Roy Models and LATE

Ivan A. Canay
Northwestern University[1]
ECON 481-3

1. Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. Econometrica, 62(2), 467-475.
2. Angrist, J. D., & Evans, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. The American Economic Review, 88(3), 450-477.

[1] Today's class is based on Alex Torgovitsky's notes. I'd like to thank him for kindly sharing them.

Topics of Part I

- Lec I: Selection on Observables
    1. Potential Outcomes vs Latent Variables
    2. Causal Inference
    3. Selection Bias
    4. Selection on Observables & Selection on Prop. Score
- Lec II: Roy Models and LATE
    1. The role of heterogeneity
    2. Multiple instruments, Covariates, and Abadie's κ
- Lec III: Marginal Treatment Effect
    1. Parameters as functions of MTEs
    2. Policy Relevant Treatment Effects
- Lec IV: Extrapolations
    1. Semi-Parametrics MTEs
    2. Weights for Target Parameters

Outline

1. **Linear IV and Heterogeneity**
2. Roy Models: parametric approach
3. LATE
4. Abadie's κ
5. Empirical Application: Angrist and Evans (98)

Instrumental Variables

- Selection on observables can be difficult to justify in economics
- Unobserved confounders tend to be the rule, not the exception
- **Instrumental variable** (IV) strategies provide an alternative

The two properties of an instrument

1. **Exogenous** — unrelated (in some sense) with potential outcomes
2. **Relevant** — related (in some sense) to treatment states

- The "in some sense" may change in different types of IV models
- But the basic idea is always the same:
  Exogenous variation in the IV changes the (endogenous) treatment
  Resulting changes in the outcome reflect only the treatment
- The first part requires relevance, so that treatment actually changes
- The second part requires exogeneity, so that nothing else changes

Linear IV Models & Heterogeneity

- You have seen linear IV models in previous courses:
  1. Exactly identified ("simple") IV
  2. Overidentified IV and the two stage least squares estimator (TSLS)
  3. Overidentified IV and (optimal) generalized method of moments (GMM)
- These models take the form:

$$\underbrace{Y}_{\text{outcome}} = \underbrace{D'\beta}_{\text{endogenous treatment}} + \underbrace{X'\gamma}_{\text{controls}} + U$$

- Unlike OLS, always intended as causal — **never descriptive**
- As such, they place strong assumptions on the potential outcomes:

$$Y_d - Y_{\widetilde{d}} = (d - \widetilde{d})'\beta$$

no treatment effect heterogeneity!

- Unobserved heterogeneity (random β) complicates things immensely
- Intuitively, we must consider *who* is affected by the instrument

IV and Potential Outcomes

Potential outcomes

- As before, we work with potential outcomes $\{Y_d : d \in \mathcal{D}\}$
- Sometimes, we also consider potential outcomes for D: $\{D_z : z \in \mathcal{Z}\}$
  D_z is what *would have* been chosen had Z been set to z
- Same relationship as for outcomes:

$$D = \sum_{z \in \mathcal{Z}} D_z I\{Z = z\}$$

Forms of exogeneity

- Exogeneity requires Z to be independent of $\{Y_d : d \in \mathcal{D}\}$ in some sense
  Could be mean independence, full independence, uncorrelated, etc.
- Stronger exogeneity conditions also require Z independent of $\{D_z : z \in \mathcal{Z}\}$
- Usually this would be joint, e.g. $Z \perp\!\!\!\perp (\{Y_d : d \in \mathcal{D}\}, \{D_z : z \in \mathcal{Z}\})$

Outline

1. Linear IV and Heterogeneity
2. **Roy Models: parametric approach**
3. LATE
4. Abadie's κ
5. Empirical Application: Angrist and Evans (98)

Roy Models

Latent variable models

- An alternative to potential outcomes for D is a latent variable model
- The leading case is binary with a separable latent variable:

$$D = I\left\{ \underbrace{U}_{\text{latent variable}} \leq \underbrace{\nu(X,Z)}_{\text{unknown function}} \right\}$$

