Instrumental Variables

Heterogeneous Effects

MIXTAPE SESSIONS

Roadmap

- The LATE Theorem
  - Potential Outcome Setup
  - Theorem and Extensions
- Characterizing Compliers
  - Outcomes
  - Covariates
- Marginal Treatment Effects
  - Continuous Instruments
  - Discrete Instruments

From Constant to Heterogeneous Effects

So far we have implicitly been considering models w/ constant effects

- $Y_i = \alpha + \beta D_i + \varepsilon_i$ implies $\partial Y / \partial D = \beta$ for all observations i
- What if this model is *misspecified*? I.e. what if $Y_i = \alpha + \beta_i D_i + \varepsilon_i$?

From Constant to Heterogeneous Effects

So far we have implicitly been considering models w/ constant effects

- $Y_i = \alpha + \beta D_i + \varepsilon_i$ implies $\partial Y / \partial D = \beta$ for all observations i
- What if this model is *misspecified*? I.e. what if $Y_i = \alpha + \beta_i D_i + \varepsilon_i$?

Intuitively, different “research designs” (e.g. instruments) may capture different effects of the same treatment — even when all are valid

- Recall charter lottery vs. takeover IVs: very different setups!

From Constant to Heterogeneous Effects

So far we have implicitly been considering models w/ constant effects

- $Y_i = \alpha + \beta D_i + \varepsilon_i$ implies $\partial Y / \partial D = \beta$ for all observations i
- What if this model is *misspecified*? I.e. what if $Y_i = \alpha + \beta_i D_i + \varepsilon_i$?

Intuitively, different "research designs" (e.g. instruments) may capture different effects of the same treatment — even when all are valid

- Recall charter lottery vs. takeover IVs: very different setups!

Formalized in the (Nobel-winning!) Imbens and Angrist '94 LATE thm.

- Using a general potential outcomes framework...

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1)$

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1) = \alpha_i + \beta_i D_i$

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1) = \alpha_i + \beta_i D_i$
- Only observe one of these POs for each i; other is a *counterfactual*
- Interested in the treatment effects $\beta_i = Y_i(1) - Y_i(0)$

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_i Y_i(1) = \alpha_i + \beta_i D_i$
- Only observe one of these POs for each i; other is a *counterfactual*
- Interested in the treatment effects $\beta_i = Y_i(1) - Y_i(0)$

Imbens-Angrist' insight: we can also do this for an IV first stage:

- Let $D_i(0)$ and $D_i(1)$ denote individual i's potential *treatment* given a binary *instrument* $Z_i \in \{0, 1\}$:

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1) = \alpha_i + \beta_i D_i$
- Only observe one of these POs for each i; other is a *counterfactual*
- Interested in the treatment effects $\beta_i = Y_i(1) - Y_i(0)$

Imbens-Angrist' insight: we can also do this for an IV first stage:

- Let $D_i(0)$ and $D_i(1)$ denote individual i's potential *treatment* given a binary *instrument* $Z_i \in \{0, 1\}$: $D_i = (1 - Z_i)D_i(0) + Z_iD_i(1)$

Potential Outcome Setup

Let $Y_i(0)$ and $Y_i(1)$ denote individual i's potential outcomes given a binary treatment $D_i \in \{0, 1\}$

- Observed outcomes: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1) = \alpha_i + \beta_i D_i$
- Only observe one of these POs for each i; other is a *counterfactual*
- Interested in the treatment effects $\beta_i = Y_i(1) - Y_i(0)$

Imbens-Angrist' insight: we can also do this for an IV first stage:

- Let $D_i(0)$ and $D_i(1)$ denote individual i's potential *treatment* given a binary *instrument* $Z_i \in \{0, 1\}$: $D_i = (1 - Z_i)D_i(0) + Z_iD_i(1)$

Under what assumptions can we causally interpret `ivreg2 Y (D=Z)`?

