Instrumental Variables and Regression Discontinuity

Giselle Montamat

Harvard University

Spring 2020

Instrumental Variables (without additional covariates)

Model:

$$Y_i = \beta_0 + \beta_1 T_{i,1} + u_i$$

- WLOG (since we include a constant) $E[u_i] = 0$
- But! $E[T_{i,1} u_i] \neq 0$ (so $Cov(T_{i,1}, u_i) \neq 0$)

β_1 is not identified by BLP of Y_i on a constant and $T_{i,1}$ (i.e., $\beta_1 \neq \frac{Cov(Y_i, T_{i,1})}{Var(T_{i,1})}$). We need something more...an instrument $Z_{i,1}$. It is a variable that satisfies:

- (Relevance) $Cov(T_{i,1}, Z_{i,1}) \neq 0$
- (Exclusion) $E[Z_{i,1} u_i] = 0$ (so $Cov(Z_{i,1}, u_i) = 0$)

Then β_1 is identified by the following "IV object" (Exercise: show!):

$$\beta_1 = \frac{Cov(Y_i, Z_{i,1})}{Cov(T_{i,1}, Z_{i,1})}$$

Instrumental Variables (without additional covariates)

More generally:

$$Y_i = T_i'\beta + u_i$$

$\Rightarrow$ If $Y_i = \beta_0 + \beta_1 T_{i,1} + u_i$ (i.e., $T_i = [1\ T_{i,1}]'$) and Z_i is $L \times 1$, with $L \geq 1$:

$$\beta_1 = \frac{E[T_{i,1}Z_i']E[Z_iZ_i']^{-1}E[Z_iY_i]}{E[T_{i,1}Z_i']E[Z_iZ_i']^{-1}E[Z_iT_{i,1}]}$$

Exercise: show that when $Z_i = [1\ Z_{i,1}]'$ (i.e, it includes a constant and a single instrument), we get $\beta_1 = \frac{Cov(Y_i, Z_{i,1})}{Cov(T_{i,1}, Z_{i,1})}$

$\Rightarrow$ If T_i is $k \times 1$ and Z_i is $L \times 1$, with $L \geq k$:

$$\beta = \left(E[T_iZ_i']E[Z_iZ_i']^{-1}E[Z_iT_i']\right)^{-1} E[T_iZ_i']E[Z_iZ_i']^{-1}E[Z_iY_i]$$

Instrumental Variables (without additional covariates)

This object can be recovered from a **two-stage procedure** (2SLS):

1. First stage: regress endogenous variable(s) on instrument(s) (include a constant in Z_i)

$$E^*[T_i|Z_i] = E[T_i Z_i']E[Z_i Z_i']^{-1} Z_i \equiv \hat{T}_i$$

(Note: in this formula, Z_i can be $L \times 1$ and T_i can be $k \times 1$. $E^*[T_i|Z_i]$ is $k \times 1$: each row contains the BLP of one endogenous variable (i.e., one component of T_i). If $k = 1$ then $E[T_i Z_i']E[Z_i Z_i']^{-1} Z_i$ is 1×1 and so it is equal to its transpose: $E^*[T_i|Z_i] = Z_i' E[Z_i Z_i']^{-1} E[Z_i T_i]$ which is the familiar notation "$X_i'\beta = X_i' E[X_i X_i']^{-1} E[X_i Y_i]$" that you're used to when there's only one dependent variable being regressed against covariates.)

2. Second stage: regress dependent variable on prediction from previous step

$$E^*[Y_i|\hat{T}_i] = \hat{T}_i' \underbrace{E[\hat{T}_i \hat{T}_i']^{-1} E[\hat{T}_i Y_i]}_{=\beta}$$

Exercise: show that $\beta = E[\hat{T}_i \hat{T}_i']^{-1} E[\hat{T}_i Y_i]$

Exercise: show how these formulas simplify for the case in which $T_i = [1 \ T_{i,1}]$ and $Z_i = [1 \ Z_{i,1}]$ (i.e., it includes a constant and a single instrument).

Instrumental Variables (without additional covariates)

IV object (2SLS estimand) can be seen as a GMM estimand derived from the following moment condition:

$$E[Z_i u_i] = 0$$

$$\underbrace{E[Z_i(Y_i - T_i'\beta)]}_{L\times 1} = 0$$

This is a linear system of L equations and k unknowns.