- Combined with $Y = Y_1 D + Y_0(1-D)$, this is called the **Roy Model**
- Refers to Roy (1951), although the name is arguably a bit of a misnomer

- Apply the usual translation: $D_z = I\{U \leq \nu(X,z)\}$
- Some advantages — explicitly models D as a choice problem
  $\nu(X,Z) - U$ is the utility of $D = 1$ vs. $D = 0$
- Economists often feel (more) comfortable with modeling choice behavior

Roy Models and Heterogeneity

- A common version of the Roy model:

$$Y_0 = X'\beta_0 + V_0 \qquad \underbrace{D = I\{U \leq W'\gamma\}}_{\text{(selection equation)}}$$
$$Y_1 = X'\beta_1 + V_1$$

where (V_0, V_1, U) are unobservable and $W \equiv (X, Z)$ are observable

- This model allows for both observed and unobserved heterogeneity:

$$Y_1 - Y_0 = \underbrace{X'(\beta_1 - \beta_0)}_{\text{observed}} + \underbrace{V_1 - V_0}_{\text{unobserved}}$$

- Implies a random coefficient specification for the observed outcome:

$$Y = DY_1 + (1 - D)Y_0 = \underbrace{(V_1 - V_0)}_{\text{random coeff.}} D + X'\beta_0 + DX'(\beta_1 - \beta_0) + V_0$$

- Selection on unobservables if U and (V_0, V_1) are dependent

Parametric Roy Models

Likelihood approaches

- Classic Roy model assumes (V_0, V_1, U) is normal, independent of (X, Z)
- Fully parametric: Point identification by well-parameterization
  Requires (X, Z) to be not perfectly collinear, as usual
- Straightforward to derive the likelihood
- Today it is also straightforward to maximize the likelihood
- However in the 1970s and 1980s it was difficult (so I'm told)

The Heckman two-step approach

- Heckman (1976) used the parametric structure to derive regressions
- These establish point identification through explicit construction
- Less efficient (statistically), but more intuitive, easier to compute
- The 1976 paper is a bit obscure — "missing data" was more standard
  But notice that causal inference *is* a missing data problem

The Heckman Two-Step Argument

- The parametric assumptions imply:

$$\begin{aligned} E[Y|W, D=1] &= X'\beta_1 + E[V_1|W, W'\gamma \geq U] \\ &= X'\beta_1 - \text{cov}(V_1, U)\lambda(W'\gamma) \end{aligned}$$

- The second equality is a property of bivariate normals:

$$\lambda(W'\gamma) \equiv \frac{\phi(W'\gamma)}{\Phi(W'\gamma)} \text{ is called the (inverse) } \textbf{Mills ratio}$$

- $\lambda(W'\gamma)$ enters as an additional regressor that controls for selection
- $\text{cov}(V_1, U)$ is an unknown parameter to be estimated
- The derivation for $D = 0$ is symmetric:

$$E[Y|W, D=0] = X'\beta_0 + \text{cov}(V_0, U)\lambda(-W'\gamma)$$

- Others versions of this idea are called **control function** approaches

Identification: Heckman Two-Step

Identification argument

- Under the parametric assumptions, the selection equation is just a probit
- γ is point identified, and hence $\lambda(W'\gamma)$ is as well
- Then identify β_d and $\text{cov}(V_d, U)$ from linear regression
- So we need X to be not perfectly collinear with $\lambda(W'\gamma)$
- Regress Y on X and $-\lambda(W'\gamma)$ among $D = 1$ to identify $\beta_1, \text{cov}(V_1, U)$
- Regress Y on X and $\lambda(-W'\gamma)$ among $D = 0$ to identify $\beta_0, \text{cov}(V_0, U)$

Point identification with an instrument

- Recall that $W \equiv (X, Z)$
- Suppose that Z helps predict D after accounting for X
- Then the component of γ corresponding to Z is non-zero in the probit
- If this is true, then X and $\lambda(W'\gamma)$ are not perfectly collinear
- This is an example of a sufficient condition for relevance

