Human Capital in the Labor Market: Education, Learning and Experience

Valdemar Neto

Labor Economics

valdemar.pinho@fgv.br

April 14, 2025

Human Capital and Education

- **First topic: Human capital and education**
- Motivated by the human capital paradigm
- Worker skills as a form of capital
- Choose how much to invest in skills, balancing:
  - Increased earnings in the future
  - Opportunity cost of earnings foregone in the present
- Key parameter: **Causal return to schooling**
- The causal return to schooling answers a counterfactual question:
  - How much more would a particular person earn if they spent more time in school?
- We will discuss such questions in the language of potential outcomes.

Potential Outcomes

- Consider a person deciding whether to attend college.
- The indicator $D_i \in \{0, 1\}$ takes a value of 1 if i attends college, and 0 otherwise.
- $Y_i(1)$ denotes i's potential earnings if she attends college.
- $Y_i(0)$ denotes i's potential earnings if she does not attend college.
- Potential outcomes are defined by a hypothetical manipulation: what would happen to a particular person in one condition or the other.
- The causal effect of college on person i's earnings is defined as:

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

- This simple model of causality is called the Rubin causal model (Holland 1986).

The Fundamental Problem of Causal Inference

- In the real world, a person either attends college, or she doesn't.
- This means only one potential outcome will ever be observed – the other is counterfactual.
- The observed outcome, Y_i, equals $Y_i(0)$ if $D_i = 0$ and $Y_i(1)$ if $D_i = 1$. We can then write:

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

- Since we can never observe both $Y_i(0)$ and $Y_i(1)$, we can't see δ_i for any individual. This is known as the fundamental problem of causal inference.
- The econometric methods we will cover can be viewed as approaches to imputing missing potential outcomes.
- We can never hope to recover δ_i for an individual person, but sometimes we can recover certain averages.

Average Treatment Effects

- The average treatment effect for a population is defined as:

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

- "Treatment effects" language is adopted from medical trials.
  - $Y_i(1)$ is i's outcome if assigned the treatment (college).
  - $Y_i(0)$ is i's outcome if assigned the control condition (no college).
  - $\delta_i = Y_i(1) - Y_i(0)$ is i's treatment effect.

- Other treatment effect parameters of interest include the effect of treatment on the treated (TOT), and the effect of treatment on the non-treated (TNT):

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

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

Treatment Effects and Selection Bias

- Consider a comparison of average observed earnings for individuals that attend college vs. those that don't:

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

- Add and subtract $E[Y_i(0) \mid D_i = 1]$ on the right-hand side:

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] = E[Y_i(1) - Y_i(0) \mid D_i = 1]$$
$$+E[Y_i(0) \mid D_i = 1] - E[Y_i(0) \mid D_i = 0]$$

Treatment Effects and Selection Bias

$$E\left[Y_i \mid D_i = 1\right] - E\left[Y_i \mid D_i = 0\right] = \underbrace{E\left[Y_i(1) - Y_i(0) \mid D_i = 1\right]}_{\text{TOT}}$$

$$+ \underbrace{E\left[Y_i(0) \mid D_i = 1\right] - E\left[Y_i(0) \mid D_i = 0\right]}_{\text{Selection Bias}}$$

- This expression decomposes the observed treatment/control difference into the TOT plus a selection bias term given by the difference in average $Y_i(0)$'s between treatment and control.
- Selection bias arises if the observed outcome for the control group fails to match the missing counterfactual for the treatment group.

The RCT Ideal

- Suppose the treatment is assigned independently of potential outcomes:

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i$$

- Then:

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

- Assigning treatment randomly as in a randomized controlled trial (RCT) guarantees independence of potential outcomes from treatment.
  - Randomization eliminates selection bias.
  - Implies treatment/control difference $= ATE = TOT = TNT$.
- Often the treatment of interest is not randomized. Other research designs aim to isolate comparisons that are as good as random.

Human Capital Investments

- Return to the idea of human capital investment.
- Start with a simple model of schooling investments, as in Card (1999).
- Individual i chooses duration of schooling S to maximize the present discounted value of earnings:

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

- The potential earnings function $Y_i(S)$ now describes i's potential earnings for every possible schooling level.
- Attending S years of school results in zero earnings until S, and then $Y_i(S)$ thereafter.
- Interest rate r_i determines how i discounts future earnings.

Optimal Schooling Choice

- Optimal schooling choice maximizes PDV:

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

.

- First-order condition:

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

.

- Marginal benefit/marginal cost formula: at any S, can invest current earnings $Y_i(S)$ and earn return r_i, or defer earnings to earn more later, with proportional return $\frac{Y_i'(S)}{Y_i(S)}$.
- Optimal schooling equalizes returns on these two investments.
- Individual i's realized earnings are $Y_i(S_i^*)$.

Ability Bias

- Empirical literature tries to estimate features of the potential earnings functions $Y_i(S)$.
- Problem: As usual, we only see one earnings level for each person, corresponding to potential earnings at his/her chosen schooling level.
- Why do people choose different levels of schooling? In the model, differences must be driven either by variation in the discount rate, or in the potential earnings function.
- "Ability bias": Individuals that choose different schooling levels may have different potential earnings functions, leading observed returns to schooling to differ from causal returns.
  - Label for selection bias in the returns to schooling context.

Observed Returns to Schooling

- Consider an ordinary least squares (OLS) regression of observed earnings Y_i on schooling S_i:

$$Y_i = a + bS_i + e_i$$

- The observed return to schooling is the OLS slope coefficient:

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

- Question: Should I be worried about whether S_i is correlated with the error term e_i?

OLS Approximates the CEF

- Answer: No. By definition, the OLS residual e_i is orthogonal to the regressor S_i:

$$\begin{aligned} \text{Cov}(e_i, S_i) &= \text{Cov}(Y_i - a - bS_i, S_i) \\ &= \text{Cov}(Y_i, S_i) - b\text{Var}(S_i) \\ &= \text{Cov}(Y_i, S_i) - \left(\frac{\text{Cov}(Y_i, S_i)}{\text{Var}(S_i)}\right)\text{Var}(S_i) \\ &= 0. \end{aligned}$$

- OLS always gives a minimum mean squared error approximation to the conditional expectation function (CEF), $E[Y_i \mid S_i]$:

$$(a, b) = \underset{(a_0, b_0)}{\arg\min}\; E\left[(E[Y_i \mid S_i] - a_0 - b_0 S_i)^2\right].$$

- OLS fits the CEF regardless of what model you have in mind. Better to ask: is the CEF economically interesting?

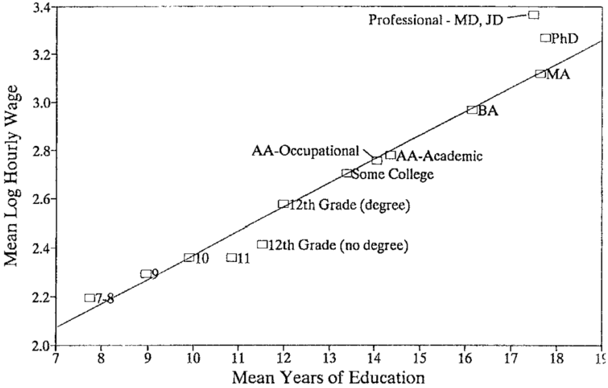


Fig. 2. Relationship between mean log hourly wages and completed education, men aged 40–45 in 1994–1996 Current Population Survey. Mean education by degree category estimated from February 1990 CPS.

Ability Bias

- Consider a constant effects potential earnings function:

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

- The causal return $\beta > 0$ is the same for all people and schooling levels.

- This model implies

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

- Suppose the interest rate r_i is the same for everyone. Is the observed return to schooling too big or too small relative to the causal return?

Negative Ability Bias

- The observed return is too small.
- When $r_i = r$ for all i, everyone earns the same amount:

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

- The observed return is therefore zero, which is less than the causal return β.
- Intuition: The primary cost of schooling is the opportunity cost of earnings foregone. Higher-ability people face higher opportunity costs and so drop out earlier.
- In this case "ability bias" is negative — the causal return exceeds the observed return.

General Ability Bias

- More generally, the observed return to schooling is

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

- Ability bias depends on variances and covariances of discount rates and ability across people.
- Direction is unclear *a priori*.
- To get positive ability bias, need another force that overrides the basic opportunity cost story.