Note: by including a constant in Z_i, we include $E[u_i] = 0$ in our system of equations.

Instrumental Variables (without additional covariates)

Let's look at the IV object if system is just-identified and over-identified:

$$\beta = \left(E[T_iZ_i']E[Z_iZ_i']^{-1}E[Z_iT_i']\right)^{-1}E[T_iZ_i']E[Z_iZ_i']^{-1}E[Z_iY_i]$$

- If $L = k$, notice that we can simplify this formula: $E[T_iZ_i']$ and $E[Z_iT_i']$ are $k \times k$ and full rank (by relevance assumption), so they are invertible (and remember: $(AB)^{-1} = B^{-1}A^{-1}$)

$$\beta = E[Z_iT_i']^{-1}\left(E[T_iZ_i']E[Z_iZ_i']^{-1}\right)^{-1}E[T_iZ_i']E[Z_iZ_i']^{-1}E[Z_iY_i]$$

$$\beta = E[Z_iT_i']^{-1}E[Z_iY_i]$$

We could arrive at this result by solving the system of equations from before:

$$E[Z_i(Y_i - T_i'\beta)] = 0$$

$$E[Z_iY_i] = E[Z_iT_i']\beta$$

Since $E[Z_iT_i']$ is invertible:

$$\beta = E[Z_iT_i']^{-1}E[Z_iY_i]$$

Instrumental Variables (without additional covariates)

Exercise: show what the moment conditions look like in the case in which $T_i = [1 \; T_{i,1}]$ and $Z_i = [1 \; Z_{i,1}]$ (i.e., one endogenous regressor and one instrument).

$$E[Z_i(Y_i - T_i'\beta)] = 0$$

$$\begin{bmatrix} E[Y_i - \beta_0 - \beta_1 T_{i,1}] \\ E[Z_{i,1}(Y_i - \beta_0 - \beta_1 T_{i,1})] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Exercise: show that $\beta_0 = E[Y_i] - \beta_1 E[T_{i,1}]$ and $\beta_1 = \frac{Cov(Y_i, Z_{i,1})}{Cov(T_{i,1}, Z_{i,1})}$.

Conclusion: $\frac{Cov(Y_i, Z_{i,1})}{Cov(T_{i,1}, Z_{i,1})}$ identifies the β_1 from a model $Y_i = \beta_0 + \beta_1 T_{i,1} + u_i$ that satisfies the moment conditions $E[u_i] = 0$ and $E[Z_{i,1} u_i] = 0$ (but doesn't satisfy the moment condition $E[T_{i,1} u_i] = 0$).

Instrumental Variables (without additional covariates)

- If $L > k$, can combine the L equations (moment conditions) and obtain k linear combinations. (A particular example of this would be to assign weight 0 to $L - k$ of them, so essentially dropping them to get a just-identified system).

  The IV object (2SLS estimand) essentially considers weight $E[T_i Z_i']E[Z_i Z_i']^{-1}$ and solves the following system of k equations and k unknowns:

$$\underbrace{E[T_i Z_i']E[Z_i Z_i']^{-1}}_{k \times L} \underbrace{E[Z_i(Y_i - T_i'\beta)]}_{L \times 1} = 0$$

Instrumental Variables (without additional covariates)

Let's look at an example:

$$Y_i = \beta_0 + \beta_1 T_{i,1} + u_i$$

Where $E[T_{i,1} u_i] \neq 0$, so $T_{i,1}$ is endogenous.

Suppose there are two instruments available:

$$E[Z_i(Y_i - T_i'\beta)] = 0$$

$$\begin{bmatrix} E[Y_i - \beta_0 - \beta_1 T_{i,1}] \\ E[Z_{i,1}(Y_i - \beta_0 - \beta_1 T_{i,1})] \\ E[Z_{i,2}(Y_i - \beta_0 - \beta_1 T_{i,1})] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Note that we are assuming that β_0 and β_1 satisfy all three equations (moment conditions).