Parametric Identification

Point identification without an instrument

- Suppose, however, there is no Z, so that $\lambda(W'\gamma) = \lambda(X'\gamma)$
- $\lambda(X'\gamma)$ will still not be perfectly collinear with X — $\lambda(\cdot)$ is nonlinear
  The fully parametric model is identified *even without any instruments*
- Although $\lambda(\cdot)$ is not very nonlinear, so can be nearly collinear in practice

The credibility of parametric identification

- This is a concerning property of this model
- Exposes the reliance of identification on the assumed parameterization
- No strong reason (except mathematical convenience) to choose normality
- So should be especially concerning that identification uses normality
- Other distributions can be used, but lead to the same issue
- Fully parametric Roy models are rarely used in top-flight research today
  When they are, they almost always have excluded instruments

Outline

1. Linear IV and Heterogeneity
2. Roy Models: parametric approach
3. **LATE**
4. Abadie's κ
5. Empirical Application: Angrist and Evans (98)

What Does Linear IV Estimate?

- Textbook linear IV models impose constant treatment effects (given X)
- Question: What is the IV estimand with unobserved heterogeneity?
- An **estimand** is the population quantity that an estimator estimates

A simple extension of the linear IV model

- Consider the random coefficients linear IV model

$$
\begin{aligned}
Y &= \alpha + BD + U && \alpha, \pi \text{ are constant} \\
D &= \pi + CZ + V && B, C, U, V \text{ are unobservable random variables}
\end{aligned}
$$

- Assume exogeneity: $Z \perp\!\!\!\perp (U, V, B, C)$ — stronger than usual
- Assume relevance: $\text{cov}(D, Z) \neq 0$
- The linear IV (slope) estimand is given by

$$
\beta_{IV} \equiv \frac{\text{cov}(Y, Z)}{\text{cov}(D, Z)}
$$

IV Estimand: Random Coeff. Model

- When B is constant, we know that $\beta_{IV} = B$
- More generally, one can show that

$$\beta_{IV} \equiv \frac{\text{cov}(Y,Z)}{\text{cov}(D,Z)} = E\left[\frac{C}{E(C)}B\right]$$

- β_{IV} is a **weighted average** of the causal effect of D on Y (i.e. of B)
- Agents more strongly impacted by Z (larger $|C|$) get more weight

- This immediately raises some concerns on interpretation
- For example, consider common instruments in the returns to schooling:
  1. Distance from college
  2. Quarter of birth (compulsory schooling laws)
  3. Tuition subsidies
- β_{IV} overweights returns to agents most affected by these instruments
  Unlikely to be representative of the overall population
- Another concern — weights can be negative for those with $C \times E[C] < 0$

Imbens and Angrist (1994)

Overview

- Provided additional conditions under which β_{IV} is easier to interpret
- Highly influential paper for both empirical and theoretical work
- Highly controversial and frequently misinterpreted or misunderstood
- We will look at results first, then discuss interpretations and controversy

Setup

- Binary treatment $D \in \{0,1\}$ and binary instrument $Z \in \{0,1\}$
- Results partially extend to multiple values of D and/or Z, with caveats
- Covariates X are conditioned on nonparametrically and implicitly

Assumptions

1. Exogeneity: $Z \perp\!\!\!\perp (Y_0, Y_1, D_0, D_1)$
2. Relevance: $\text{cov}(D, Z) \neq 0$
3. "**Monotonicity**": $D_1 \geq D_0$ a.s. — a new condition

Local Average Treatment Effect

- Since both D and Z are binary, one has

$$\beta_{IV} \equiv \frac{\text{cov}(Y,Z)}{\text{cov}(D,Z)} = \underbrace{\frac{E[Y|Z=1]-E[Y|Z=0]}{E[D|Z=1]-E[D|Z=0]}}_{\textbf{Wald estimand}}$$

- The main result is the following (shown in 480):

$$\beta_{IV} = \underbrace{E[Y_1 - Y_0}_{\text{average treatment effect}}\underbrace{|D_1 = 1, D_0 = 0]}_{\text{local}} \equiv \text{LATE}$$