Estimating Causal Returns

- The observed return to schooling may be contaminated by ability bias of unclear sign and magnitude. How can we estimate the causal return?
- Maintain the simple constant-effects model for potential earnings:

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

- We can then write observed earnings as

$$Y_i = \bar{\alpha} + \beta S_i + \epsilon_i.$$

- Here $\bar{\alpha} = E[\alpha_i]$ and $\epsilon_i = \alpha_i - \bar{\alpha}$.
- Question: Should I be worried about whether S_i is correlated with the error term ϵ_i?

Observed and Causal Returns

$$Y_i = \bar{\alpha} + \beta S_i + \epsilon_i.$$

- Answer: Yes. The coefficient β is now defined as a parameter from a causal (potential outcomes) model, so there is no guarantee that $\text{Cov}(S_i, \epsilon_i) = 0$.
- Schooling is not randomly assigned, so it may not be independent of potential outcomes, summarized here by ϵ_i.
- This means the OLS slope coefficient b may not coincide with the causal effect β.
- **Instrumental variables** (IV) is a common research design that seeks to eliminate selection bias in nonexperimental data.

Instrumental Variables

$$Y_i = \bar{\alpha} + \beta S_i + \epsilon_i.$$

- Suppose we have a third variable, Z_i, that satisfies two conditions:
  1. **First stage:** $\text{Cov}(S_i, Z_i) \neq 0$.
  2. **Exclusion restriction:** $\text{Cov}(\epsilon_i, Z_i) = 0$.
- First stage requires Z_i (the instrument) to be correlated with S_i (the endogenous variable).
- Exclusion requires the instrument to be uncorrelated with potential outcomes (here, ϵ_i).
  - Z_i must be as good as randomly assigned.
  - Z_i cannot affect Y_i through channels other than S_i.

The Population IV Coefficient

- Covariance between outcome and instrument:

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

$$= \beta \text{Cov}(S_i, Z_i) + \text{Cov}(\epsilon_i, Z_i)$$

- Exclusion implies the second term is zero, so $\text{Cov}(Y_i, Z_i) = \beta \text{Cov}(S_i, Z_i)$.

- First stage implies $\text{Cov}(S_i, Z_i) \neq 0$, so we can divide by this covariance to solve for β:

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

- The ratio of covariances on the left is the population instrumental variables coefficient, β_{IV}.

IV Interpretation

- Divide the top and bottom of the IV coefficient by $\text{Var}(Z_i)$ to obtain:

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

- The IV coefficient is a ratio of two regression coefficients:
  - The **reduced form** regression of Y_i on Z_i.
  - The **first stage** regression of S_i on Z_i.

- Suppose Z_i is binary. Then the IV coefficient becomes a Wald ratio of two differences in means:

$$\beta_{IV} = \frac{E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]}{E[S_i \mid Z_i = 1] - E[S_i \mid Z_i = 0]}.$$

IV Estimates of the Return to Schooling: Angrist and Krueger (1991)

- Angrist and Krueger (QJE 1991): classic study reporting IV estimates of the return to education.
- Instrumental variables strategy motivated by interaction between compulsory schooling and age-at-entry laws.
  - Students can typically drop out of school on the day they turn 16.
  - Birth date cutoff for starting age: Students usually start school in the fall of the calendar year in which they turn six.
- Generates differences in mean educational attainment by date of birth.

Birth date	School start date	Dropout date	Schooling at dropout date
January 2, 1930	September 1, 1936	January 2, 1946	9.5 years
December 31, 1930	September 1, 1936	December 31, 1946	10.5 years

QOB Instruments

- AK's instrument is date of birth.
- Operationalize using quarter of birth (QOB), which is available in US Census data.
  - $Z_i = 1\{i \text{ was born in first quarter}\}$
- What do the first stage and exclusion restriction assumptions mean for a QOB instrument?

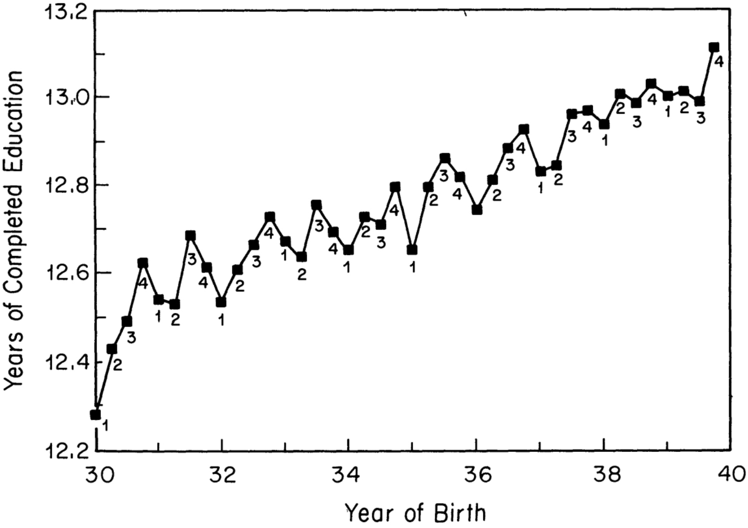


FIGURE I
Years of Education and Season of Birth
1980 Census
Note. Quarter of birth is listed below each observation.

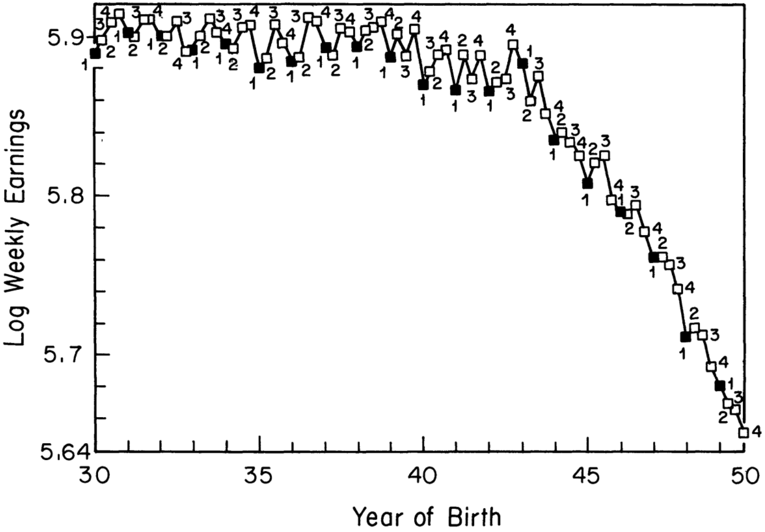


FIGURE V
Mean Log Weekly Wage, by Quarter of Birth
All Men Born 1930–1949; 1980 Census

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)

a. The sample size is 247,199 in Panel A, and 327,509 in Panel B. Each sample consists of males born in the United States who had positive earnings in the year preceding the survey. The 1980 Census sample is drawn from the 5 percent sample, and the 1970 Census sample is from the State, County, and Neighborhoods 1 percent samples.

b. The OLS return to education was estimated from a bivariate regression of log weekly earnings on years of education.

QOB Interpretation

- IV estimates based on QOB suggest a return to schooling of 7-10% per year.
- IV estimates are comparable to or bigger than corresponding OLS estimates.
- Card (1999) finds a similar pattern for other IV strategies.
- In our simple model, this suggests negative ability bias: people with lower earnings potential attend school for longer.
- Other interpretations?

Heterogeneous Treatment Effects

- Our simple model assumed constant effects of schooling across people.
- Return to general potential outcomes model with binary treatment D_i and potential outcomes $Y_i(1)$ and $Y_i(0)$.
- Suppose we have a binary instrument Z_i, and consider two new potential outcomes defined by a hypothetical manipulation of Z_i:
  - $D_i(1)$: i's treatment status if $Z_i = 1$.
  - $D_i(0)$: i's treatment status if $Z_i = 0$.
- Observed treatment is $D_i = D_i(0) + (D_i(1) - D_i(0))Z_i$.

IV Assumptions

- IV assumptions in a heterogeneous treatment effects world:
  1. **Independence/exclusion:** $(Y_i(1), Y_i(0), D_i(1), D_i(0)) \perp\!\!\!\perp Z_i$
  2. **First stage:** $\Pr[D_i = 1 \mid Z_i = 1] > \Pr[D_i = 1 \mid Z_i = 0]$
  3. **Monotonicity:** $D_i(1) \geq D_i(0) \ \forall i$
- Relative to our constant effects IV setup, monotonicity is the novel assumption.
- Monotonicity requires the instrument to affect everyone's treatment status in the same direction.