Instrumental Variables (without additional covariates)

The issue is that, while this is true in population moments, when we replace these equations with their sample analogs (means instead of expectations) in order to obtain estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ then we (most likely) end up with three linear equations where none is a linear combination of the others, so the system will have no solution. Remember:

System of linear equations: suppose we have L equations and k unknowns.

- If $L = k$ (just-identified system): one solution
- If $L > k$ (over-identified system): no solution
- If $L < k$ (under-identified system): ∞ solutions

So we can first work with the population system of moment conditions to derive a just-identified system and then motivate our estimators as a solution to the sample analog of that system.

Instrumental Variables (without additional covariates)

We first combine the equations available to create a system of k equations. This is achieved by pre-multiplying our system by a matrix D (nonrandom) that is $k \times L$. The β_0 and β_1 that we are trying to retrieve satisfy this new system.

$$D \times E[Z_i(Y_i - T_i'\beta)] = 0$$

$$\begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \end{bmatrix} \begin{bmatrix} E[Y_i - \beta_0 - \beta_1 T_{i,1}] \\ E[Z_{i,1}(Y_i - \beta_0 - \beta_1 T_{i,1})] \\ E[Z_{i,2}(Y_i - \beta_0 - \beta_1 T_{i,1})] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\beta = \left(DE[Z_i T_i']\right)^{-1} \left(DE[Z_i Y_i]\right)$$

For instance, 2SLS considers $D = E[T_i Z_i']E[Z_i Z_i']^{-1}$.

This new system motivates an estimator $\hat{\beta}$ for β by replacing D with a $\hat{D}$ (that is consistent) and the expectations with means.

We require that $DE[Z_i T_i']$ is nonsingular, and that there is a $\hat{D}$ that converges in probability to D and $\hat{D}\frac{1}{n}\sum Z_i T_i'$ is nonsingular. While any D under these conditions will motivate an estimator of β that is consistent ($\hat{\beta} \xrightarrow{p} \beta$), some will be "better" than others (example: efficiency). For instance, the 2SLS estimator is the efficient GMM estimator under homoskedasticity.

Instrumental Variables

Model (same but adding covariates):

$$Y_i = X_i'\alpha + \beta T_i + u_i$$

- WLOG (because X_i includes a constant) $E[u_i] = 0$
- $E[X_i u_i] = 0$ (so $Cov(X_{i,j}, u_i) = 0$ for all the other j components of X_i)
- But! $E[T_i u_i] \neq 0$ (so $Cov(T_i, u_i) \neq 0$)

β is not identified by BLP of Y_i on X_i and T_i. We need something more...an instrument Z_i. It is a variable that satisfies:

- (Relevance) $Cov(T_i, Z_i) \neq 0$
- (Exclusion) $E[Z_i u_i] = 0$ (so $Cov(Z_i, u_i) = 0$)

Then β is identified by the following "IV object" (Exercise: show!):

$$\beta = \frac{Cov(Y_i, \tilde{Z}_i)}{Cov(T_i, \tilde{Z}_i)} = \frac{E[Y_i \tilde{Z}_i]}{E[T_i \tilde{Z}_i]}$$

Where $\tilde{Z}_i$ is the residual from regressing Z_i on X_i.

Instrumental Variables

More generally:

$$Y_i = X_i'\alpha + T_i'\beta + u_i$$

$\Rightarrow$ If $Y_i = X_i'\alpha + \beta T_i + u_i$ and Z_i is $L \times 1$, with $L \geq 1$:

$$\beta = \frac{E[T_i\tilde{Z}_i']E[\tilde{Z}_i\tilde{Z}_i']^{-1}E[\tilde{Z}_iY_i]}{E[T_i\tilde{Z}_i']E[\tilde{Z}_i\tilde{Z}_i']^{-1}E[\tilde{Z}_iT_i]}$$

$\Rightarrow$ If T_i is $k \times 1$ and Z_i is $L \times 1$, with $L \geq k$:

$$\beta = \left(E[T_i\tilde{Z}_i']E[\tilde{Z}_i\tilde{Z}_i']^{-1}E[\tilde{Z}_iT_i']\right)^{-1} E[T_i\tilde{Z}_i']E[\tilde{Z}_i\tilde{Z}_i']^{-1}E[\tilde{Z}_iY_i]$$

Where $\tilde{Z}_i$ is the residual from regressing Z_i on X_i.