Interpretation of the LATE

- Average treatment effect for the **compliers** $= [D_1 = 1, D_0 = 0]$
- Name motivated by randomized experiment with partial compliance
- Subgroup whose treatment is affected by the instrument
- The **defiers** are $[D_1 = 0, D_0 = 1]$, so monotonicity $\Leftrightarrow$ "no defiers"

One-Sided Non-Compliance

- Suppose that only agents with $Z = 1$ can have $D = 1$
- **One-sided non-compliance**: $[Z = 1, D = 0]$ ok, but not $[Z = 0, D = 1]$
- Leading example is an experiment with $D = 1$ unavailable elsewhere

- Under one-sided non-compliance we know $D_0 = 0$ always, hence

$$\underbrace{[D = 1]}_{\text{treated}} = \underbrace{[D_0 = 1, D_1 = 1]}_{\text{probability 0}} + \underbrace{[D_0 = 0, D_1 = 1]}_{\text{compliers}}$$

- So in this case, LATE $=$ ATT, and hence $\beta_{IV} =$ ATT

- Opposite case is when *all* agents with $Z = 1$ have $D = 1$ ($D_1 = 1$):

$$\underbrace{[D = 0]}_{\text{untreated}} = \underbrace{[D_0 = 0, D_1 = 1]}_{\text{compliers}} + \underbrace{[D_0 = 0, D_1 = 0]}_{\text{probability 0}}$$

- So here LATE $=$ ATU, and hence $\beta_{IV} =$ ATU

Dist. of Outcomes for Compliers

Choice types

- **compliers:** $[D_1 > D_0] = [D_1 = 1, D_0 = 0]$
- **defiers:** $[D_0 > D_1]$ — probability 0 under monotonicity
- **always-takers:** $[D_0 = D_1 = 1]$
- **never-takers:** $[D_0 = D_1 = 0]$
- Let $T \in \{a, n, c\}$ denote a choice type

- Complier outcome distributions are point identified:

$$F_{Y_0|T}(y|c) = F_{Y|DZ}(y|0,0)\frac{P[T \in \{c,n\}]}{P[T = c]} - F_{Y|DZ}(y|0,1)\frac{P[T = n]}{P[T = c]}$$
$$F_{Y_1|T}(y|c) = F_{Y|DZ}(y|1,1)\frac{P[T \in \{c,a\}]}{P[T = c]} - F_{Y|DZ}(y|1,0)\frac{P[T = a]}{P[T = c]}$$

- The intuition is that we can difference out the always (or never) takers:

$$[D = 1, Z = 1] \setminus [D = 1, Z = 0] = [T \in \{a, c\}] \setminus [T \in \{a\}] = [T \in \{c\}]$$

Extension to Multiple Instruments

Adjusted setup and assumptions

- Suppose $Z \in \{0, 1, \ldots, L\}$ with $L \geq 2$ — multiple discrete instruments
- Adjust **exogeneity** to be: $Z \perp\!\!\!\perp (D_0, D_1, \ldots, D_L, Y_0, Y_1)$
- Adjust **monotonicity** to be: $D_{z'} \geq D_z$ (or conversely) for all $z' \geq z$

Ordering the instruments

- Notice that the direction of monotonicity is point identified
- Let $p(z) \equiv P\{D = 1 | Z = z\}$ — our old friend the **propensity score**
- By the exogeneity condition, we know that $p(z) = P\{D_z = 1\}$
- Under monotonicity, $D_{z'} \geq D_z$ if and only if $p(z') \geq p(z)$
- So we can relabel Z such that D_z and $p(z)$ are increasing:

$$\begin{array}{ccccc} D_0 & \leq D_1 & \leq D_2 & \leq \cdots \leq & D_L \\ p(z_0) & \leq p(z_1) & \leq p(z_2) & \leq \cdots \leq & p(z_L) \end{array}$$

