# What Does IV Estimate?

# "Estimating Marginal Returns to Education"

Pedro Carneiro, James J. Heckman and Edward Vytlacil

*American Economic Review* 101(6): 2754–2871 (2011).

James J. Heckman

Econ 312, Spring 2023

THE UNIVERSITY OF CHICAGO

- Marginal treatment effect (MTE):

$$\text{MTE}(x, u_S) \equiv E(\beta \mid \boldsymbol{X} = x, U_S = u_S),$$

- Mean return to schooling for individuals with characteristics $\boldsymbol{X} = x$ and $U_S = u_S$.

$$\mathrm{IV}_k(x) \equiv \frac{\mathrm{Cov}(Y, Z^k|\boldsymbol{X} = x)}{\mathrm{Cov}(S, Z^k|\boldsymbol{X} = x)} = \int_0^1 \mathrm{MTE}(x, u_S) h_k(x, u_S)\, du_S, \quad (1)$$

where again the weights can be consistently estimated from sample data.

# Policy Relevant Treatment Effects and Marginal Policy Relevant Treatment Effects

- Let $S^*$ be the treatment choice that would be made after the policy change.
- Let $P^*$ be the corresponding probability that $S^* = 1$ after the policy change.
- $S^*$ is defined by $S^* = \mathbf{1}[P^* \geq U_S]$.
- Let $Y^* = S^* Y_1 + (1 - S^*) Y_0$ be the outcome under the alternative policy.

- Policy Relevant Treatment Effect (PRTE), defined when $E(S) \neq E(S^*)$ as

$$\frac{E(Y \mid \text{Alternative Policy}) - E(Y \mid \text{Baseline Policy})}{E(S^* \mid \text{Alternative Policy}) - E(S \mid \text{Baseline Policy})}$$
$$= \frac{E(Y^*) - E(Y)}{E(S^*) - E(S)} = \int_0^1 \text{MTE}(u_S)\, \omega_{\text{PRTE}}(u_S)\, du_S$$

where

$$\omega_{\text{PRTE}}(u_S) = \frac{F_P(u_S) - F_{P^*}(u_S)}{E_{F_{P^*}}(P) - E_{F_P}(P)}$$

where $F_{P^*}$ and $F_P$ are the distributions of $P^*$ and $P$, respectively

- Suppress $\boldsymbol{X}$ to simplify notation.

## Marginal MPRTE

- MPRTE:
- Consider a sequence of policies indexed by a scalar variable $\alpha$, with $\alpha = 0$ denoting the baseline, status quo policy.
- We associate with each policy $\alpha$ the corresponding fitted probability of schooling $P_\alpha$, where $P_0 = P(\boldsymbol{Z})$, the baseline propensity score.
- For each policy $\alpha$ we define the corresponding PRTE parameter for going from the baseline status quo to policy $\alpha$.
- We define the MPRTE as the limit of such a sequence of PRTEs as $\alpha$ goes to zero.

- We will consider the following examples of such sequences of policies:
  1. a policy that increases the probability of attending college by an amount $\alpha$, so that $P_\alpha = P_0 + \alpha$ and $F_\alpha(t) = F_0(t - \alpha)$;
  2. a policy that changes each person's probability of attending college by the proportion $(1 + \alpha)$, so that $P_\alpha = (1 + \alpha)P_0$ and $F_\alpha(t) = F_0(\frac{t}{1+\alpha})$; and
  3. a policy intervention that has an effect similar to a shift in one of the components of $\boldsymbol{Z}$, say $Z^{[k]}$, so that $Z_\alpha^{[k]} = Z^{[k]} + \alpha$ and $Z_\alpha^{[j]} = Z^{[j]}$ for $j \neq k$.
- For example, the $k$th element of $\boldsymbol{Z}$ might be college tuition, and the policy under consideration subsidizes college tuition by the fixed amount $\alpha$.
- In each of these three cases, we consider the corresponding PRTE for going from the status quo to policy $\alpha$, and consider the limit of such PRTEs as $\alpha$ goes to zero.

- These limits differ from IV estimates in general.
- Just as IV is a weighted average of the MTE, as in equation (1), there is a similar expression for average marginal policy changes that weights up the MTE by the proportion of persons induced to change by the policy.
- In general, the weights are different for IV and MPRTE.
- (Compare the IV weights in Table A-1B in the Appendix and the weights in Table 1 in the text.)