Compliance Groups

- Under monotonicity, we can partition the population into three groups defined by their behavioral responses to the instrument (Angrist, Imbens, and Rubin 1996):

  1. **Always takers:** $D_i(1) = D_i(0) = 1$
  2. **Never takers:** $D_i(1) = D_i(0) = 0$
  3. **Compliers:** $D_i(1) = 1,\ D_i(0) = 0$

- Compliers have $D_i(1) > D_i(0)$: their treatment status increases with the instrument.

- Monotonicity rules out *defiers* with $D_i(1) = 0,\ D_i(0) = 1$.

Local Average Treatment Effects

- Under these assumptions, IV identifies a **local average treatment effect (LATE)**:

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

- This is the LATE theorem of Imbens and Angrist (1994).
- LATE is the average treatment effect for compliers – individuals whose treatment status is determined by the instrument.

LATE Theorem: Proof

- Note that $Y_i = Y_i(D_i) = Y_i(D_i(Z_i))$, so by independence

$$E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0] = E[Y_i(D_i(1)) \mid Z_i = 1] - E[Y_i(D_i(0)) \mid Z_i = 0]$$
$$= E[Y_i(D_i(1)) - Y_i(D_i(0))].$$

- By monotonicity we either have $D_i(1) = D_i(0)$ or $D_i(1) > D_i(0)$, so

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

- The same logic implies

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

so

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

Interpreting IV Estimates

- LATE interpretation suggests that QOB instrument identifies the causal effect of extra schooling for individuals on the margin of dropping out early around mid-century.
- Next, we will consider more recent evidence looking at other schooling margins.

Returns to College for Marginal Students: Zimmerman (2014)

- Observed return to college has increased dramatically in recent decades.
  - College wage premium rose from 50% to 97% between 1980 and 2008 (Acemoglu and Autor, 2011).
  - May reflect fast growth of skill demand coupled with slow growth of skill supply (Goldin and Katz, 2008).
- At the same time, many students in the US start college but don't finish.
  - 62% of students attending four-year colleges graduate within 6 years (NCES, 2020).
  - Does college attendance improve earnings for academically marginal students?
- Zimmerman (JOLE 2014) leverages a **regression discontinuity design** to study returns for students on the margin of four-year college enrollment.

Regression Discontinuity Designs

- Consider a setting with a binary treatment $D_i \in \{0, 1\}$, and potential outcomes $Y_i(1)$ and $Y_i(0)$.
- Suppose the treatment is a deterministic and discontinuous function of an observed covariate R_i, such that

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

- R_i is called the **running variable** or **forcing variable**.
- This is a **sharp RD** because the probability of treatment switches from zero to one at the threshold.
- Zimmerman (2014): GPA cutoff for admission to state universities in Florida.

Regression Discontinuity Designs

- We get to observe $Y_i(1)$ when $R_i > c$ and $Y_i(0)$ when $R_i \leq c$.
- Basic idea of the RD design: Compare observations just above and just below the threshold to infer treatment effect.
- Intuitively, the treatment may be as good as randomly assigned for individuals in the neighborhood of $R_i = c$, so comparing treated and nontreated near c reveals a treatment effect.

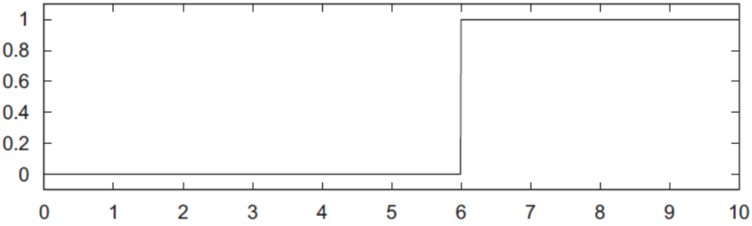


Fig. 1. Assignment probabilities (SRD).

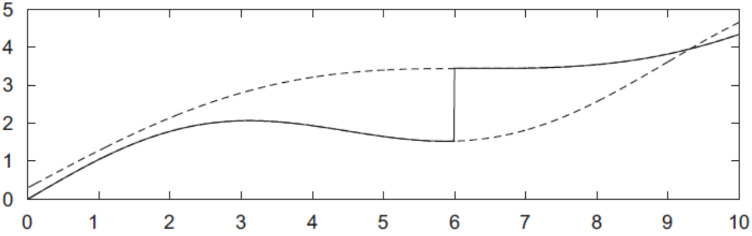


Fig. 2. Potential and observed outcome regression functions.

RD Identification

- Key assumption: potential outcomes are smooth at the threshold.
- Formally:

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

- Potential outcome CEFs must be continuous at the threshold.
- The population just below must not be discretely different from the population just above.

RD Identification

- If this assumption holds we have

$$
\begin{aligned}
&\lim_{r \to c^+} E\left[Y_i \mid R_i = r\right] - \lim_{r \to c^-} E\left[Y_i \mid R_i = r\right] \\
&= \lim_{r \to c^+} E\left[Y_i(1) \mid R_i = r\right] - \lim_{r \to c^-} E\left[Y_i(0) \mid R_i = r\right] \\
&= E\left[Y_i(1) \mid R_i = c\right] - E\left[Y_i(0) \mid R_i = c\right] \\
&= E\left[Y_i(1) - Y_i(0) \mid R_i = c\right]
\end{aligned}
$$

- When potential outcomes are smooth around the threshold, a comparison of individuals just above and just below yields the average treatment effect for those at the threshold.

- Identification argument is nonparametric: we don't need to assume anything about the distribution of potential outcomes other than continuity of CEFs.

RD Interpretation

- Core RD intuition: for those near the threshold, things could have gone either way.
- Interpret RD as a local randomized trial among those sufficiently close to $R_i = c$.
- Explains why RD evidence can be especially compelling relative to other research designs – close to RCT ideal.
- "Local randomization" view motivates common RD diagnostics.
  - Check balance of pre-determined characteristics for observations above and below the threshold.
  - Look for anomalies in the distribution of the running variable around the threshold, which may indicate manipulation (McCrary, 2008).

Fuzzy RD

- Sometimes treatment is generated by a discontinuous assignment rule that isn't deterministic.
- Suppose that

$$\lim_{r \to c^-} \Pr\left[D_i = 1 \mid R_i = r\right] < \lim_{r \to c^+} \Pr\left[D_i = 1 \mid R_i = r\right]$$

- The probability of treatment jumps at $R_i = c$, but not necessarily from zero to one.
- This is a **fuzzy RD** scenario because treatment is only partly determined by the threshold.
- Zimmerman (2014): Students above GPA cutoff are eligible for admission, but not guaranteed.

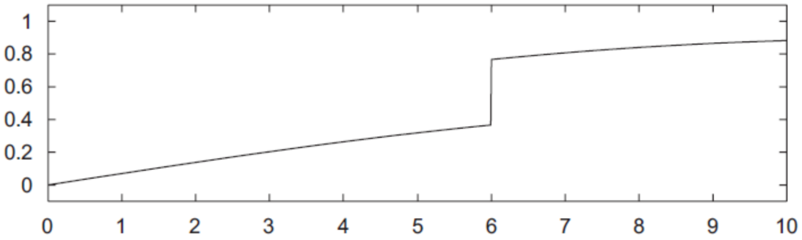


Fig. 3. Assignment probabilities (FRD).

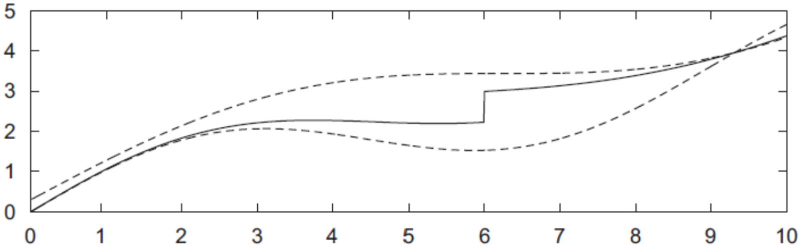


Fig. 4. Potential and observed outcome regression (FRD).