TSLS Estimand: Multiple Instruments

- Consider the TSLS estimator using $\{I\{Z = z\} : z = 0, \ldots, L\}$ as instruments
- The population first stage coefficient on $I\{Z = z_l\}$ is $p(z_l)$
- Hence, the slope coefficient estimand on D is given by

$$\beta_{TSLS} \equiv \frac{\mathrm{cov}(Y, p(Z))}{\mathrm{cov}(D, p(Z))}$$

- The following is derived in the problem set:

$$\beta_{TSLS} = \sum_{m=1}^{L} \lambda_m \text{LATE}_{m-1}^{m} \quad \text{where} \quad \text{LATE}_{m-1}^{m} \equiv E[Y_1 - Y_0 | D_m = 1, D_{m-1} = 0]$$

$$\text{and} \quad \lambda_m \equiv \frac{[p(z_m) - p(z_{m-1})] \left(\sum_{l=m}^{L} (p(z_l) - E[p(Z)]) P[Z = z_l] \right)}{\sum_{n=1}^{L} (p(z_n) - p(z_{n-1})) \left(\sum_{l=n}^{L} (p(z_l) - E[p(Z)]) P[Z = z_l] \right)}.$$

- So the TSLS estimand is a weighted average of pairwise LATEs
- The pairwise LATEs represent $Z = m - 1$ to $Z = m$ compliers
- The weights are positive (this is important) and sum to 1

Covariates — Nonparametric

- Often, one wants covariates X to help justify the exogeneity of Z
- And/or to reduce residual noise in Y
- And/or to look at observed heterogeneity in treatment effects

Adjust the assumptions to be conditional on X

1. Exogeneity: $(Y_0, Y_1, D_0, D_1) \perp\!\!\!\perp Z | X$
2. Relevance: $P\{D = 1 | X, Z = 1\} \neq P\{D = 1 | X, Z = 0\}$ a.s.
3. Monotonicity: $P\{D_1 \geq D_0 | X\} = 1$ a.s.
4. Overlap: $P\{Z = 1 | X\} \in (0, 1)$ a.s.

- Same argument point identifies $E[Y_1 - Y_0 | T = c, X = x] \equiv \text{LATE}(x)$
- In addition, one could aggregate these into:

$$E[Y_1 - Y_0 | T = c] = E\left[\frac{\text{LATE}(X) P[T = c | X]}{P[T = c]}\right]$$

Covariates — TSLS

- Usual curse of dimensionality from conditioning on $X = x$
- A more standard procedure with covariates is IV/TSLS
- Can these estimands be interpreted as (positively-weighted) LATEs?
- In general, the **answer is no**

A specific case where the answer is yes

- Suppose that X is discrete and write it as a set of binary indicators
- Included instruments: X — a full set of dummies
- Excluded instruments: X and XZ — suppose $Z \in \{0, 1\}$ for simplicity
- Not nonparametric (fully saturated), since no XD in the outcome
- Then the TSLS estimand of coefficient on D is:

$$\beta_{TSLS} = E\left[\text{LATE}(X) \frac{\text{var}(p(X,Z)|X)}{E[\text{var}(p(X,Z)|X)]}\right]$$

- Weighted average of LATEs – more weight the more residual Z variation

Interpreting Estimands

IV

- The IV estimand in the binary D, binary Z case is the LATE
- This parameter is easy to interpret as the average effect for compliers
- It could be quite relevant for a policy intervention that affects compliers

TSLS

- In contrast, the TSLS estimand is a mess, even in specialized cases
- A weighted average of several different complier groups
- When would these weights be useful to inform a counterfactual?

Reverse engineering

- These results are motivated by a backward (literally) thought process
- Start with a common estimator, then interpret the estimand
- Why not *start* with a parameter of interest and then create an estimator?