Imbens and Angrist (1994) Assumptions

1. *As-good-as-random assignment*: $Z_i \perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$
   - → Consider the Angrist draft lottery, or Angrist-Krueger's QoB IV

Imbens and Angrist (1994) Assumptions

1. *As-good-as-random assignment*: $Z_i \perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$
   - → Consider the Angrist draft lottery, or Angrist-Krueger's QoB IV
2. *Exclusion*: Z_i only affects Y_i through its effect on D_i
   - → Implicit in our potential outcomes notation: $Y_i(d)$ not indexed by Z_i

Imbens and Angrist (1994) Assumptions

1. *As-good-as-random assignment*: $Z_i \perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$
   - $\rightarrow$ Consider the Angrist draft lottery, or Angrist-Krueger's QoB IV
2. *Exclusion*: Z_i only affects Y_i through its effect on D_i
   - $\rightarrow$ Implicit in our potential outcomes notation: $Y_i(d)$ not indexed by Z_i
3. *Relevance*: Z_i is correlated with D_i
   - $\rightarrow$ Equivalently, given Assumption 1, $E[D_i(1) - D_i(0)] \neq 0$

Imbens and Angrist (1994) Assumptions

1. *As-good-as-random assignment*: $Z_i \perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$

   → Consider the Angrist draft lottery, or Angrist-Krueger's QoB IV

2. *Exclusion*: Z_i only affects Y_i through its effect on D_i

   → Implicit in our potential outcomes notation: $Y_i(d)$ not indexed by Z_i

3. *Relevance*: Z_i is correlated with D_i

   → Equivalently, given Assumption 1, $E[D_i(1) - D_i(0)] \neq 0$

4. *Monotonicity*: $D_i(1) \geq D_i(0)$ for all i (i.e., almost-surely)

   → The instrument can only shift the treatment in one direction

Local Average Treatment Effect (LATE) Identification

Imbens and Angrist showed that under these assumptions:

$$\beta^{IV} = E[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0)]$$

The IV estimand β^{IV} identifies a LATE: the average treatment effect $Y_i(1) - Y_i(0)$ among *compliers*: those with $1 = D_i(1) > D_i(0) = 0$

Local Average Treatment Effect (LATE) Identification

Imbens and Angrist showed that under these assumptions:

$$\beta^{IV} = E[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0)]$$

The IV estimand β^{IV} identifies a LATE: the average treatment effect $Y_i(1) - Y_i(0)$ among *compliers*: those with $1 = D_i(1) > D_i(0) = 0$

- Intuitively, IV can't tell us anything about the treatment effects of *never-takers* $D_i(1) = D_i(0) = 0$ / *always-takers* $D_i(1) = D_i(0) = 1$]

Local Average Treatment Effect (LATE) Identification

Imbens and Angrist showed that under these assumptions:

$$\beta^{IV} = E[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0)]$$

The IV estimand β^{IV} identifies a LATE: the average treatment effect $Y_i(1) - Y_i(0)$ among *compliers*: those with $1 = D_i(1) > D_i(0) = 0$

- Intuitively, IV can't tell us anything about the treatment effects of *never-takers* $D_i(1) = D_i(0) = 0$ / *always-takers* $D_i(1) = D_i(0) = 1$]
- Monotonicity rules out the presence of *defiers*, with $D_i(1) < D_i(0)$

Local Average Treatment Effect (LATE) Identification

Imbens and Angrist showed that under these assumptions:

$$\beta^{IV} = E[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0)]$$

The IV estimand β^{IV} identifies a LATE: the average treatment effect $Y_i(1) - Y_i(0)$ among *compliers*: those with $1 = D_i(1) > D_i(0) = 0$

- Intuitively, IV can't tell us anything about the treatment effects of *never-takers* $D_i(1) = D_i(0) = 0$ / *always-takers* $D_i(1) = D_i(0) = 1$]
- Monotonicity rules out the presence of *defiers*, with $D_i(1) < D_i(0)$

$\Longrightarrow$ Different (valid) IVs can identify different LATEs!

What Does This Mean *Practically*?

Two conceptually distinct considerations: *internal* vs. *external* validity

- Context of an IV, and who the compliers likely are, may matter
- Usual “overidentification” test logic fails: two valid IVs may have different estimands! (see Kitagawa (2015) for alternative tests)

What Does This Mean *Practically*?

Two conceptually distinct considerations: *internal* vs. *external* validity

- Context of an IV, and who the compliers likely are, may matter
- Usual "overidentification" test logic fails: two valid IVs may have different estimands! (see Kitagawa (2015) for alternative tests)

In addition to as-good-as-random assignment / exclusion, we may need to worry about monotonicity when we do IV

- Sensible in earlier lottery / natural experiment / panel examples
- Maybe questionable in judge IVs (coming soon!)