Fuzzy RD Assumptions

- As before, assume the distributions of $Y_i(1)$ and $Y_i(0)$ are smooth around the threshold.
- Let $D_i(1)$ and $D_i(0)$ denote potential treatment statuses for individual i if s/he were located above and below the threshold. Assume these are also smooth across the threshold, and

$$D_i(1) \geq D_i(0) \quad \forall i$$

- Crossing the threshold weakly increases the likelihood of treatment for everyone.

Fuzzy RD

- Under these assumptions, we have

$$\frac{\lim_{r \to c^+} E\left[Y_i \mid R_i = r\right] - \lim_{r \to c^-} E\left[Y_i \mid R_i = r\right]}{\lim_{r \to c^+} E\left[D_i \mid R_i = r\right] - \lim_{r \to c^-} E\left[D_i \mid R_i = r\right]}$$

$$= E\left[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0), R_i = c\right]$$

- The numerator on the left is the jump in outcomes at the threshold, as in a sharp RD.
- The denominator is the change in the probability of treatment at the threshold.
- The ratio of the jump in the outcome to the jump in the treatment probability identifies an average treatment effect for individuals who switch treatment status at the threshold.
- Sound familiar?

Fuzzy RD is IV

- Fuzzy RD is IV using a threshold indicator $Z_i = 1\{R_i > c\}$ as an instrument for treatment in the neighborhood of the threshold.
- Think of fuzzy RD as a local randomized trial with non-compliance.
- Implies fuzzy RD estimates are local in two senses:
  - Local to the threshold, $R_i = c$ (also applies to sharp RD).
  - Only apply to compliers at the threshold (that's the "local" in LATE).

RD Implementation

- Implementing RD requires estimating the left- and right-hand limits of average outcomes and treatment probabilities.
- Bias/variance tradeoff: using data away from the threshold increases sample size, but may introduce bias if potential outcomes are related to the running variable.
- In practice RD is usually implemented with local linear regression.
  - Regress outcome on the running variable among observations within a small bandwidth of the threshold, with weights that decline with distance to threshold.
  - RD estimate is difference in fitted regression functions above and below the threshold.
- Recent econometric literature proposes optimal bandwidths that balance bias and variance to minimize mean squared error, automated in `rdrobust` Stata command (Imbens and Kalyanaraman, 2011; Calonico et al., 2014).

Returns to College for Marginal Students: Zimmerman (2014)

- Zimmerman (2014) uses a GPA cutoff to estimate the returns to four-year college admission at public institutions in Florida.
- Students above the cutoff are eligible for admission to schools in the Florida State University System (SUS).
- In practice, the cutoff is relevant for admission to Florida International University (FIU), a large SUS campus in Miami.
- Population around the FIU admission cutoff has relatively low SAT scores (21st percentile nationwide) and low graduation rates.
- Estimates are therefore informative about returns to college for marginal students.

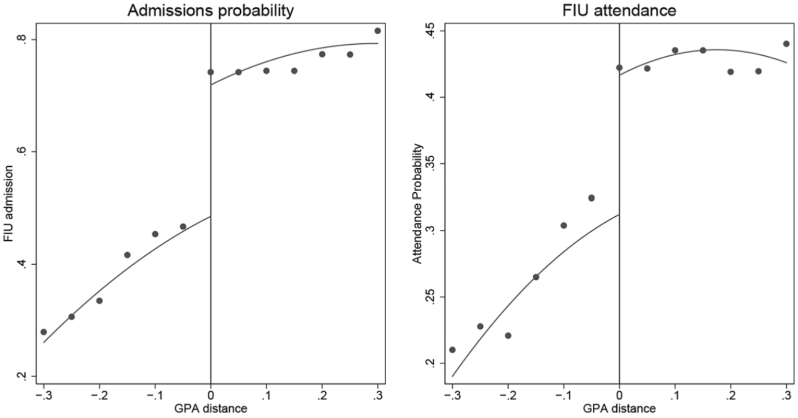


FIG. 4. Admissions and FIU attendance. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

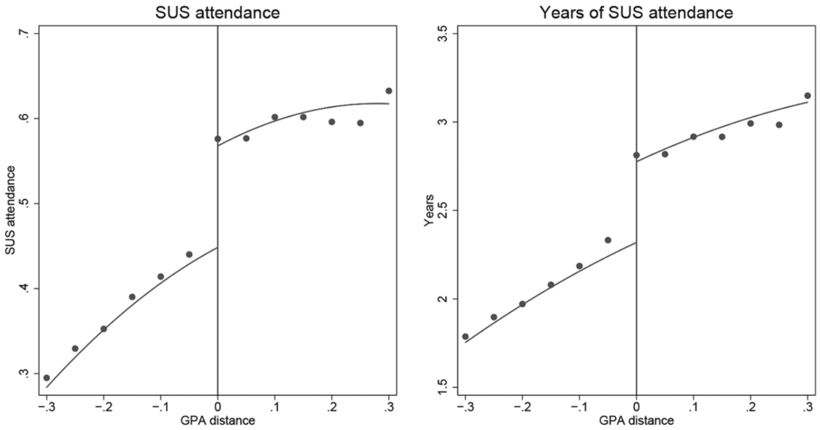


FIG. 5. SUS attendance and persistence. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

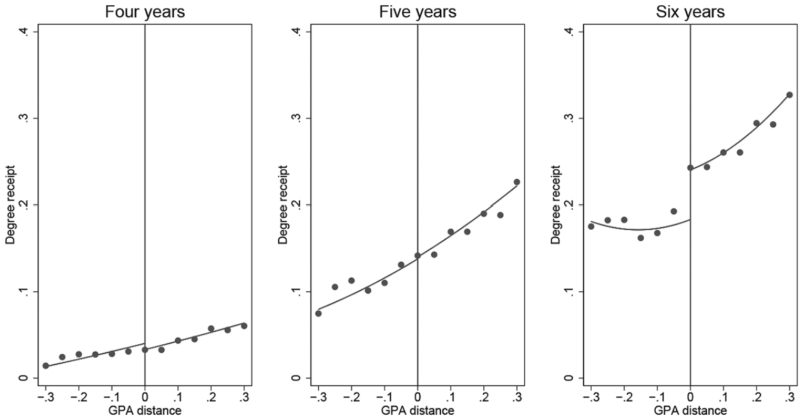


FIG. 6. SUS BA receipt by years elapsed since high school. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

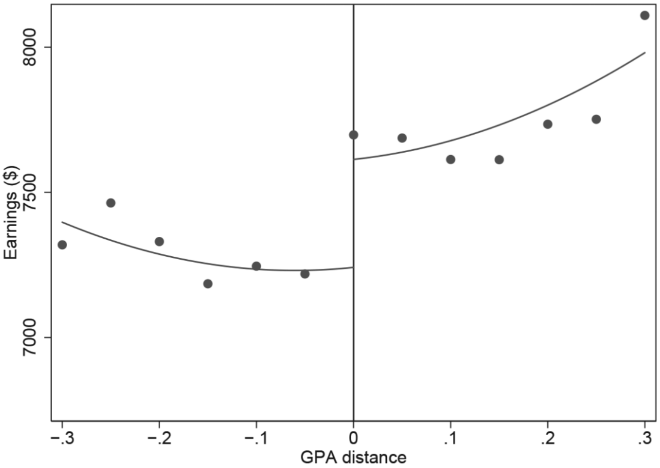


FIG. 8. Quarterly earnings by distance from GPA cutoff. Lines are fitted values based on the main specification. Dots, shown every .05 grade points, are rolling averages of values within .05 grade points on either side that have the same value of the threshold-crossing dummy.

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.

Human Capital vs. Signaling

- Evidence so far suggests that education increases earnings.
- Conventional human capital view is that schooling investments raise earnings by boosting productivity.
- **Signaling models** (Spence, 1973) provide an alternative explanation for the return to schooling.
  - If employers cannot observe ability, schooling may serve as a costly signal that separates low- and high-ability types, rather than increasing productivity.
  - Implies schooling is pure social waste: burns resources to create inequality.
- Distinguishing between human capital and signaling views is essential for education policy.
- Signaling models provide an explanation for the fact that observed return to schooling is especially large for grade 12.

