Lecture 10: Instrumental Variables I

POL-GA 1251
Quantitative Political Analysis II
Prof. Cyrus Samii
NYU Politics

Feburary 27, 2018

For today:

- Basics of IV for causal effects from a semi-parametric regression perspective.

For next class:

- IV from a non-parametric, potential outcomes perspective.

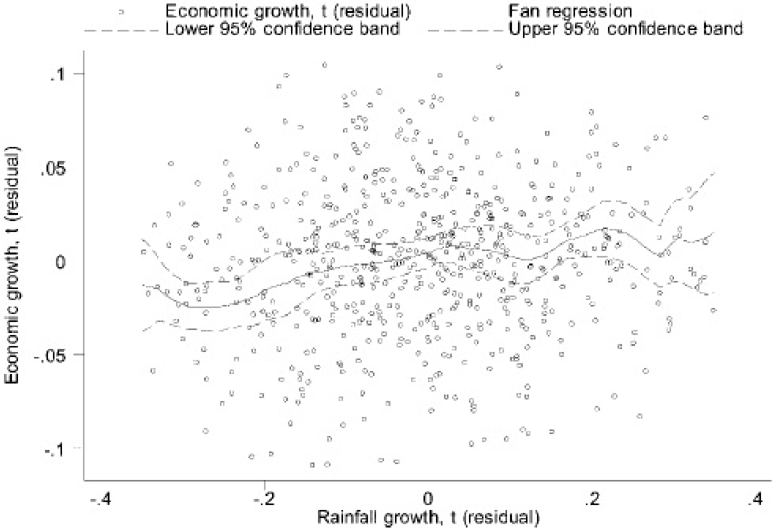


FIG. 1.—Current economic growth rate on current rainfall growth. Nonparametric Fan regression, conditional on country fixed effects and country-specific time trends.

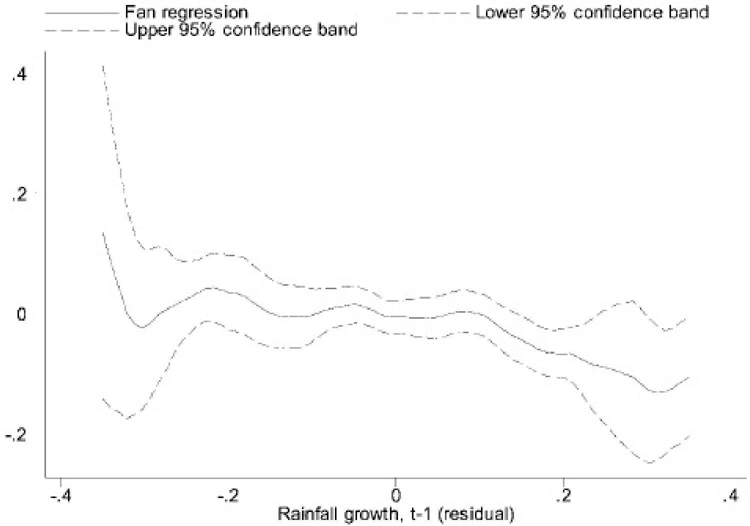
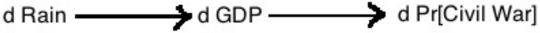
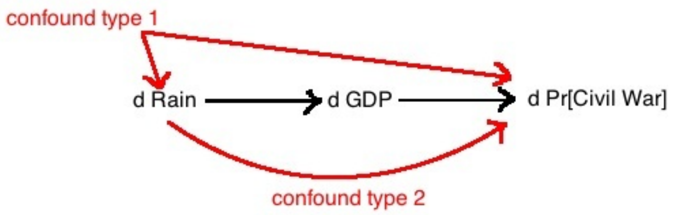
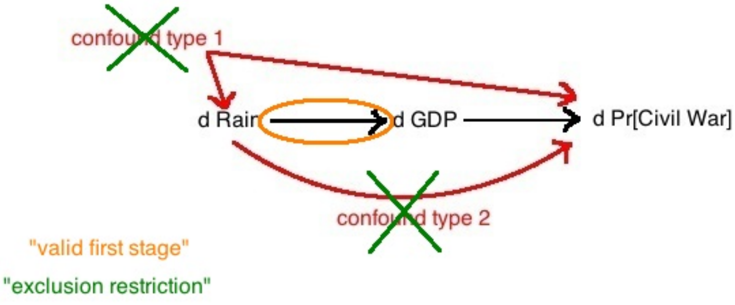