Instrumental Variables

In what cases do we have a model where T_i is "endogenous"? (And so its coefficient can't be identified by BLP).

IV estimator motivated in different contexts:

1. Simulatenous equation bias ("reverse causation")
2. Measurement error bias
3. Omitted variable bias (OVB)

That is, people have shown that in these contexts, the "IV object" can identify the desired structural parameter (the coefficient on T_i).

But basically, all these can be seen as an OVB problem.

Instrumental Variables

OBV:

$$Y_i = \alpha + \beta T_i + \underbrace{\gamma A_i + v_i}_{u_i}$$

- WLOG (since we include a constant) $E[v_i] = 0$
- $E[T_i v_i] = 0$ (so $Cov(T_i, v_i) = 0$)
- $E[A_i v_i] = 0$ (so $Cov(A_i, v_i) = 0$)

If we could regress Y_i on a constant, T_i and A_i (ECON2120: "long regression"), then β is identified. Assumptions imply that BLP recovers the coefficients of the model: $E^*[Y_i|1, T_i, A_i] = \alpha + \beta T_i + \gamma A_i$

Instrumental Variables

But suppose that instead we regress Y_i on a constant and T_i only (ECON2120: "short regression"):

$$E^*[Y_i|1, T_i] = E^*[\alpha + \beta T_i + \gamma A_i + v_i|1, T_i]$$
$$= \alpha + \beta T_i + \gamma E^*[A_i|1, T_i]$$

Auxiliary regression: $E^*[A_i|1, T_i] = \phi_0 + \phi_1 T_i$ where $\phi_1 = \frac{Cov(A_i, T_i)}{Var(T_i)}$

$$E^*[Y_i|1, T_i] = \alpha + \beta T_i + \gamma E^*[A_i|1, T_i]$$
$$= \alpha + \gamma\phi_0 + (\beta + \gamma\phi_1) T_i$$

So unless $Cov(T_i, A_i) = 0$, the BLP $E^*[Y_i|1, T_i]$ doesn't allow to identify β. Instead, it identifies $\beta + \gamma\phi_1$, where $\gamma\phi_1$ is the "OVB".

Instrumental Variables

IV to the rescue! β can be identified by the "IV object" $\frac{Cov(Y_i,Z_i)}{Cov(Y_i,T_i)}$ for a Z_i that satisfies $Cov(Z_i, \underbrace{\gamma A_i + v_i}_{u_i}) = 0$ (exclusion) and $Cov(T_i, Z_i) \neq 0$ (relevance):

$$Cov(Y_i, Z_i) = Cov(\alpha + \beta T_i + \gamma A_i + v_i, Z_i) = \beta Cov(T_i, Z_i)$$

So far we haven't said anything about causality. Let's see a model that introduces the notion of causality and how the "IV object" helps identify a causal effect.

Linear model of constant causal effects

Simple model with causal interpretation:

$$Y_i = Y_i(T_i) = \alpha + \beta T_i + u_i$$

In particular, if $T_i \in \{0, 1\}$:

$$Y_i(0) = \alpha + u_i$$

$$Y_i(1) = \alpha + \beta + u_i$$

$$\Rightarrow TE \equiv Y_i(1) - Y_i(0) = \beta$$

$$Y_i = Y_i(1)T_i + Y_i(0)(1 - T_i) = Y_i(0) + (Y_i(1) - Y_i(0))T_i = \alpha + \beta T_i + u_i$$

So model assumes <u>no</u> heterogeneity of treatment effects.

Goal is to identify $\beta = TE = Y_i(1) - Y_i(0)$ (constant for every i).
Note: TE = ATE because TE is constant.