Table 1: Weights for MPRTE

| Measure of Distance for People Near the Margin | Definition of Policy Change | Weight |
|---|---|---|
| $\lvert \mu_S(\boldsymbol{Z}) - V \rvert < e$ | $Z^k_\alpha = Z^k + \alpha$ | $h_{MPRTE}(\boldsymbol{x}, u_S) = \frac{f_{P\mid\boldsymbol{X}}(u_S) f_{V\mid\boldsymbol{X}}(F^{-1}_{V\mid\boldsymbol{X}}(u_S))}{E(f_{V\mid\boldsymbol{X}}(\mu_S(\boldsymbol{Z}))\mid\boldsymbol{X})}$ |
| $\lvert P - U \rvert < e$ | $P_\alpha = P + \alpha$ | $h_{MPRTE}(\boldsymbol{x}, u_S) = f_{P\mid\boldsymbol{X}}(u_S)$ |
| $\left\lvert \frac{P}{U} - 1 \right\rvert < e$ | $P_\alpha = (1+\alpha) P$ | $h_{MPRTE}(\boldsymbol{x}, u_S) = \frac{u_S f_{P\mid\boldsymbol{X}}(u_S)}{E(P\mid\boldsymbol{X})}$ |

Source: Carneiro, Heckman, and Vytlacil (2010).

- The MPRTE is the appropriate parameter with which to conduct cost-benefit analysis of marginal policy changes.

## Average Marginal Treatment Effect

- The effect of a marginal policy change for a particular perturbation of $P(\boldsymbol{Z})$ is the same as the average effect of treatment for those who are arbitrarily close to being indifferent between treatment or not, using a metric $m(P, U_S)$ measuring the distance between $P(\boldsymbol{Z})$ and $U_S$.
- This parameter is defined as $AMTE = \lim_{e \to 0} E[Y_1 - Y_0 | m(P, U_S) \leq e]$.

- For the three examples of MPRTE previously discussed, the corresponding metrics defining the AMTE are respectively:
  1. $m(P, U_S) = \left| F_V^{-1}(P) - F_V^{-1}(U_S) \right| = |\mu_S(\boldsymbol{Z}) - V|$;
  2. $m(P, U_S) = |P - U_S|$;
  3. $m(P, U_S) = \left| \frac{P}{U_S} - 1 \right|$.
- Table 1 shows the different weights associated with the different definitions of the Average Marginal Treatment Effect and the associated MPRTE.

## Table 2: Definitions of the Variables Used in the Empirical Analysis

| Variable | Definition |
|---|---|
| $Y$ | Log Wage in 1991 (average of all non-missing wages between 1989 and 1993) |
| $S$=1 | If ever Enrolled in College by 1991: zero otherwise |
| $\boldsymbol{X}$ | AFQT,[A] Mother's Education, Number of Siblings, Average Log Earnings 1979-2000 in County of Residence at 17, Average Unemployment 1979-2000 in State of Residence at 17, Urban Residence at 14, Cohort Dummies, Years of Experience in 1991, Average Local Log Earnings in 1991, Local Unemployment in 1991. |
| $\boldsymbol{Z}\backslash\boldsymbol{X}$[B] | Presence of a College at Age 14 (Cameron and Taber, 2004; Card, 1993), Local Earnings at 17 (Cameron and Heckman, 1998; Cameron and Taber, 2004), Local Unemployment at 17 (Cameron and Heckman, 1998), Local Tuition in Public 4 Year Colleges at 17 (Kane and Rouse, 1995). |

Notes: [A]We use a measure of this score corrected for the effect of schooling attained by the participant at the date of the test, since at the date the test was taken, in 1981, different individuals have different amounts of schooling and the effect of schooling on AFQT scores is important. We use a correction based on the method developed in Hansen, Heckman, and Mullen (2004). We take the sample of white males, perform this correction and then standardize the AFQT to have mean 0 and variance 1 within this sample. See Table A-2. [B]The papers in parentheses are papers that previously used these instruments.

- Distance to college was first used as an instrument for schooling by Card (1995) and was subsequently used by Kane and Rouse (1995), Kling (2001), Currie and Moretti (2003) and Cameron and Taber (2004).
- Cameron and Taber (2004) and Carneiro and Heckman (2002) show that distance to college in the NLSY79 is correlated with a measure of ability (AFQT).
- In this paper, we include this measure of ability in the outcome equation.