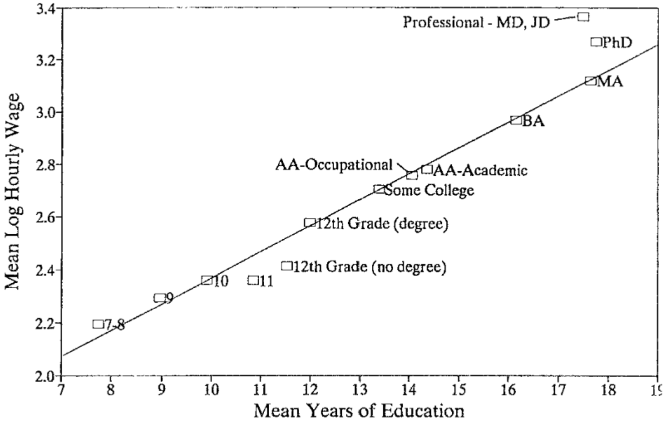


Fig. 2. Relationship between mean log hourly wages and completed education, men aged 40-45 in 1994 1996 Current Population Survey. Mean education by degree category estimated from February 1990 CPS.

Signaling Value of a High School Diploma: Clark and Martorell (2014)

- Clark and Martorell (JPE 2014) use an RD design to estimate the causal effect of high school graduation on earnings.
- CM use the fact that students in Texas must pass exams before graduating high school.
- Testing starts in 10th grade and students can try multiple times, but eventually face a "last chance" exam at the end of 12th grade.
- Students who just barely fail vs. barely pass should have similar human capital, but differ in educational credentials.
- RD therefore plausibly identifies the signaling value of a diploma.
- The last-chance exams are not completely deterministic – so the RD is fuzzy.

TABLE 6

ASSOCIATIONS BETWEEN DIPLOMA AND TEST SCORES AND EARNINGS

	A. MEAN DIFFERENCES BY DIPLOMA STATUS				
	Last-Chance Sample	Complete Grade 12, No College			
		All	T1	T2	T3
	(1)	(2)	(3)	(4)	(5)
Earnings years 7–11	1,814.7	2,867.8	1,780.3	1,752.0	2,385.3
	(138.1)	(79.3)	(111.8)	(176.1)	(228.5)
Observations	128,460	992,031	210,793	193,970	194,896
Mean earnings without diploma	12,400	12,673	11,858	13,301	13,538
Difference (%)	14.6	22.6	15.0	13.2	17.6
PDV earnings	8,054.5	8,731.0	7,280.7	7,459.4	10,546.3
	(632.3)	(341.9)	(501.9)	(779.8)	(951.4)
Observations	37,571	340,028	74,490	63,652	64,548
Mean earnings without diploma	70,280	69,992	66,466	74,216	73,860
Difference (%)	11.5	12.5	11.0	10.1	14.3

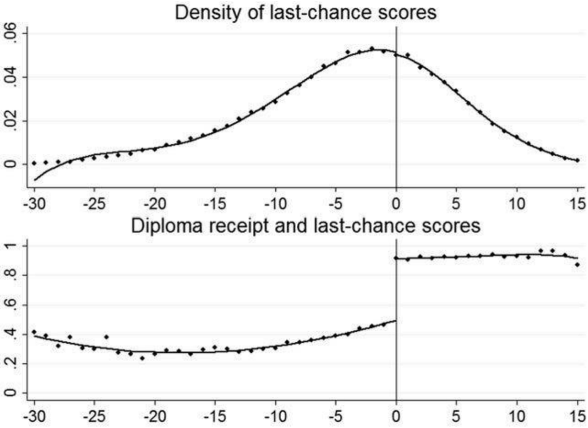
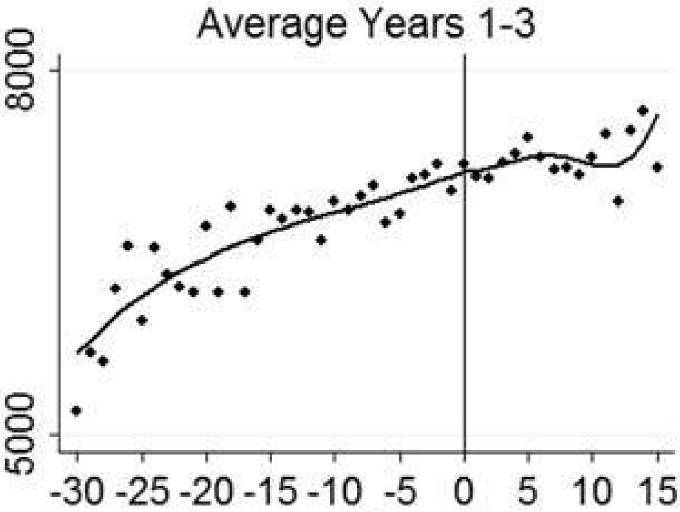
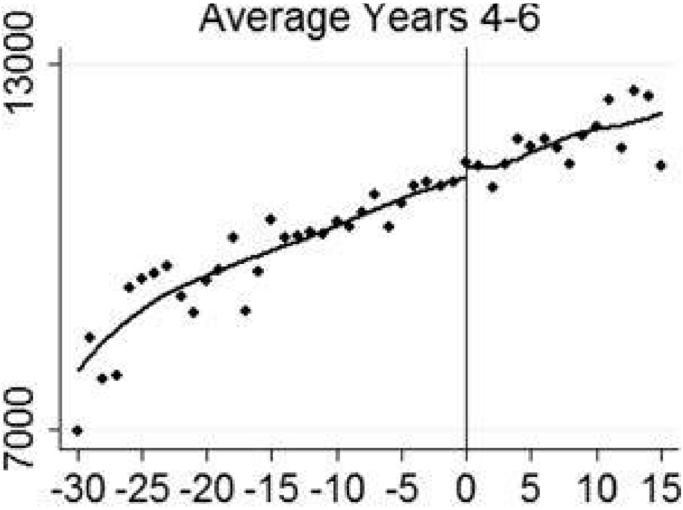
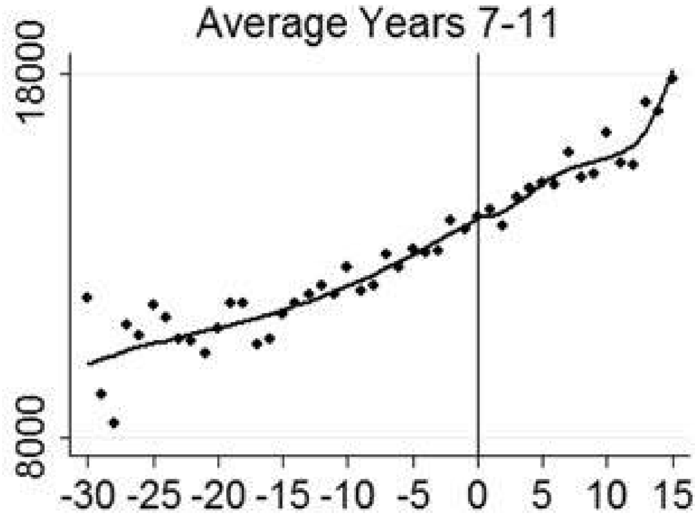
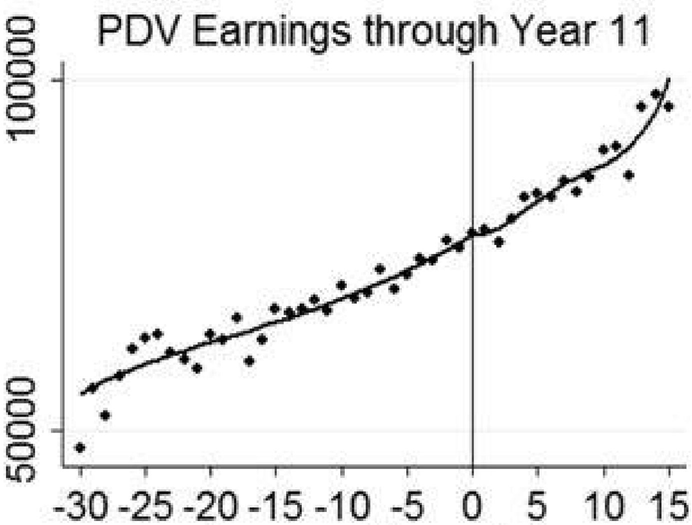


FIG. 1. Last-chance exam scores and diploma receipt. The graphs are based on the lastchance sample. See table 1 and the text. Dots are test score cell means. The scores on the x - axis are the minimum of the section scores recentered to be zero at the passing cutoff that are taken in the last-chance exam. Lines are fourth-order polynomials fitted separately on either side of the passing threshold.