Outline

1. Linear IV and Heterogeneity
2. Roy Models: parametric approach
3. LATE
4. **Abadie's κ**
5. Empirical Application: Angrist and Evans (98)

Abadie's (2003) κ

- For covariates (but D, Z binary) there exists a more elegant approach
- Idea is to run regressions only on the compliers
- Compliers aren't directly observable, but they can be weighted
- Abadie showed that for any function $G = g(Y, X, D)$

$$E[G|T = c] = \frac{1}{P\{T = c\}} E[\kappa G],$$

$$\text{where} \quad \kappa \equiv 1 - \frac{D(1-Z)}{P\{Z = 0|X\}} - \frac{(1-D)Z}{P\{Z = 1|X\}}$$

Intuition

- Complier $= 1 -$ Always Taker $-$ Never Taker
- On average, κ only applies positive weights to compliers:

$$E[\kappa|T = t, X, D, Y] = I\{t = c\} \qquad \text{for } t = c, a, n$$

- So on average, κG is only positive for compliers

Using Abadie's κ

Linear/nonlinear regression

- For example, take $g(Y, X, D) = (Y - \alpha D - X'\beta)^2$ then:

$$\min_{\alpha,\beta} E[(Y - \alpha D - X'\beta)^2 | T = c] = \min_{\alpha,\beta} E\left[\kappa(Y - \alpha D - X'\beta)^2\right]$$

- Estimate α, β by solving a sample analog of the second problem
  This is just a weighted regression, with estimated weights (κ)
- Result is general enough to use for many other estimators (e.g. MLE)
- Specify X however you like — still picks out the compliers

Estimating κ

- To implement the result one must estimate κ, hence $P\{Z = 1|X\}$
- If $P\{Z = 1|X\}$ is linear, the κ-weighted (linear) regression **equals** TSLS
- Of course, Z is binary, so $P\{Z = 1|X\}$ typically won't be exactly linear
- Logit/probit often close to linear, so in practice may be close anyway

Outline

1. Linear IV and Heterogeneity
2. Roy Models: parametric approach
3. LATE
4. Abadie's κ
5. **Empirical Application: Angrist and Evans (98)**

Angrist and Evans (98, "AE")

Motivation

- Relationship between fertility decisions and female labor supply?
- Strong negative correlation, but these are joint *choices*
- Leads to many possible endogeneity stories, here's just one:
  High earning women have fewer children due to higher opportunity cost

Empirical strategy

- Y is a labor market outcome for the woman (or her husband)
- Restrict the sample to only women (or couples) with 2 or more children
- D is an indicator for having more than 2 children (vs. exactly 2)
- $Z = 1$ if first two children had the same sex
  Based on the idea that there is preference to have a mix of boys and girls
- Also consider $Z = 1$ if the second birth was a twin
  Twins are primarily for comparison — used before this paper

Assumptions in AE

Exogeneity

- Requires the assumption that sex at birth is randomly assigned
- Authors conduct balance tests to support this (next slide)
- The twins instrument is less compelling
- First, well-known that older women have twins more (see next slide)
  More subtly, it impacts both the number and *spacing* of children

Monotonicity

- Monotonicity restricts preference heterogeneity in unattractive ways
  Some families may want two boys or girls (then stop)
- No discussion of this in the paper — unfortunately common practice

Evidence in Support of Exogeneity

Table 4—Differences in Means for Demographic Variables by *Same Sex* and *Twins-2*

Variable	Difference in means (standard error)		
	By *Same sex*		By *Twins-2*
	1980 PUMS	1990 PUMS	1980 PUMS
Age	−0.0147 (0.0112)	0.0174 (0.0112)	0.2505 (0.0607)
Age at first birth	0.0162 (0.0094)	−0.0074 (0.0114)	0.2233 (0.0510)
Black	0.0003 (0.0010)	0.0021 (0.0011)	0.0300 (0.0056)
White	0.0003 (0.0012)	−0.0006 (0.0013)	−0.0210 (0.0066)
Other race	−0.0006 (0.0005)	−0.0014 (0.0009)	−0.0090 (0.0041)
Hispanic	−0.0014 (0.0009)	−0.0007 (0.0010)	−0.0069 (0.0047)
Years of education	−0.0028 (0.0076)	0.0100 (0.0074)	0.0940 (0.0415)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

- Same sex is uncorrelated with a variety of observed confounders
- Twins is well-known to be correlated with age (so, education) and race

Wald Estimates

Table 5—Wald Estimates of Labor-Supply Models

	1980 PUMS			1990 PUMS			1980 PUMS		
		Wald estimate using as covariate:			Wald estimate using as covariate:			Wald estimate using as covariate:	
Variable	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Twins-2*	*More than 2 children*	*Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