- Cameron and Heckman (1998, 2001) and the papers they cite emphasize the importance of controlling for local labor market characteristics (see also Cameron and Taber, 2004).
- If local unemployment and local earnings at age 17 are correlated with the unobservables in the earnings equations in the adult years, our measures of local labor market conditions would not be valid instruments.
- To mitigate this concern, we have included measures of permanent local labor market conditions (which we define as the average earnings and unemployment between 1973 and 2000 for each location of residence at 17) both in the selection and outcome equations.

- Effectively, we only use the innovations in the local labor market variables as instruments.
- This is similar to the procedure used by Cameron and Taber (2004).
- Further, in the outcome equations we also include the average log earnings in the county of residence in 1991, and the average unemployment rate in the state of residence in 1991.

- Tuition is used to predict college attendance in Cameron and Heckman (1998, 2001) and Kane and Rouse (1995).
- We control for AFQT and maternal education in all of our models.
- These variables are likely to be highly correlated with college quality.
- We use these variables to account for any correlation between our measure of tuition (which corresponds only to 4 year public colleges) and college quality.
- In order to examine the sensitivity of our estimates to the choice of instruments, we estimate models with and without tuition.

# Table 3: College Decision Model - Average Marginal Derivatives

| | Average Derivative |
|---|---|
| CONTROLS ($\boldsymbol{X}$): | |
| Corrected AFQT | 0.2826 |
| | (0.0114)*** |
| Mother's Years of Schooling | 0.0441 |
| | (0.0059)*** |
| Number of Siblings | -0.0233 |
| | (0.0068)*** |
| Urban Residence at 14 | 0.0340 |
| | (0.0274) |
| "Permanent" Local Log Earnings at 17 | 0.1820 |
| | (0.0941)** |
| "Permanent" State Unemployment Rate at 17 | 0.0058 |
| | (0.0165) |
| INSTRUMENTS ($\boldsymbol{Z}$): | |
| Presence of a College at 14 | 0.0529 |
| | (0.0273)** |
| Local Log Earnings at 17 | -0.2687 |
| | (0.1008)*** |
| Local Unemployment Rate at 17 | 0.0149 |
| (in %) | (0.0100) |
| Tuition in 4 Year Public Colleges at 17 | -0.0027 |
| (in $100) | (0.0017)* |
| Test for joint significance of instruments: $p$-value | 0.0001 |

Table 4: Test of linearity of $E\left(Y|\boldsymbol{X}, P=p\right)$ using models with different orders of polynomials in $P^{\text{A}}$

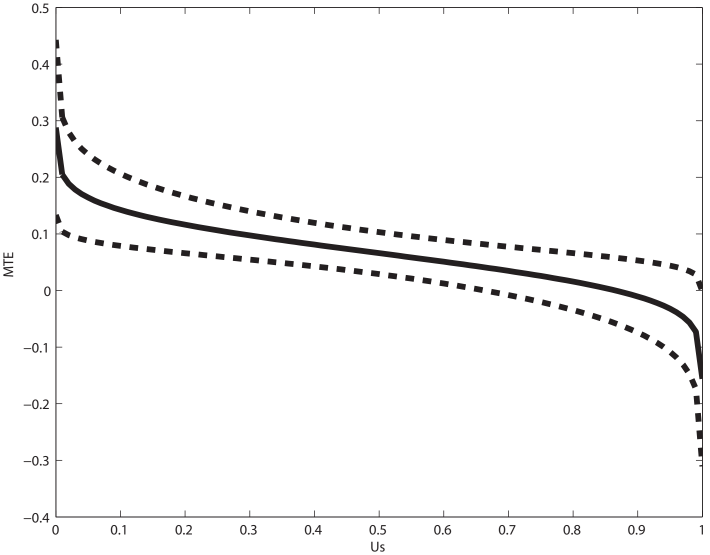
| | | | | |
|---|---|---|---|---|
| Degree of Polynomial for model | 2 | 3 | 4 | 5 |
| $p$-value of joint test of nonlinear terms | 0.035 | 0.049 | 0.086 | 0.122 |
| Adjusted critical value | | 0.057 | | |
| Outcome of test: | | Reject | | |