Extensions

Angrist and Imbens worked out the original LATE theroem for binary D_i, discrete Z_i, and no included controls

Extensions

Angrist and Imbens worked out the original LATE theroem for binary D_i, discrete Z_i, and no included controls... but it extends

Extensions

Angrist and Imbens worked out the original LATE theroem for binary D_i, discrete Z_i, and no included controls... but it extends

- Angrist/Imbens '95: multivalued (ordered) D_i, saturated covariates
- Angrist/Graddy/Imbens '00: continuous D_i (supply/demand setup)
- Heckman/Vytlicil '05: continuous Z_i (more on this soon)
- Multiple unordered treatments is harder (e.g. Behaghel et al. 2013)

Extensions

Angrist and Imbens worked out the original LATE theroem for binary D_i, discrete Z_i, and no included controls... but it extends

- Angrist/Imbens '95: multivalued (ordered) D_i, saturated covariates
- Angrist/Graddy/Imbens '00: continuous D_i (supply/demand setup)
- Heckman/Vytlicil '05: continuous Z_i (more on this soon)
- Multiple unordered treatments is harder (e.g. Behaghel et al. 2013)

Recent discussions highlight importance of including flexible controls

- E.g. Słoczyński '20, Borusyak and Hull '21, Mogstad et al. '22
- If monotonicity only holds conditional on X_i, may need Z_i-by-X_i interactions (which may lead to many-weak problems...)

Roadmap

The LATE Theorem

- Potential Outcome Setup
- Theorem and Extensions

Characterizing Compliers

- Outcomes
- Covariates

Marginal Treatment Effects

- Continuous Instruments
- Discrete Instruments

Who Are the Compliers?

Characterizing the i that make up the IV estimand (w/$D_i(1) > D_i(0)$) is key for understanding internal vs. external validity

- Unfortunately we can't identify compliers directly: we only observe $D_i(1)$ (when $Z_i = 1$) or $D_i(0)$ (when $Z_i = 0$), not both together!

Who Are the Compliers?

Characterizing the i that make up the IV estimand (w/$D_i(1) > D_i(0)$) is key for understanding internal vs. external validity

- Unfortunately we can't identify compliers directly: we only observe $D_i(1)$ (when $Z_i = 1$) or $D_i(0)$ (when $Z_i = 0$), not both together!

It turns out we can still characterize compliers by their outcomes ($Y_i(0)$ and $Y_i(1)$) and other observables X_i

- Comparing $E[X_i \mid D_i(1) > D_i(0)]$ to $E[X_i]$ can maybe shed light on how $E[Y_i(1) - Y_i(0) \mid D_i(1) > D_i(0)]$ compares to $E[Y_i(1) - Y_i(0)]$

Outcomes

Computing $E[Y_i(1) \mid D_i(1) > D_i(0)]$ is surprisingly easy in the IA setup

- Define $W_i = Y_i D_i$, and note that this new outcome has potentials with respect to D_i of $W_i(1) = Y_i(1)$ and $W_i(0) = 0$

Outcomes

Computing $E[Y_i(1) \mid D_i(1) > D_i(0)]$ is surprisingly easy in the IA setup

- Define $W_i = Y_i D_i$, and note that this new outcome has potentials with respect to D_i of $W_i(1) = Y_i(1)$ and $W_i(0) = 0$
- Thus IV with W_i as the outcome identifies $E[W_i(1) - W_i(0) \mid D_i(1) > D_i(0)] = E[Y_i(1) \mid D_i(1) > D_i(0)]$

Outcomes

Computing $E[Y_i(1) \mid D_i(1) > D_i(0)]$ is surprisingly easy in the IA setup

- Define $W_i = Y_i D_i$, and note that this new outcome has potentials with respect to D_i of $W_i(1) = Y_i(1)$ and $W_i(0) = 0$
- Thus IV with W_i as the outcome identifies $E[W_i(1) - W_i(0) \mid D_i(1) > D_i(0)] = E[Y_i(1) \mid D_i(1) > D_i(0)]$

Similar logic shows that IV with $Y_i(1 - D_i)$ as the outcome and $1 - D_i$ as the treatment identifies $E[Y_i(0) \mid D_i(1) > D_i(0)]$

- So easy to do! And extends to covariates / multiple IVs...