First stage (denominator of Wald) for two measures of fertility

Wald Estimates

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

	1980 PUMS			1990 PUMS			1980 PUMS		
		Wald estimate using as covariate:			Wald estimate using as covariate:			Wald estimate using as covariate:	
Variable	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Twins-2*	*More than 2 children*	*Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Reduced form (numerator of Wald) for several labor market outcomes

Wald Estimates

Table 5—Wald Estimates of Labor-Supply Models

Variable	1980 PUMS: Mean difference by *Same sex*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*	1990 PUMS: Mean difference by *Same sex*	1990 PUMS: Wald estimate using as covariate: *More than 2 children*	1990 PUMS: Wald estimate using as covariate: *Number of children*	1980 PUMS: Mean difference by *Twins-2*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

IV (Wald) estimator, e.g. $-.133 \approx -.008/.060$ — these are LATEs

Two Stage Least Squares Estimates

Table 7—OLS and 2SLS Estimates of Labor-Supply Models Using 1980 Census Data

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.176 (0.002)	−0.120 (0.025)	−0.113 (0.025) [0.013]	−0.167 (0.002)	−0.120 (0.028)	−0.113 (0.028) [0.013]	−0.008 (0.001)	0.004 (0.009)	0.001 (0.008) [0.013]
Weeks worked	−8.97 (0.07)	−5.66 (1.11)	−5.37 (1.10) [0.017]	−8.05 (0.09)	−5.40 (1.20)	−5.16 (1.20) [0.071]	−0.82 (0.04)	0.59 (0.60)	0.45 (0.59) [0.030]
Hours/week	−6.66 (0.06)	−4.59 (0.95)	−4.37 (0.94) [0.030]	−6.02 (0.08)	−4.83 (1.02)	−4.61 (1.01) [0.049]	0.25 (0.05)	0.56 (0.70)	0.50 (0.69) [0.71]
Labor income	−3768.2 (35.4)	−1960.5 (541.5)	−1870.4 (538.5) [0.126]	−3165.7 (42.0)	−1344.8 (569.2)	−1321.2 (565.9) [0.703]	−1505.5 (103.5)	−1248.1 (1397.8)	−1382.3 (1388.9) (0.549)
ln(*Family income*)	−0.126 (0.004)	−0.038 (0.064)	−0.045 (0.064) [0.319]	−0.132 (0.004)	−0.051 (0.056)	−0.053 (0.056) [0.743]	—	—	—
ln(*Non-wife income*)	—	—	—	−0.053 (0.005)	0.023 (0.066)	0.016 (0.066) [0.297]	—	—	—

Notes: The table reports estimates of the coefficient on the *More than 2 children* variable in equations (4) and (6) in the text. Other covariates in the models are *Age, Age at first birth,* plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from equation (6). The *p*-value for the test of overidentifying restrictions associated with equation (6) is shown in brackets. Standard errors are reported in parentheses.

OLS is quite different from IV — consistent with endogeneity (selection)

Two Stage Least Squares Estimates

TABLE 7—OLS AND 2SLS ESTIMATES OF LABOR-SUPPLY MODELS USING 1980 CENSUS DATA

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.176 (0.002)	−0.120 (0.025)	−0.113 (0.025) [0.013]	−0.167 (0.002)	−0.120 (0.028)	−0.113 (0.028) [0.013]	−0.008 (0.001)	0.004 (0.009)	0.001 (0.008) [0.013]
Weeks worked	−8.97 (0.07)	−5.66 (1.11)	−5.37 (1.10) [0.017]	−8.05 (0.09)	−5.40 (1.20)	−5.16 (1.20) [0.071]	−0.82 (0.04)	0.59 (0.60)	0.45 (0.59) [0.030]
Hours/week	−6.66 (0.06)	−4.59 (0.95)	−4.37 (0.94) [0.030]	−6.02 (0.08)	−4.83 (1.02)	−4.61 (1.01) [0.049]	0.25 (0.05)	0.56 (0.70)	0.50 (0.69) [0.71]
Labor income	−3768.2 (35.4)	−1960.5 (541.5)	−1870.4 (538.5) [0.126]	−3165.7 (42.0)	−1344.8 (569.2)	−1321.2 (565.9) [0.703]	−1505.5 (103.5)	−1248.1 (1397.8)	−1382.3 (1388.9) (0.549)
ln(*Family income*)	−0.126 (0.004)	−0.038 (0.064)	−0.045 (0.064) [0.319]	−0.132 (0.004)	−0.051 (0.056)	−0.053 (0.056) [0.743]	—	—	—
ln(*Non-wife income*)	—	—	—	−0.053 (0.005)	0.023 (0.066)	0.016 (0.066) [0.297]	—	—	—