## Table 5: Test of equality of LATEs ($H_0 : LATE_1 \left(U_S^{1L}, U_S^{1H}\right) - LATE_1 \left(U_S^{2L}, U_S^{2H}\right) = 0$) - baseline model[B]

| | | | | | | |
|---|---|---|---|---|---|---|
| Ranges of $U_S$ for $\text{LATE}_1$<br>Ranges of $U_S$ for $\text{LATE}_2$ | (0,0.04)-<br>-(0.08,0.12) | (0.08,0.12)-<br>-(0.16,0.20) | (0.12,0.20)-<br>-(0.24,0.28) | (0.24,0.28)-<br>-(0.32,0.36) | (0.32,0.36)-<br>-(0.40,0.44) | (0.40,0.44)-<br>-(0.48,0.52) |
| Difference in LATEs | 0.0689 | 0.0629 | 0.0577 | 0.0531 | 0.0492 | 0.0459 |
| $p$-value | 0.0240 | 0.0280 | 0.0280 | 0.0320 | 0.0320 | 0.0520 |
| Ranges of $U_S$ for $\text{LATE}_1$<br>Ranges of $U_S$ for $\text{LATE}_2$ | (0.48,0.52)-<br>-(0.56,0.60) | (0.56,0.60)-<br>-(0.64,0.68) | (0.64,0.68)-<br>-(0.72,0.76) | (0.72,0.76)-<br>-(0.80,0.84) | (0.80,0.84)-<br>-(0.88,0.92) | (0.88,0.92)-<br>-(0.96,1) |
| Difference in LATEs | 0.0431 | 0.0408 | 0.0385 | 0.0364 | 0.0339 | 0.0311 |
| $p$-value | 0.0520 | 0.0760 | 0.0960 | 0.1320 | 0.1800 | 0.2400 |
| Joint $p$-value | 0.0520 | | | | | |

- Figure 1 plots the estimated MTE with 90% confidence bands, evaluated at mean values of $\boldsymbol{X}$ (we obtain annualized estimates of the returns to college by dividing the MTE by 4, which is the average difference in years of schooling for those with $S = 1$ and those with $S = 0$).

# Figure 1: MTE estimated from a normal selection model



- The people with the highest high gross returns are more likely to go to college (have low $U_S$).
- Individuals choose the schooling sector in which they have comparative advantage.

Notes: To estimate the function plotted here we estimate a parametric normal selection model by maximum likelihood. The figure is computed using the following formula:

$$\Delta^{\mathrm{MTE}}(x, u_S) = \mu_1(x) - \mu_0(x) - (\sigma_{1V} - \sigma_{0V})\,\Phi^{-1}(u_S)$$

## Table 6: Returns to a Year of College

| Model | | Normal | Semi-Parametric |
|---|---|---|---|
| ATE= $E(\beta)$ | | 0.0670 | Not Identified |
| | | (0.0378) | |
| TT= $E(\beta \mid S=1)$ | | 0.1433 | Not Identified |
| | | (0.0346) | |
| TUT= $E(\beta \mid S=0)$ | | -0.0066 | Not Identified |
| | | (0.0707) | |
| MPRTE | | | |
| Policy Perturbation | Metric | | |
| $Z_\alpha^k = Z^k + \alpha$ | $\lvert Z\gamma - V \rvert < e$ | 0.0662 | 0.0802 |
| | | (0.0373) | (0.0424) |
| $P_\alpha = P + \alpha$ | $\lvert P - U \rvert < e$ | 0.0637 | 0.0865 |
| | | (0.0379) | (0.0455) |
| $P_\alpha = (1+\alpha)P$ | $\left\lvert \frac{P}{U} - 1 \right\rvert < e$ | 0.0363 | 0.0148 |
| | | (0.0569) | (0.0589) |
| Linear IV (Using $P(\boldsymbol{Z})$ as the instrument) | | 0.0951 | |
| | | (0.0386) | |
| OLS | | 0.0836 | |
| | | (0.0068) | |