Linear model of constant causal effects

1) Easy case:

- WLOG (since we introduce a constant) $E(u_i) = 0$
- Assume $E[T_i u_i] = 0$ (so $Cov(T_i, u_i) = 0$) (treatment is independent of potential outcomes)

$\Rightarrow$ under these identifiying assumptions, β identified by BLP:

$$\beta = TE = ATE = \frac{Cov(Y_i, T_i)}{Var(T_i)}$$

Exercise: Show that $\frac{Cov(Y_i, T_i)}{Var(T_i)} = E[Y_i | T_i = 1] - E[Y_i | T_i = 0]$

Linear model of constant causal effects

2) Adding some more structure to the model:

$$Y_i = Y_i(T_i) = \alpha + \beta T_i + \underbrace{\gamma A_i + v_i}_{\equiv u_i}$$

- A_i is additional control that affects potential outcomes
- WLOG (since we introduce a constant) $E(v_i) = 0$
- Assume $E[A_i v_i] = 0$ (so $Cov(A_i, v_i) = 0$)
- Assume $E[T_i v_i] = 0$ (so $Cov(T_i, v_i) = 0$) and A_i **observed** (treatment is independent of the unobservable stuff that affects potential outcomes, and the observable stuff we can control for because...it's observed.)

Linear model of constant causal effects

$\Rightarrow$ under these identifiying assumptions, β is identified by BLP of Y_i on a constant and two regressors T_i, A_i.

$$\beta = TE = ATE = \left\{E(X_i X_i')^{-1} E(X_i Y_i)\right\}_{2,2} = \frac{Cov(Y_i, \tilde{T}_i)}{Var(\tilde{T}_i)} = \frac{E(Y_i \tilde{T}_i)}{E(\tilde{T}_i^2)}$$

Where $X_i = [1 \ T_i \ A_i]$ and $\tilde{T}_i$ is the residual from regressing T_i against a constant and A_i (remember Frisch-Waugh-Lovell?).

3) But! Suppose A_i is **unobserved** and $Cov(T_i, A_i) \neq 0$ (example: A_i is ability, T_i is education, and smarter kids select into better schools). In other words, treatment is not independent of unobserved stuff that affects potential outcomes.

$\Rightarrow$ BLP of Y_i on a constant and T_i can't identify β. There is OVB.

Linear model of constant causal effects

Instrumental variables: suppose there is an instrument $Z_i \in \{0,1\}$ that satisfies:

- (Relevance) $Cov(T_i, Z_i) \neq 0$
- (Exclusion) $Cov(u_i, Z_i) \neq 0$ (it only impacts observed outcome Y_i through T_i; that is, it doesn't affect potential outcomes)

$\Rightarrow \beta$ is identified by:

$$\beta = TE = ATE = \frac{Cov(Y_i, Z_i)}{Cov(T_i, Z_i)}$$

Exercise: Show that $\frac{Cov(Y_i,Z_i)}{Cov(T_i,Z_i)} = \frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[T_i|Z_i=1]-E[T_i|Z_i=0]}$

Conclusion: in a linear model of constant causal effects with a binary endogenous treatment and a binary instrument, $\frac{Cov(Y_i,Z_i)}{Cov(T_i,Z_i)}$ identifies the TE (which is also the ATE).

Nonparametric model of heterogenous treatment effects

More general model with causal interpretation (we still assume $T_i \in \{0, 1\}$):

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

We don't want to make any further assumptions about $(Y_i(1) - Y_i(0))$. So model assumes heterogeneity of treatment effects.

Goal is to identify $ATE = E[Y_i(1) - Y_i(0)]$.

1) Easy case: $T_i \perp (Y_i(0), Y_i(1))$
$\Rightarrow$ under these identifiying assumptions, ATE identified by:

$$ATE = \frac{Cov(Y_i, T_i)}{Var(T_i)}$$

Exercise: Show that $\frac{Cov(Y_i,T_i)}{Var(T_i)} = E[Y_i | T_i = 1] - E[Y_i | T_i = 0]$

Nonparametric model of heterogenous treatment effects

2) $T_i \not\perp (Y_i(0), Y_i(1))$

Instrumental variables: suppose there is an instrument $Z_i \in \{0, 1\}$ that satisfies:

- (Relevance) $Cov(T_i, Z_i) \neq 0$
- (Exclusion) + (Independence) Y_i is not a function of Z_i and $Z_i \perp Y_i(1), Y_i(0)$

Also: $Z_i \perp T_i(1), T_i(0)$

$\Rightarrow$ ATE is identified for compliers (LATE) -under monotonicity- by:

$$LATE = \frac{Cov(Y_i, Z_i)}{Cov(T_i, Z_i)}$$

Exercise: Show that $\frac{Cov(Y_i,Z_i)}{Cov(T_i,Z_i)} = \frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[T_i|Z_i=1]-E[T_i|Z_i=0]}$

Note: if we use another instrument, then we uncover another LATE in the sense that there will be another set of compliers for that other instrument.