Characterizing Charter Lottery Complier $Y_i(0)$'s

TABLE 6—POTENTIAL-OUTCOME GAPS IN URBAN AND NONURBAN AREAS

	Urban				Nonurban			
Subject	Treatment effect (1)	$E_u[Y_0 \mid D=0]$ (2)	λ_0^u (3)	λ_1^u (4)	Treatment effect (5)	$E_n[Y_0 \mid D=0]$ (6)	λ_0^n (7)	λ_1^n (8)
Panel A. Middle school								
Math	0.483***	−0.399***	0.077	0.560***	−0.177**	0.236***	0.010	−0.143***
	(0.074)	(0.011)	(0.049)	(0.054)	(0.074)	(0.007)	(0.061)	(0.042)
N	4,858				2,239			
ELA	0.188***	−0.422***	0.118**	0.306***	−0.148***	0.260***	0.102**	−0.086***
	(0.064)	(0.012)	(0.054)	(0.049)	(0.048)	(0.007)	(0.050)	(0.030)
N	4,551				2,323			
Panel B. High school								
Math	0.557***	−0.371***	0.074	0.602***	0.065	0.241***	0.207	0.271***
	(0.164)	(0.021)	(0.099)	(0.151)	(0.146)	(0.008)	(0.145)	(0.041)
N	3,743				432			
ELA	0.417***	−0.369***	−0.004	0.410***	0.064	0.250***	0.237	0.301***
	(0.140)	(0.018)	(0.096)	(0.119)	(0.151)	(0.008)	(0.152)	(0.051)
N	4,858				435			

Source: Angrist, Pathak, and Walters (2013)

Covariates

For covariates X_i (not affected by D_i) we can follow a similar trick:

- Either IV'ing X_iD_i on D_i or IV'ing $X_i(1-D_i)$ on $1-D_i$ identifies complier characteristics $E[X_i \mid D_i(1) > D_i(0)]$
- Shouldn't be very different (implicit balance test); can be averaged

Covariates

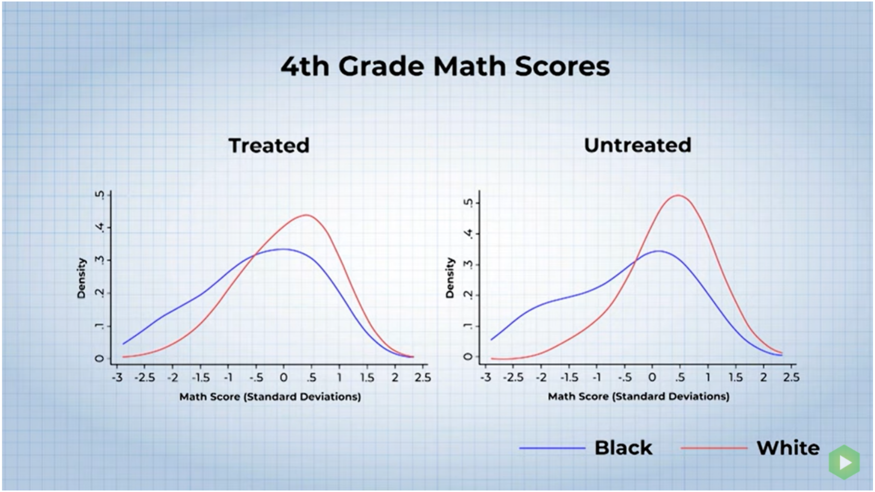
For covariates X_i (not affected by D_i) we can follow a similar trick:

- Either IV'ing X_iD_i on D_i or IV'ing $X_i(1 - D_i)$ on $1 - D_i$ identifies complier characteristics $E[X_i \mid D_i(1) > D_i(0)]$
- Shouldn't be very different (implicit balance test); can be averaged

Abadie (2003) gives a slicker (but a bit more involved) approach to estimating any function of $(Y_i(0), Y_i(1), X_i)$ for compliers