Average Years 1-3
8000
5000
-30 -25 -20 -15 -10 -5 0 5 10 15

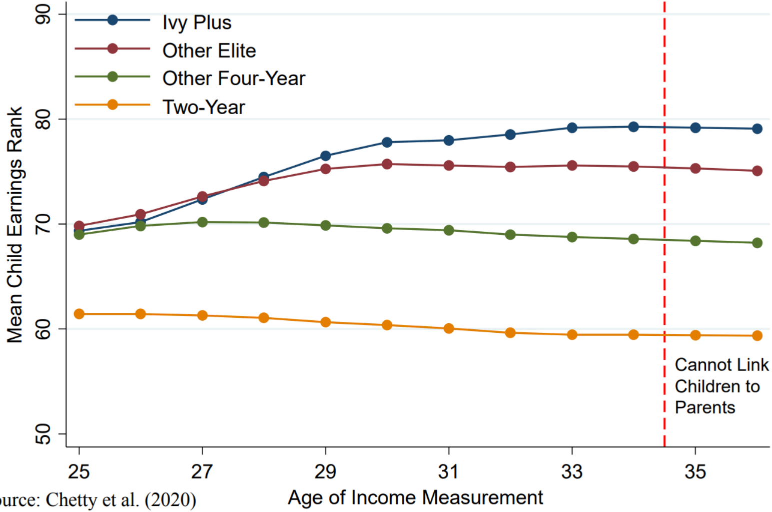

Average Years 4-6
13000
7000
-30 -25 -20 -15 -10 -5 0 5 10 15


Average Years 7-11
18000
8000
-30
-25
-20
-15
-10
-5
0
5
10
15

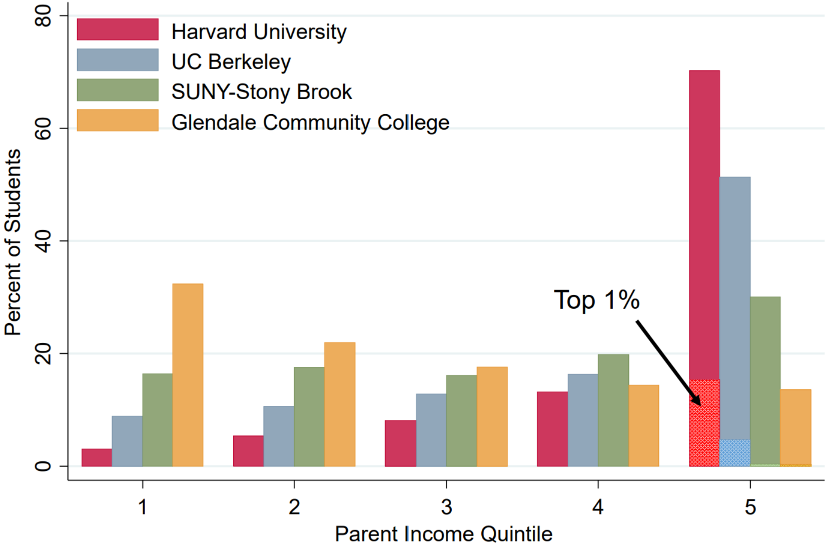

PDV Earnings through Year 11
100000
50000
-30 -25 -20 -15 -10 -5 0 5 10 15

Returns to College Selectivity

- For many students the relevant choice margin is which college to attend rather than years of schooling or college vs. no college.
- Very large differences in earnings between students attending different US colleges.
- But there is also a lot of selection into college choice.



Source: Chetty et al. (2020)


Harvard University
UC Berkeley
SUNY-Stony Brook
Glendale Community College
Top 1%
Percent of Students
0
20
40
60
80
1
2
3
4
5
Parent Income Quintile

Returns to College Selectivity

- Hard to find good experiments and quasi-experiments that induce variation in attendance at more vs. less selective colleges.
- Dale and Krueger (2002, 2014) use a **matching** approach that compares outcomes for students who applied and were admitted to the same sets of colleges, but attended different schools.
- Based on a **selection on observables** assumption: college choice is independent of potential outcomes conditional on a set of observed covariates.

Potential Outcomes Model

- Return to our causal model with binary treatment $D_i \in \{0, 1\}$ and potential outcomes $Y_i(1)$ and $Y_i(0)$.
- Suppose treatment isn't randomly assigned.
- As we've seen, the observed difference between average outcomes for individuals with $D_i = 1$ and $D_i = 0$ may be contaminated by selection bias.
- Suppose we also have data on a vector of observed covariates X_i.
- Dale and Krueger: D_i is attending a more selective college, and X_i is the list of colleges where a student applied and was admitted.

Selection on Observables

- Selection-on-observables approaches are based on a **conditional independence assumption (CIA)**:

$$(Y_i(1), Y_i(0)) \perp D_i \mid X_i$$

- CIA is also called "unconfoundedness," "ignorability," "exogeneity".
- The idea is that while potential outcomes and treatment may not be independent in general, they are independent conditional on a set of observed covariates - treatment is as good as random conditional on X_i.
- CIA necessarily holds in stratified RCTs, and may hold in non-experimental data with the right controls.

Full Covariate Matching

- Under CIA an obvious approach is to simply compare treatment and control groups conditional on the covariates.
- Let $\Delta(x)$ denote the observed treatment/control difference for a particular value of the covariates:

$$\Delta(x) \equiv E\left[Y_i \mid D_i = 1, X_i = x\right] - E\left[Y_i \mid D_i = 0, X_i = x\right]$$

- CIA implies

$$\begin{aligned}\Delta(x) &= E\left[Y_i(1) \mid D_i = 1, X_i = x\right] - E\left[Y_i(0) \mid D_i = 0, X_i = x\right] \\ &= E\left[Y_i(1) - Y_i(0) \mid X_i = x\right] \equiv ATE(x)\end{aligned}$$

- Covariate-specific treatment/control contrasts capture conditional average treatment effects.
- By computing $\Delta(x)$ for every value of x and then weighting appropriately, we can obtain any causal effect of interest. This is **full covariate matching**.

Computing Treatment Effects

- Under CIA, we can use full covariate matching to compute average treatment effects:

$$ATE = \sum_x Pr[X_i = x]\Delta(x)$$

$$TOT = \sum_x Pr[X_i = x \mid D_i = 1]\Delta(x)$$

$$TNT = \sum_x Pr[X_i = x \mid D_i = 0]\Delta(x)$$

OLS Regression as Matching

- Consider an OLS regression of outcomes on a treatment indicator, controlling for indicators for every value of the covariates X_i:

$$Y_i = a + bD_i + \sum_x \pi_x 1\{X_i = x\} + e_i$$

- This regression is **saturated** in the controls: there is a different coefficient for every value of X_i.
- With saturated controls, the OLS coefficient is

$$b = \sum_x \left(\frac{Pr[X_i = x] \text{Var}(D_i \mid X_i = x)}{\sum_{x'} Pr[X_i = x'] \text{Var}(D_i \mid X_i = x')} \right) \Delta(x)$$

- OLS with saturated controls is a version of full covariate matching.
  - "Saturate-and-weight" theorem (Angrist and Pischke, 2009).
  - Under CIA, generates a variance-weighted average treatment effect.

CIA Methods

- In practice, full covariate matching may not be feasible (e.g. many-valued or continuous controls).
- There are a variety of approaches to controlling for X_i in such cases:
  - OLS with additive controls.
  - Nearest-neighbor or kernel matching.
  - Propensity score matching/reweighting.
- These methods are not qualitatively different.
  - All are approaches to adjusting for covariates.
  - Coincide when the controls are flexible enough.
- Key to the research design is the underlying CIA assumption, not the particular method used to control for X_i.