# Estimating the MTE and Marginal Policy Effects using Local Instrumental Variables

Figure 2: Support of $P$ conditional on $\boldsymbol{X}$

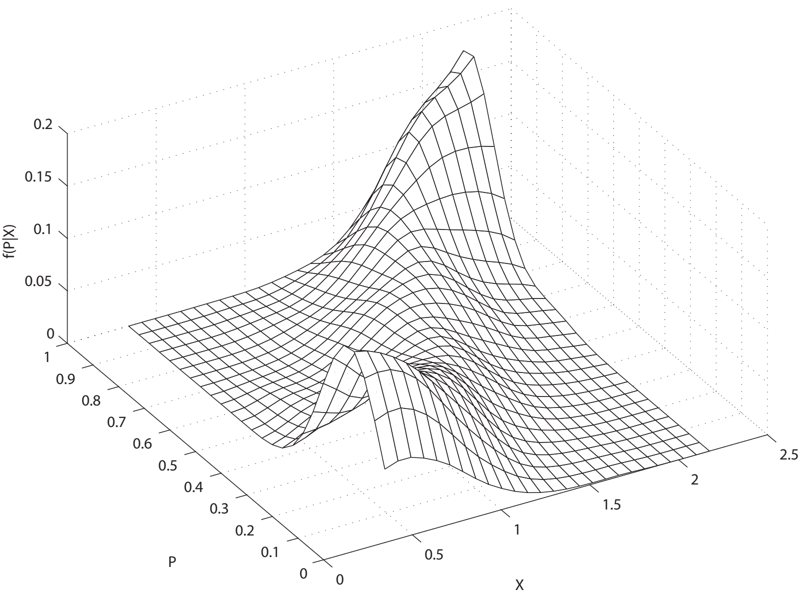
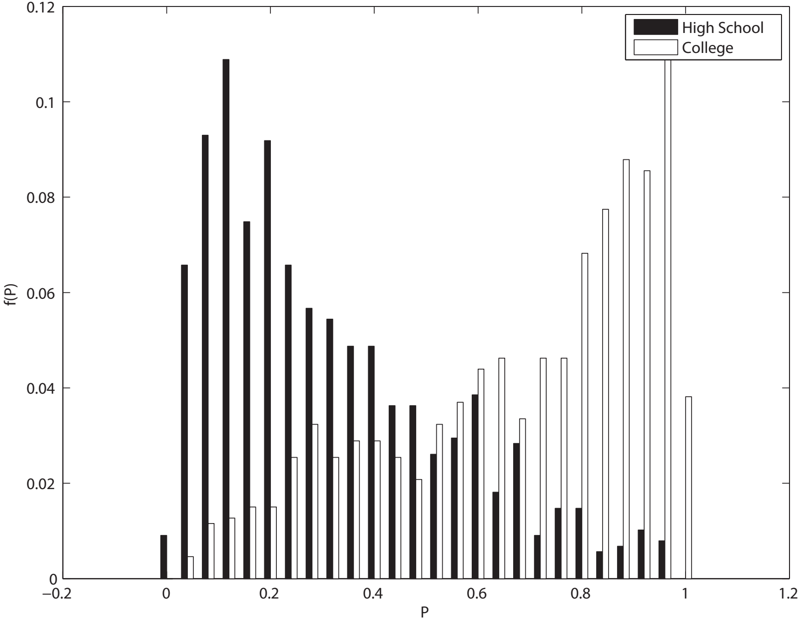
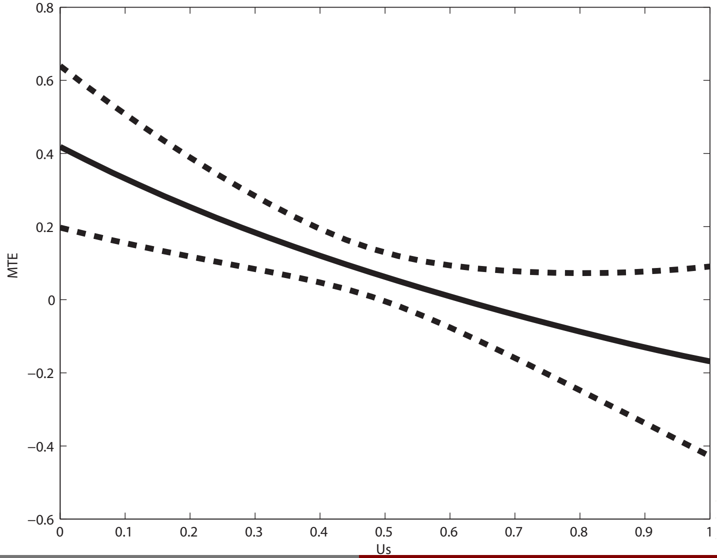


Figure 3: Support of $P$ for $S = 0$ and $S = 1$



- Figure 4 plots the component of the MTE that depends on $U_S$, with 90% confidence bands computed from the bootstrap.
- We fix the components of $\boldsymbol{X}$ at their mean values in the sample.
- We annualize the MTE.
- Our estimates show that, in agreement with the normal model, $E(U_1 - U_0 \mid U_S = u_S)$ is declining in $u_S$, i.e., that students with high values of $U_S$ have lower returns than those with low values of $U_S$.

