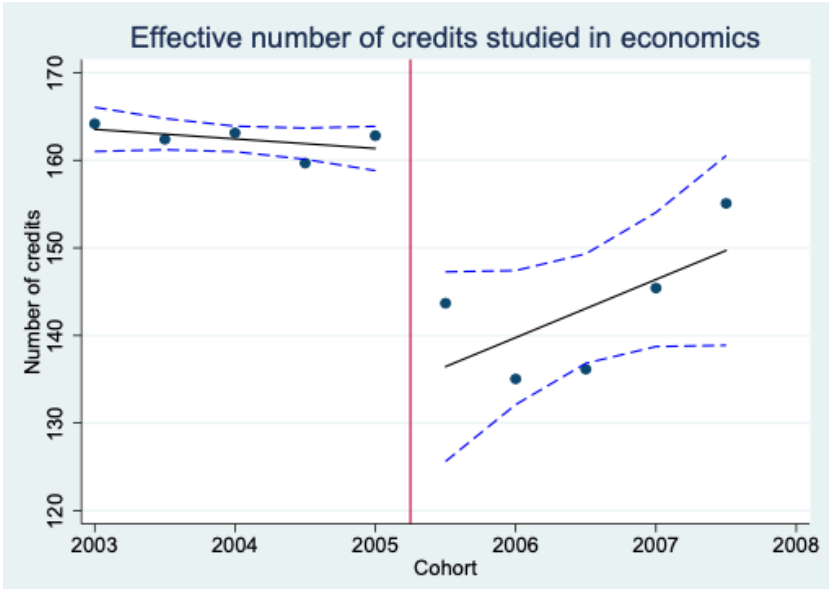
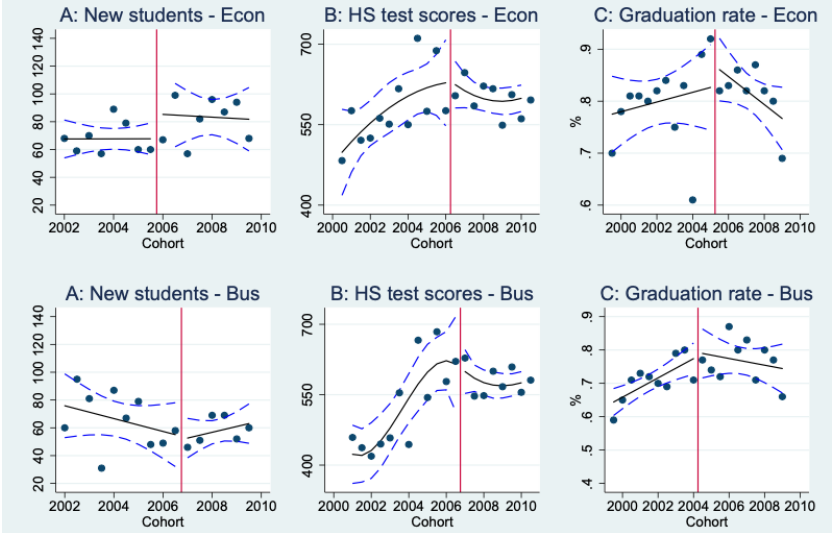

Effective number of credits studied in economics
Number of credits
120
130
140
150
160
170
2003
2004
2005
2006
2007
2008
Cohort

Table 1
First stage – The effect of the reform on instruction and class quality.
Source: Annual bulletin – Universidad de los Andes & Department of economics.

Dep variable	Degree duration		No. of credits	Class size		HS test scores		Graduation rates	
	Econ	Buss	Econ	Econ	Buss	Econ	Buss	Econ	Buss
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post	−1.038**	−0.916***	−24.37**	8.56	−22.15	−1.396	36.42	0.0192	0.00097
	[0.367]	[0.262]	[6.751]	[14.74]	[32.20]	[40.97]	[71.51]	[0.0635]	[0.0496]
Trend pre	0.0943	−0.0821	3.317*	1.024	2.086	−6.272	−1.655	−0.0135	−0.00503
	[0.119]	[0.0528]	[1.637]	[2.248]	[4.124]	[5.546]	[7.913]	[0.0110]	[0.00606]
Trend post	−0.121***	−0.0309	−0.545	0.0952	−2.309	12.90***	19.93***	0.00692	0.0145**
	[0.0280]	[0.0266]	[1.637]	[2.248]	[1.899]	[3.755]	[3.801]	[0.00596]	[0.00606]
Constant	9.716***	10.51***	160.8***	68.08***	64.35***	637.6***	557.0***	0.842***	0.789***
	[0.222]	[0.181]	[5.430]	[9.403]	[5.537]	[21.28]	[16.13]	[0.0438]	[0.0376]
Obs	18	18	10	16	16	21	21	20	20
R squared	0.868	0.881	0.867	0.233	0.171	0.45	0.672	0.163	0.427

Standard errors in brackets below the coefficients.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.