Returns to College Selectivity: Dale and Krueger

- Dale and Krueger (QJE 2002, JHR 2014) take a matching/selection on observables approach to estimating the returns to college selectivity.
- Research design: compare students who applied to, and were admitted by, the same colleges, but chose to attend different schools.
- Intuition:
  Application choices capture a lot of students' information about their own ability, while admission decisions capture a lot of colleges' information about student ability.
- Data: College and Beyond (C&B).
  - Survey of students enrolled at 34 colleges, more selective than the US average.
  - 2014 paper matches C&B to administrative earnings data from the Social Security Administration (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 SAT/100	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)
	{.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	

Updating Dale/Krueger: Mountjoy and Hickman (2020)

- A recent paper by Mountjoy and Hickman (2020) updates the Dale/Krueger strategy using administrative data from Texas.
- Rather than looking at overall return to selectivity, estimate a "value-added" model with a different effect for every college, conditioning on DK application/admission controls.
- Relate college value-added to selectivity and other institution characteristics.
- Consistent with DK, Mountjoy and Hickman find limited returns to selectivity.
- Estimated college value-added is positively correlated with other inputs like instructional expenditures and faculty/student ratio.

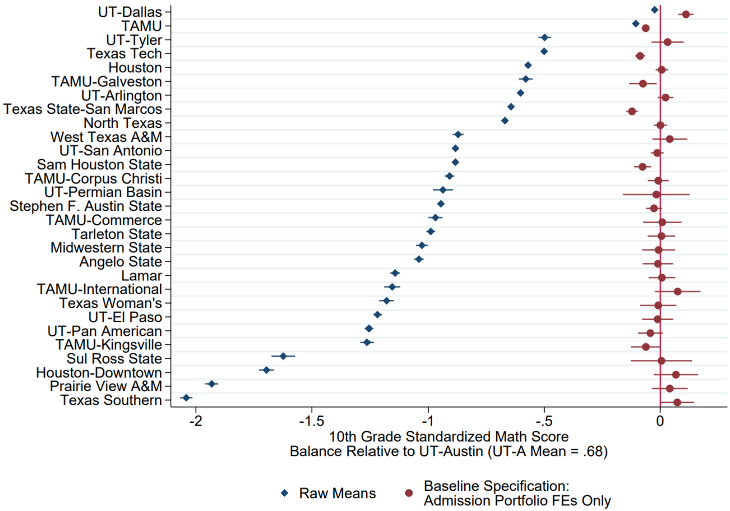
Relative College Value-Added and Economic Outcomes

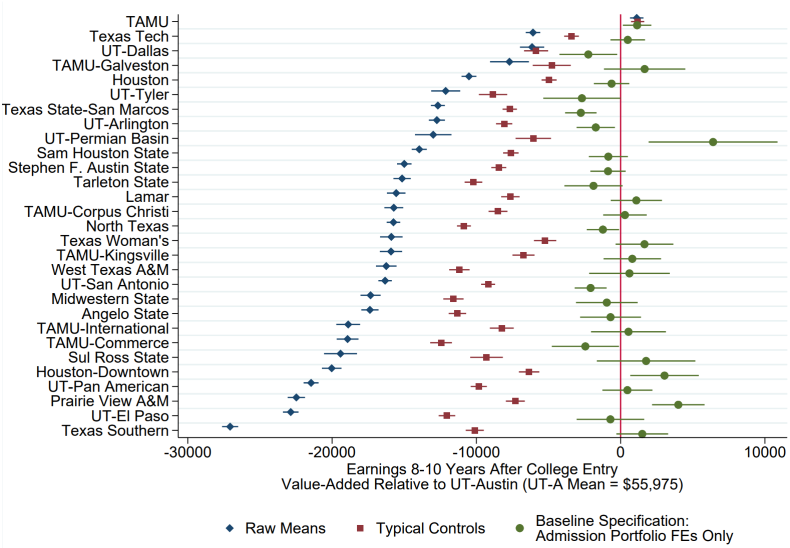
- Vast disparities in economic outcomes exist for students attending different U.S. colleges.
- Study focuses on relative value-added of colleges within student choice sets, using data linking high school, college, and earnings in Texas.
- Methodology:
  - Identify college value-added by comparing outcomes of students admitted to the same set of institutions.
  - Framework developed to account for match effects and sorting gains.

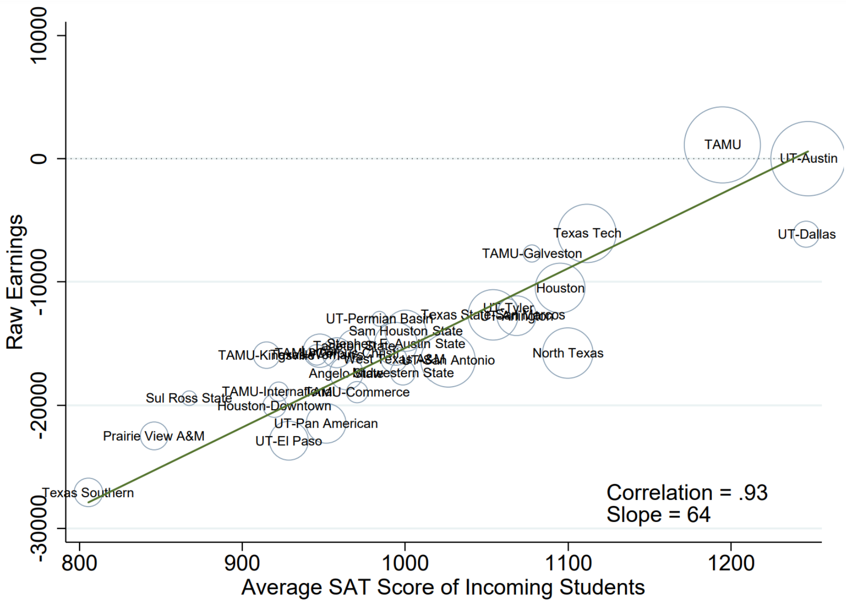
$$Y_{ij} = \underbrace{\alpha_i}_{\substack{\text{Student effect} \\ \text{(Ability, ambition, advantage)}}} + \underbrace{\eta_{ij}}_{\substack{\text{School effect} \\ \text{(Value-added)}}}$$

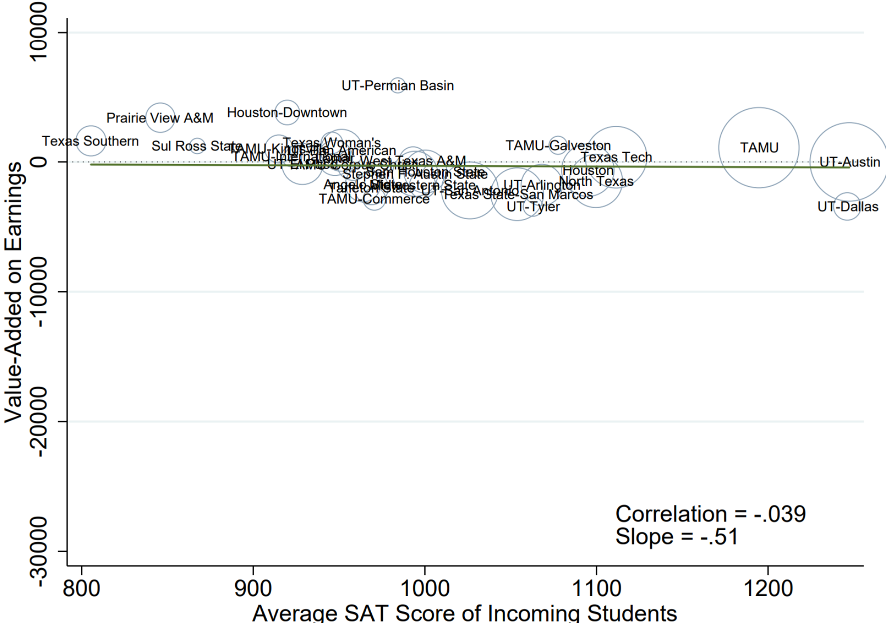
- Key Findings:
  - Selectivity is a weak predictor of long-term earnings.
  - Instructional spending is a stronger indicator of value-added.
  - Potential for mismatch effects when value-added varies by student characteristics.