# Figure 4: $E(Y_1 - Y_0|\boldsymbol{X}, U_S)$ with 90 percent confidence interval–locally quadratic regression estimates



- These are local average treatment effects (LATEs) for different sections of the MTE.
- We compare adjacent LATEs.
- Table 7 reports the outcome of these comparisons.
- For example, the first column reports that

$$\begin{aligned} & E\left(Y_1 - Y_0 \mid \boldsymbol{X} = \overline{x}, 0 \leq U_S \leq 0.04\right) \\ & \quad - E\left(Y_1 - Y_0 | \boldsymbol{X} = \overline{x}, 0.08 \leq U_S \leq 0.12\right) = 0.0689. \end{aligned}$$

Table 7: Test of equality of LATEs
$(H_0 : LATE_1\left(U_S^{1L}, U_S^{1H}\right) - LATE_1\left(U_S^{2L}, U_S^{2H}\right) = 0)$ - baseline model[B]

| | | | | | | |
|---|---|---|---|---|---|---|
| Ranges of $U_S$ for $LATE_1$<br>Ranges of $U_S$ for $LATE_2$ | (0,0.04)-<br>-(0.08,0.12) | (0.08,0.12)-<br>-(0.16,0.20) | (0.12,0.20)-<br>-(0.24,0.28) | (0.24,0.28)-<br>-(0.32,0.36) | (0.32,0.36)-<br>-(0.40,0.44) | (0.40,0.44)-<br>-(0.48,0.52) |
| Difference in LATEs | 0.0689 | 0.0629 | 0.0577 | 0.0531 | 0.0492 | 0.0459 |
| $p$-value | 0.0240 | 0.0280 | 0.0280 | 0.0320 | 0.0320 | 0.0520 |
| Ranges of $U_S$ for $LATE_1$<br>Ranges of $U_S$ for $LATE_2$ | (0.48,0.52)-<br>-(0.56,0.60) | (0.56,0.60)-<br>-(0.64,0.68) | (0.64,0.68)-<br>-(0.72,0.76) | (0.72,0.76)-<br>-(0.80,0.84) | (0.80,0.84)-<br>-(0.88,0.92) | (0.88,0.92)-<br>-(0.96,1) |
| Difference in LATEs | 0.0431 | 0.0408 | 0.0385 | 0.0364 | 0.0339 | 0.0311 |
| $p$-value | 0.0520 | 0.0760 | 0.0960 | 0.1320 | 0.1800 | 0.2400 |
| Joint $p$-value | 0.0520 | | | | | |

- The *p*-value of the test of the hypothesis that this difference is equal to zero is reported below this number and is 0.0240, which implies that we reject this hypothesis at conventional levels of significance.

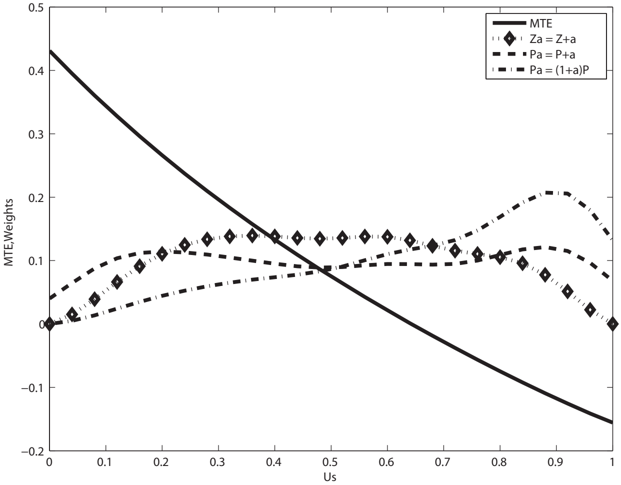
- This table shows that the slope of the MTE is negative and statistically significant (at a 10% level of significance) for values of $U_S$ up to 0.76 ($p$-values are reported in the bottom row of the table), and it remains negative but statistically insignificant after that.
- This is further evidence that individuals select into college based on heterogeneous returns in realized outcomes, although the rejection is only strong in the left tail of the estimated MTEs.
- A joint test that the difference across all adjacent LATEs is different from zero has a $p$-value of 0.0520.

- $P$ only has support between 0.0324 and 0.9775, and thus it is not possible to estimate parameters which require full support such as $E(\beta)$, $E(\beta \mid S = 1)$ and $E(\beta \mid S = 0)$.
- Estimation of such parameters is possible in the normal model only because of its parametric assumptions.
- Analysts often define such parameters as the objects of interest, even though they are very hard to identify, and even though they are often not economically interesting.