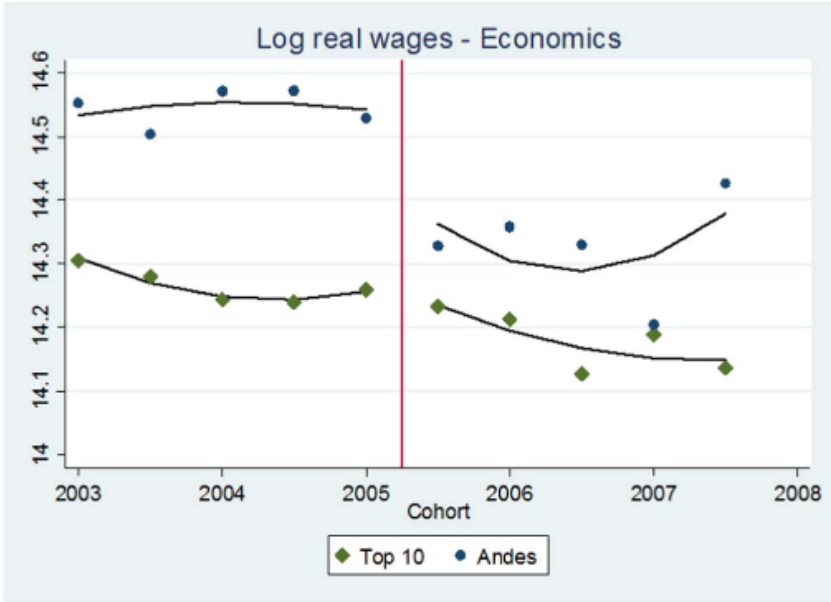
Source: Universidad de los Andes. The solid lines are the fitted values and dashed lines the 95% CI.

Table 2
Summary statistics.
Source: Ministry of Education, Colombia.

	Real wage	Experience	Age	Female	HS test	Family income[a]	Obs
Andes economics	3,017,001	2.6	25.8	0.46	58.1	5.93	1736
	1,776,674	1.9	2.2	0.50	5.5	1.44	
Top 10	2,119,275	2.98	26.26	0.59	51.28	3.75	3580
	1,457,070	1.98	2.83	0.49	6.01	1.76	
Andes business	3,192,033	2.5	25.8	0.46	58.1	5.93	2659
	1,959,143	1.8	2.2	0.50	5.5	1.44	
Top 10	2,141,599	2.90	26.24	0.59	51.33	3.82	22505
	1,522,623	2.01	2.79	0.49	6.03	1.76	
Other majors at Los Andes	2,482,154	2.66	25.8	0.55	57.6	5.87	6069
	1,695,091	1.99	2.2	0.50	5.4	1.53	

Note: Top rows show means and bottom standard deviation. Data from all students who started college between 2002 and 2007, and graduated after 2004. The top 10 universities were chosen using SABER PRO scores for schools of at least 1000 students.

[a] Based on a classification over 9 categories of income.


Log real wages - Economics
14.6
14.5
14.4
14.3
14.2
14.1
14
2003
2004
2005
2006
2007
2008
Cohort
Top 10
Andes


Wages - Business
Log real wages
14
14.1
14.2
14.3
14.4
14.5
14.6
2002
2004
2006
2008
Cohort
Top 10
Andes

Table 3
Baseline results. Effect of the reform on wages.
Source: Ministry of Education OLE and SPADIES.

Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Panel a: Economics						
Post*Andes	−0.163**	−0.161**	−0.167***	−0.164**	−0.164**	−0.161**
	[0.0500]	[0.0501]	[0.0505]	[0.0505]	[0.0501]	[0.0501]
Post	0.0817**	0.0819**	0.0721*	0.0744*	0.0810*	0.0865*
	[0.0293]	[0.0292]	[0.0311]	[0.0310]	[0.0366]	[0.0360]
Andes	0.312***	0.301***	0.312***	0.300***	0.311***	0.300***
	[0.0304]	[0.0301]	[0.0304]	[0.0301]	[0.0304]	[0.0301]
Experience	0.135***	0.154***	0.137***	0.154***	0.135***	0.156***
	[0.00842]	[0.0173]	[0.00841]	[0.0173]	[0.0127]	[0.0188]
Experience squared		−0.00424		−0.004		−0.00429
		[0.00431]		[0.00429]		[0.00431]
Female		−0.0912***		−0.0908***		−0.0914***
		[0.0223]		[0.0223]		[0.0224]
Constant	14.16***	14.20***	14.13***	14.17***	13.96***	14.19***
	[0.0197]	[0.0238]	[0.0495]	[0.0511]	[0.0846]	[0.0383]
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
Obs	3,621	3,621	3,621	3,621	3,621	3,621
R − sq	0.157	0.165	0.157	0.165	0.159	0.167
Panel b: Business						
Post*Andes	−0.136***	−0.136***	−0.141***	−0.141***	−0.135**	−0.136**
	[0.0410]	[0.0413]	[0.0412]	[0.0414]	[0.0412]	[0.0414]
Post	0.0952***	0.0940***	0.0555**	0.0558**	0.0971***	0.0991***
	[0.0153]	[0.0152]	[0.0185]	[0.0185]	[0.0189]	[0.0188]
Andes	0.429***	0.423***	0.432***	0.427***	0.429***	0.423***
	[0.0312]	[0.0316]	[0.0312]	[0.0316]	[0.0312]	[0.0315]
Experience	0.124***	0.145***	0.128***	0.147***	0.125***	0.151***
	[0.00517]	[0.0115]	[0.00512]	[0.0115]	[0.00782]	[0.0120]
Experience squared		−0.00525		−0.00481		−0.00635*
		[0.00303]		[0.00303]		[0.00302]
Female		−0.0976***		−0.0969***		−0.0979***
		[0.0147]		[0.0147]		[0.0148]
Constant	14.06***	14.11***	13.96***	14.00***	14.15***	14.10***
	[0.0129]	[0.0160]	[0.0317]	[0.0337]	[0.0968]	[0.0243]
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
N	10,348	10,348	10,348	10,348	10,348	10,348
R − sq	0.122	0.130	0.124	0.132	0.122	0.131