Recap

We've studied the potential outcomes model:

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

(Implicit is SUTVA assumption)

Identifying assumptions on how treatment is assigned - cases discussed:

1) T is randomly assigned: $\{Y_i(1), Y_i(0)\} \perp T_i$
2) T is *not* randomly assigned: $\{Y_i(1), Y_i(0)\} \not\perp T_i$
3) T is not randomly assigned but there is random assignment of an instrument Z: $\{Y_i(1), Y_i(0), T_i(1), T_i(0)\} \perp Z_i$
4) T is randomly assigned conditional on a set of observable characteristics X: $\{Y_i(1), Y_i(0)\} \perp T_i \mid X_i$
5) T is not randomly assigned but there is random assignment of an instrument Z if we condition on a set of observables X: $\{Y_i(1), Y_i(0), T_i(1), T_i(0)\} \perp Z_i \mid X_i$

Regression discontinuity

Still within the framework of potential outcomes model, but now instead of relying on random assignment of treatment (or instrument) as identifying assumption, the key assumption is that there is some "running" variable according to which treatment (or instrument) is assigned.

Sharp RD: running variable determines if you received treatment or not

$$T_i = T_i(R) = \begin{cases} 0 & \text{if } R < r^* \\ 1 & \text{if } R \geq r^* \end{cases}$$

$$Y_i = Y_i(1)1\{R \geq r^*\} + Y_i(0)1\{R < r^*\}$$

Fuzzy RD: running variable determines if you received instrument or not

$$Z_i = Z_i(R) = \begin{cases} 0 & \text{if } R < r^* \\ 1 & \text{if } R \geq r^* \end{cases}$$

$$T_i = T_i(1)1\{R \geq r^*\} + T_i(0)1\{R < r^*\}$$

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

Analogies to help you remember

Random assignment of T (Case 1)):

1. Key assumption: $Y_i(0), Y_i(1) \perp T_i$
2. Treatment is independent of potential outcomes, so people are "intrinsically" the same (on average), except that some got treatment and some didn't. So compare those in treated v non-traded groups to identify causal effect of treatment.
3. Can identify ATE:

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

Sharp RD:

1. Key assumption: $E[Y_i(0)|R_i]$ and $E[Y_i(1)|R_i]$ continuous at $R_i = r^*$.
2. People "close to" $R = r^*$ are "intrinsically" the same (on average), except some got treatment and some didn't. So compare their outcomes to identify causal effect of treatment.
3. Can identify a ATE at $R = r^*$:

$$E[Y_i(1)|R_i = r^*] - E[Y_i(0)|R_i = r^*] = \lim_{r \downarrow r^*} E[Y_i|R_i = r] - \lim_{r \uparrow r^*} E[Y_i|R_i = r]$$

Analogies to help you remember

Random assignment of Z (Case 3)):

1. Key assumption: $Y_i(0), Y_i(1), T_i(0), T_i(1) \perp Z_i$
2. + exlcusion, monotonicity, first stage
3. Can identify LATE:

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

Fuzzy RD:

1. Key assumption: $E[Y_i(1)|R_i]$, $E[Y_i(0)|R_i]$, $E[T_i(1)|R_i]$ and $E[T_i(0)|R_i]$ are continuous at $R_i = r^*$
2. + exclusion, monotonicity, first stage, all conditional on $R_i = r^*$
3. Can identify a LATE at $R = r^*$:

$$E[Y_i(1) - Y_i(0) | T_i(1) > T_i(0), R_i = r^*] = \frac{\lim_{r \downarrow r^*} E[Y_i|R_i = r] - \lim_{r \uparrow r^*} E[Y_i|R_i = r]}{\lim_{r \downarrow r^*} E[T_i|R_i = r] - \lim_{r \uparrow r^*} E[T_i|R_i = r]}$$