- In contrast, the MPRTE parameter not only answers interesting economic questions about the marginal gains of specific policies, but it is also identified without strong support assumptions since it only requires estimating the MTE within the support of the data (see Carneiro, Heckman, and Vytlacil, 2010).
- The second column of Table 6 presents estimates of three different versions of the MPRTE where the policy considered is either a marginal change in tuition or a marginal change in $P$.
- The estimates are obtained in the following way.

- Figure 5 graphs the weights on $E(Y_1 - Y_0|\boldsymbol{X}, U_S = u_S)$ for the three MPRTE parameters with estimates reported in Table 6, all evaluated at the mean of $\boldsymbol{X}$.
- While the MPRTE weights for the first two policies ($Z_\alpha^k = Z^k + \alpha$ and $P_\alpha = P + \alpha$) weight mainly the middle section of the MTE, the third policy ($P_\alpha = (1 + \alpha)P$) overweights individuals with high levels of $U_S$ because its effect on enrollment is larger for those with already high levels of $P$.

# Figure 5: Weights for Three Different Versions of the MPRTE

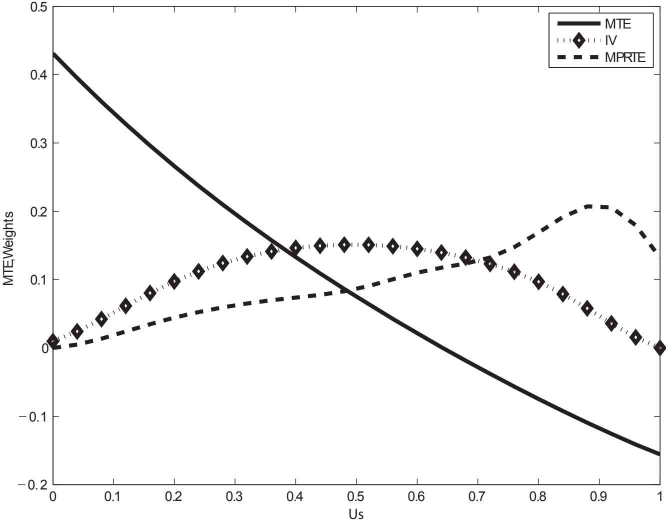



Notes: The scale of the y-axis is the scale of the MTE, not the scale of the weights, which are scaled to fit the picture.

- The IV estimate, presented in the second to the last row of Table 6, is 0.0951.
- We use $P(\boldsymbol{Z})$ as the instrument, but it is possible to construct IV estimates for other combinations of instruments.
- IV does not correspond to any of the MPRTE parameters that we consider, and it is particularly far from MPRTE in the case of the third policy.
- Figure A-1 in the Web Appendix shows the sharp difference in the MTE weights for IV/LATE and the third of the MPRTE parameters, evaluated at mean $\boldsymbol{X}$.

- Notice that both MPRTE and LATE correspond to some marginal effect (see Imbens (2010), for arguments for using LATE).
- However, LATE only estimates the policy effect **of interest** if the instrument variation corresponds exactly to the policy variation.
- For a specific policy of interest LATE can be wildly off the mark.
- Figure 6 plots the IV weights (using $P$ as the instrument) together with the MPRTE weights for the marginal policy defined by $P_\alpha = (1 + \alpha)P$.
- These two weights are dramatically different from each other.

# Figure 6: Weights for IV and MPRTE

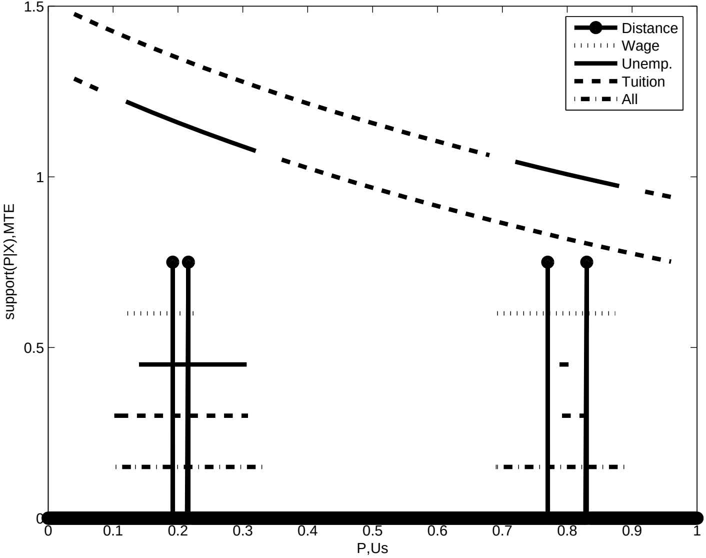



Notes: The scale of the $y$-axis is the scale of the MTE, not the scale of the weights, which are scaled to fit the picture.