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



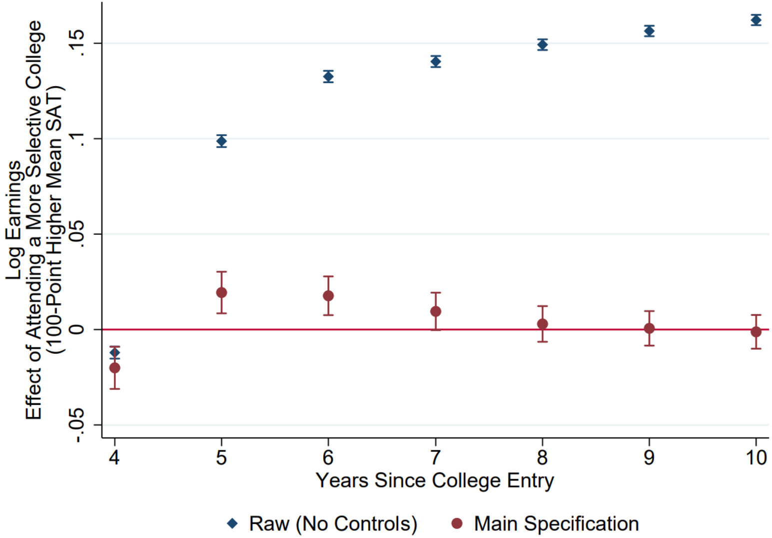

TAMU
Texas Tech
UT-Dallas
TAMU-Galveston
Houston
UT-Tyler
Texas State-San Marcos
UT-Arlington
UT-Permian Basin
Sam Houston State
Stephen F. Austin State
Tarleton State
Lamar
TAMU-Corpus Christi
North Texas
Texas Woman's
TAMU-Kingsville
West Texas A&M
UT-San Antonio
Midwestern State
Angelo State
TAMU-International
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

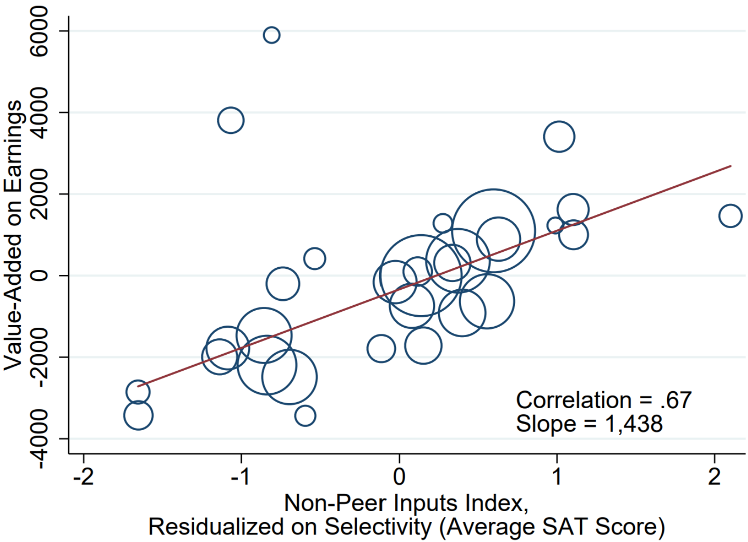

Raw Earnings
10000
0
-10000
-20000
-30000
800
900
1000
1100
1200
Average SAT Score of Incoming Students
TAMU
UT-Austin
Texas Tech
UT-Dallas
TAMU-Galveston
Houston
UT-Permian Basin
Sam Houston State
North Texas
Angelo State
UT-Pan American
UT-El Paso
Houston-Downtown
Sul Ross State
Prairie View A&M
Texas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



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

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

Learning and Experience

- Searching for job security and the consequences of job loss
  - JAROSCH, Gregor. Searching for job security and the consequences of job loss. Econometrica, 2023.
- Human Capital Depreciation and Returns to Experience
  - DINERSTEIN, Michael; MEGALOKONOMOU, Rigissa; YANNELIS, Constantine. Human capital depreciation and returns to experience. American Economic Review, 2022.
- Duration Dependence and Labor Market Conditions
  - KROFT, Kory; LANGE, Fabian; NOTOWIDIGDO, Matthew J. Duration dependence and labor market conditions: Evidence from a field experiment. The Quarterly Journal of Economics, 2013.

The Consequences of Job Loss. Jarosch, Gregor(2023)

- **Issue:** Job loss leads to significant earnings losses, challenging traditional labor market models.
- **Model:** Proposes a frictional labor market model where:
  - Jobs vary in unemployment risk.
- **Human Capital:** Accounts for effects of time spent out of work on skill depreciation.
- **Empirical Framework:** Estimated using German Social Security data.
- **Findings:**
  - Model captures wage, employment, and unemployment risk responses to job loss.
  - Interaction between job security and human capital evolution intensifies the "unemployment scar."
  - **Conclusion:** Improving job security and better employment search are crucial to mitigating job loss consequences.

Human Capital Depreciation and Returns to Experience. Dinerstein et al. (2022)

- **Topic:** Examines how unused skills lead to human capital depreciation.
- **Challenge:** Measuring skill depreciation is difficult due to:
  - Difficulty in quantifying worker skills.
  - Selection bias, as less productive workers tend to spend more time in nonemployment.
- **Methodology:**
  - Uses administrative data on teacher assignments and student outcomes in Greece.
  - Employs quasi-random variation from the teacher assignment process.
- **Findings:**
  - A one-year increase in time without formal employment reduces students' test scores by 0.05 standard deviations.
  - Estimated skill depreciation rate: 4.3
  - Estimated returns to experience: 6.8
  - **Conclusion:** Nonemployment periods lead to significant human capital depreciation, impacting productivity.

Duration Dependence and Labor Market Conditions. Kroft et al. (2013)

- **Research Focus:** Investigates employer behavior's role in creating "negative duration dependence" — the negative impact of longer unemployment spells.
- **Methodology:**
  - Sent fictitious résumés to real job postings in 100 U.S. cities.
  - Analyzed callback rates based on unemployment spell length.
- **Key Findings:**
  - Callback likelihood decreases significantly with longer unemployment spells.
  - Majority of the decline occurs within the first eight months of unemployment.
- **Conclusion:**
  - Results support screening models where employers use unemployment duration as a signal of productivity.

Conclusion

Thank you for your attention!

For any questions or suggestions, please contact:
valdemar.pinho@fgv.br

References

- Acemoglu, D., and Autor, D. (2011). "Skills, tasks, and technologies: implications for employment and earnings." *Handbook of Labor Economics*, Volume 4b.
- Angrist, J., Imbens, G., and Rubin, D. (1996). "Identification of causal effects using instrumental variables." *Journal of the American Statistical Association* 91(434).
- Angrist, J., and Krueger, A. (1991). "Does compulsory school attendance affect schooling and earnings?" *Quarterly Journal of Economics* 106(4).
- Angrist, J., and Pischke, J. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Calonico, S., Cattaneo, M., and Titiunik, R. (2014). "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6).
- Card, D. (1999). "The causal effect of education on earnings." *Handbook of Labor Economics*, Volume 3.
- Chetty, R., Friedman, J., Saez, E., Turner, N., and Yagan, D. (2020). "Income segregation and intergenerational mobility across colleges in the United States." *Quarterly Journal of Economics* 135(3).
- Clark, D., and Martorell, P. (2014). "The signaling value of a high school diploma." *Journal of Political Economy* 122(2).
- Dale, S., and Krueger, A. (2002). "Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables." *Quarterly Journal of Economics* 117(4).

References

- Dale, S., and Krueger, A. (2014). "Estimating the effects of college characteristics over the career using administrative earnings data." *Journal of Human Resources* 49(2).
- Goldin, C., and Katz, L. (2008). *The Race Between Education and Technology*. Harvard University Press.
- Holland, P. (1986). "Statistics and causal inference." *Journal of the American Statistical Association* 81(396).
- Imbens, G., and Angrist, J. (1994). "Identification and estimation of local average treatment effects." *Econometrica* 62(2).
- Imbens, G., and Kalyanaraman, K. (2011). "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies* 79(3).
- Imben, G., and Lemieux, T. (2008). "Regression discontinuity designs: a guide to practice." *Journal of Econometrics* 142(2).
- McCrary, J. (2008). "Manipulation of the running variable in the regression discontinuity design: a density test." *Journal of Econometrics* 142(2).
- Mountjoy, J., and Hickman, B. (2020). "The returns to college(s): estimating value-added and match effects in higher education." Working paper.
- Spence, M. (1973). "Job market signaling." *Quarterly Journal of Economics* 87(3).
- Zimmerman, S. (2014). "The returns to college admission for academically marginal students." *Journal of Labor Economics* 32(4).