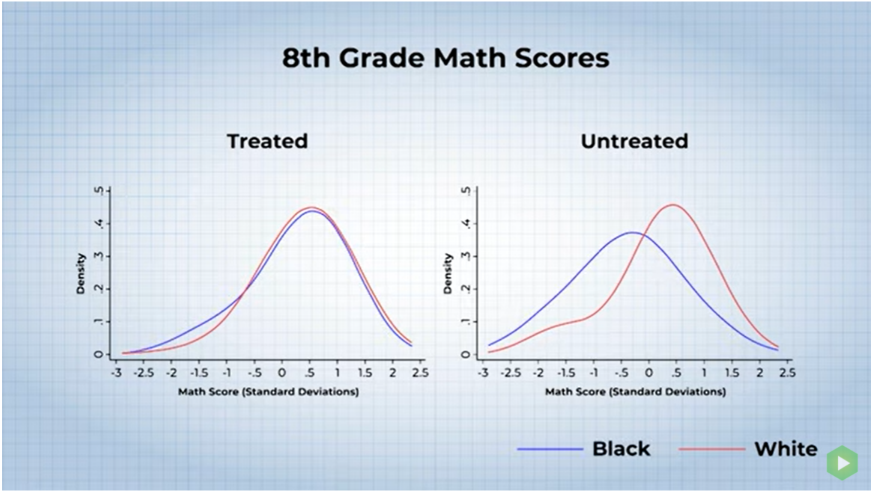
- Involves weighting by $\kappa = 1 - \frac{D_i(1-Z_i)}{1-E[Z_i|W_i]} - \frac{(1-D_i)Z_i}{E[Z_i|W_i]}$ where W_i are any necessary "design controls" (e.g. lottery risk sets)
- You can do some really cool stuff with this!

Black/White Potential Outcomes, Pre-Charter



Source: Josh Angrist Nobel Lecture (2021)

Black/White Potential Outcomes, Post-Charter



Source: Josh Angrist Nobel Lecture (2021)

Roadmap

The LATE Theorem

- Potential Outcome Setup
- Theorem and Extensions

Characterizing Compliers

- Outcomes
- Covariates

Marginal Treatment Effects

- Continuous Instruments
- Discrete Instruments

Heckman and Vytlicil (2005, 2007, 2010, 2013...)

If we have a Z_i that varies continuously, we might learn more about how treatment effects vary with compliance

- Different types of i may “respond” at different margins of Z_i

Heckman and Vytlicil (2005, 2007, 2010, 2013...)

If we have a Z_i that varies continuously, we might learn more about how treatment effects vary with compliance

- Different types of i may "respond" at different margins of Z_i

Heckman-Vytlicil write $D_i = \mathbf{1}[p(Z_i) \geq U_i]$, with $U_i \mid Z_i \sim U(0,1)$

- $p(z) = Pr(D_i = 1 \mid Z_i = z)$ is the treatment propensity score
- U_i indexes treatment "resistance" (i.e. types of compliers); Vytlacil (2002) shows model is equivalent to IA's monotonicity w/ binary Z_i

Heckman and Vytlicil (2005, 2007, 2010, 2013...)

If we have a Z_i that varies continuously, we might learn more about how treatment effects vary with compliance

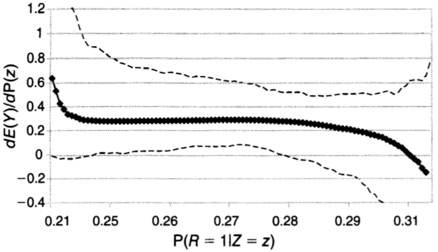
- Different types of i may "respond" at different margins of Z_i

Heckman-Vytlicil write $D_i = \mathbf{1}[p(Z_i) \geq U_i]$, with $U_i \mid Z_i \sim U(0,1)$

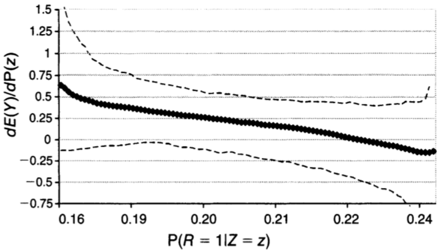
- $p(z) = Pr(D_i = 1 \mid Z_i = z)$ is the treatment propensity score
- U_i indexes treatment "resistance" (i.e. types of compliers); Vytlacil (2002) shows model is equivalent to IA's monotonicity w/ binary Z_i

Now we can consider how $Y_i(1) - Y_i(0)$ varies continuously with U_i ...