- One benefit of this approach over the LATE approach is that it enables us to determine what portion of the MTE each instrument in $\boldsymbol{Z} \setminus \boldsymbol{X}$ identifies, i.e., it enables analysts to identify the quantiles of $U_S$ that each instrument traces out.
- Thus in our approach, the margin traced out by variation in each instrument is clearly identified.
- In the LATE approach that does not specify an explicit choice equation, the margin identified by variation in an instrument is not clearly defined.

- Figure 7 shows the support of $P(\boldsymbol{Z})$ when we fix the variables in $\boldsymbol{X}$ at two different values and vary the instruments one at a time.

# Figure 7: Support of $P$ for fixed $X$



- This is the approach required to secure estimates if we condition on $\boldsymbol{X}$ and do not invoke independence between $\boldsymbol{X}$ and $(U_0, U_1, V)$.
- This exercise informs us about what margin each instrument identifies under more general conditions.
- It also shows how far we expand the support of $P(\boldsymbol{Z})$ (and therefore, the support of $U_S$ over which we can estimate the MTE) by using multiple instruments simultaneously, as opposed to using them one at a time.

- There are two curves in the picture corresponding to the MTE evaluated at different values of $\boldsymbol{X}$.
- The two lines correspond to the 25th (bottom) and 75th (top) percentiles of the distribution of $\boldsymbol{X}(\delta_1 - \delta_0)$.
- The curves have dashed and solid segments.
- The solid segment represents the portion of the MTE we can identify at each value of $\boldsymbol{X}$ if we do not invoke independence between $\boldsymbol{X}$ and $(U_0, U_1, V)$.
- To be precise, we find values of $\boldsymbol{X}$ for which $\boldsymbol{X}(\delta_1 - \delta_0)$ is at the 25th percentile of its distribution (we pick values of this index between the 24th and 26th percentile of its distribution and compute mean $\boldsymbol{X}$ in this interval).

- Then we vary the instruments within this range and we trace out the corresponding support of $P$ for $\boldsymbol{X}$ fixed at this value.
- Then we take our estimate of the MTE and select only the segment which is contained within the support of $P$ for fixed $\boldsymbol{X}$.
- We do the same for values of $\boldsymbol{X}$ for which $\boldsymbol{X}(\delta_1 - \delta_0)$ is at the 75th percentile of its distribution.
- The dashed segments in each curve correspond to the additional portions of the MTE that we identify if instead we assume independence between $\boldsymbol{X}$ and $(U_0, U_1, V)$.

- It is informative to know not only what section of the MTE is identified at each value of $\boldsymbol{X}$, but also what section of the MTE is identified by varying each instrument at a time.
- To generate the graph labeled "Distance," we not only fix $\boldsymbol{X}$ at the two values referred to above, but we also fix all the other instruments at the corresponding mean values for each of the two percentiles of the distribution of $\boldsymbol{X}\,(\delta_1 - \delta_0)$ that we consider.
- Because the distance variable only takes two values for each $\boldsymbol{X}$, the support of $P(\boldsymbol{Z})$ in this case only has two points.

- The line labeled "Wage" corresponds to the support of $P(\boldsymbol{Z})$ we obtain when all variables except local wage at 17 are kept at their mean values (conditional on a given percentile of $\boldsymbol{X}(\delta_1 - \delta_0)$), the line labeled "Unemp." is generated by varying only local unemployment at 17, and the line labeled "Tuition" is generated by varying only local tuition at 17.
- Finally, the line labeled "All" is the support of $P(\boldsymbol{Z})$ when all the instruments are allowed to vary and the variables in $\boldsymbol{X}$ are fixed.

- Each instrument has different support, so if we were to use each instrument in isolation at mean $\boldsymbol{X}$ we would only be able to identify a small section of the MTE.
- When we allow all the instruments to vary simultaneously we get larger support for the MTE, but it is still not close to the full unit interval.
- This analysis makes clear which instruments contribute to identifying which portions of the MTE.