Notes: The table reports estimates of the coefficient on the *More than 2 children* variable in equations (4) and (6) in the text. Other covariates in the models are *Age, Age at first birth,* plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from equation (6). The *p*-value for the test of overidentifying restrictions associated with equation (6) is shown in brackets. Standard errors are reported in parentheses.

Break same-sex into two instruments — two boys vs. two girls

Two Stage Least Squares Estimates

Table 7—OLS and 2SLS Estimates of Labor-Supply Models Using 1980 Census Data

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.176	−0.120	−0.113	−0.167	−0.120	−0.113	−0.008	0.004	0.001
	(0.002)	(0.025)	(0.025)	(0.002)	(0.028)	(0.028)	(0.001)	(0.009)	(0.008)
			[0.013]			[0.013]			[0.013]
Weeks worked	−8.97	−5.66	−5.37	−8.05	−5.40	−5.16	−0.82	0.59	0.45
	(0.07)	(1.11)	(1.10)	(0.09)	(1.20)	(1.20)	(0.04)	(0.60)	(0.59)
			[0.017]			[0.071]			[0.030]
Hours/week	−6.66	−4.59	−4.37	−6.02	−4.83	−4.61	0.25	0.56	0.50
	(0.06)	(0.95)	(0.94)	(0.08)	(1.02)	(1.01)	(0.05)	(0.70)	(0.69)
			[0.030]			[0.049]			[0.71]
Labor income	−3768.2	−1960.5	−1870.4	−3165.7	−1344.8	−1321.2	−1505.5	−1248.1	−1382.3
	(35.4)	(541.5)	(538.5)	(42.0)	(569.2)	(565.9)	(103.5)	(1397.8)	(1388.9)
			[0.126]			[0.703]			(0.549)
ln(*Family income*)	−0.126	−0.038	−0.045	−0.132	−0.051	−0.053	—	—	—
	(0.004)	(0.064)	(0.064)	(0.004)	(0.056)	(0.056)			
			[0.319]			[0.743]			
ln(*Non-wife income*)	—	—	—	−0.053	0.023	0.016	—	—	—
				(0.005)	(0.066)	(0.066)			
						[0.297]			

Notes: The table reports estimates of the coefficient on the *More than 2 children* variable in equations (4) and (6) in the text. Other covariates in the models are *Age, Age at first birth,* plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from equation (6). The *p*-value for the test of overidentifying restrictions associated with equation (6) is shown in brackets. Standard errors are reported in parentheses.

Overid test p-values — many interpretations with heterogeneity

Controversy and Debate

Controversy

- Angrist, Imbens and Rubin (1996, "AIR") — special issue of JASA
- Article criticized "econometric approaches" (latent variable notation)
- Advocated potential outcome "approach" (notation) as more credible
- LATE held up as an example of the fruits of potential outcomes
- Clearly struck a nerve with some (see commenting articles)

Debate

- Many economists are skeptical of the relevance of LATEs
- The *definition* of LATE depends on the instrument — external validity? What do same-sex compliers tell us about policy?
- LATE as a battle in a broader debate: Internal vs. external validity
- This debate continues in economics — see symposia in the JoE, JEP, JEL
- Camps sometimes roughly described as "structural" or "reduced form"
- Both groups make good points — why not combine their best elements?