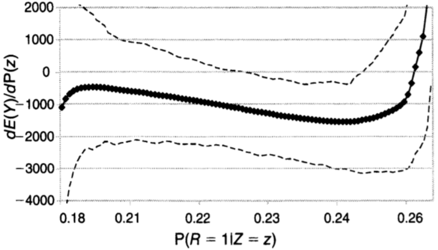
Doyle (2007): MTEs of Foster Care Removal



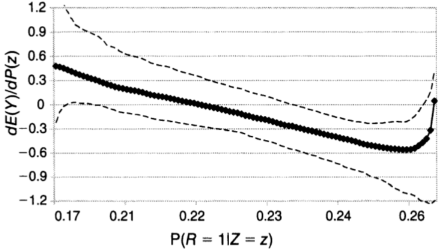
A. Delinquency MTE



B. Teen Motherhood MTE



C. Earnings MTE



D. Employment MTE

Local Instrumental Variables

Heckman (2000) shows that MTEs are identified by “local IV”:

$$E[Y_i(1) - Y_i(0) \mid U_i = p] = \frac{\partial E[Y_i \mid p(Z_i) = p]}{\partial p}$$

under natural extensions of Imbens and Angrist (1994)

Local Instrumental Variables

Heckman (2000) shows that MTEs are identified by "local IV":

$$E[Y_i(1) - Y_i(0) \mid U_i = p] = \frac{\partial E[Y_i \mid p(Z_i) = p]}{\partial p}$$

under natural extensions of Imbens and Angrist (1994)

- Suggests we flexibly estimate $p(z) = Pr(D_i = 1 \mid Z_i = z)$, $E[Y_i \mid p(Z_i)]$, and then take the derivative of the latter
- In practice this is often done parametrically, and with controls

What if We Don't Have Continuous Instruments?

A fascinating recent literature considers intermediate cases of Imbens-Angrist and Heckman-Vytlacil:

- Discrete (binary/multivalued) Z_i, with parametric/shape restrictions to trace out (or maybe bound) the MTE curve
- Effectively using a model to "extrapolate" from local variation, maybe to identify more policy-relevant parameters

What if We Don't Have Continuous Instruments?

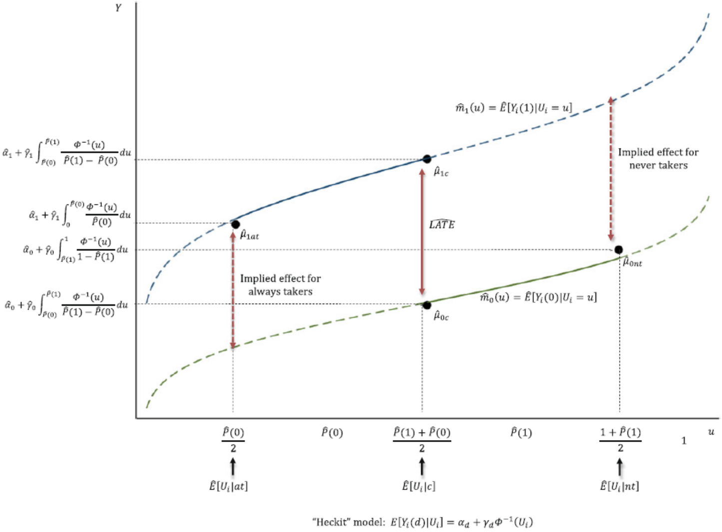
A fascinating recent literature considers intermediate cases of Imbens-Angrist and Heckman-Vytlacil:

- Discrete (binary/multivalued) Z_i, with parametric/shape restrictions to trace out (or maybe bound) the MTE curve
- Effectively using a model to "extrapolate" from local variation, maybe to identify more policy-relevant parameters

Some examples: Brinch et al. (2017), Mogstad et al. (2018), Kline and Walters (2019), Hull (2020), Arnold et al. (2021), Kowalski (2022)...

- Lots more to do here (especially on the practical side)

How Parametric "Heckit" Models Extrapolate LATEs



Source: Kline and Walters (2019)