FIG. 2.—Current likelihood of civil conflict (≥25 battle deaths) on lagged rainfall growth. Nonparametric Fan regression, conditional on current rainfall growth, country fixed effects, and country-specific time trends.


d Rain
d GDP
d Pr[Civil War]


confound type 1
d Rain
d GDP
d Pr[Civil War]
confound type 2


confound type 1
d Rain
d GDP
d Pr[Civil War]
confound type 2
"valid first stage"
"exclusion restriction"

Some IV "Greatest Hits"

- Draft lottery numbers $\rightarrow$ military service $\rightarrow$ income (Angrist 1990)
- Quarter of birth $\rightarrow$ schooling $\rightarrow$ income (Angrist & Krueger 1991)
- Election year $\rightarrow$ number of police $\rightarrow$ crime (Levitt 1997)
- Sibling sex composition $\rightarrow$ number of children $\rightarrow$ labor supply (Angrist & Evans 1998)
- Settler mortality $\rightarrow$ investment in institutions $\rightarrow$ avg. income (Acemoglu et al. 2001)
- Rain $\rightarrow$ avg. income $\rightarrow$ civil war (Miguel et al. 2004)

Formal presentation follows MHE (Ch. 4). We start with a simplified regression framework to convey some key intuitions.

"An initial focus on constant effects allows us to explain the mechanics of IV with a minimum of fuss" (115).

Setting

- For treatment condition $S_i = s$, define i's response function in terms of constant effects,

$$Y_{si} = f_i(s) = \alpha + \rho s + \eta_i,$$

and so the outcome we observe is $Y_i = \alpha + \rho S_i + \eta_i$.

Setting

- For treatment condition $S_i = s$, define i's response function in terms of constant effects,

$$Y_{si} = f_i(s) = \alpha + \rho s + \eta_i,$$

and so the outcome we observe is $Y_i = \alpha + \rho S_i + \eta_i$.

- Suppose the idiosyncratic determinants of Y_{si} can be written as,

$$\eta_i = A_i'\gamma + \nu_i,$$

where A_i are potentially observable control variables and γ are coefficients from the population OLS regression, in which case $\text{Cor}(A_i, \nu_i) = 0$ by definition.

Setting

- For treatment condition $S_i = s$, define i's response function in terms of constant effects,

$$Y_{si} = f_i(s) = \alpha + \rho s + \eta_i,$$

  and so the outcome we observe is $Y_i = \alpha + \rho S_i + \eta_i$.
- Suppose the idiosyncratic determinants of Y_{si} can be written as,

$$\eta_i = A_i' \gamma + \nu_i,$$

  where A_i are potentially observable control variables and γ are coefficients from the population OLS regression, in which case $\mathrm{Cor}(A_i, \nu_i) = 0$ by definition.
- Finally, suppose that S_i is assigned such that there is no confounding after controlling for A_i, and so $\mathrm{E}[S_i \nu_i] = 0$.

Setting

Given these assumptions, if we observe A_i, we can use OLS to estimate the slopes in the long regression,

$$Y_i = \alpha + \rho S_i + A_i'\gamma + v_i,$$

which yields a consistent estimate of ρ, the causal effect of going from $S_i = s$ to $s+1$.

Setting

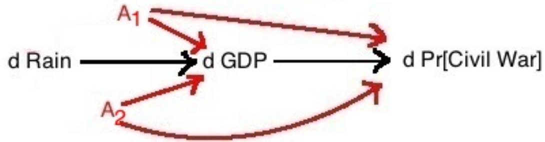
Given these assumptions, if we observe A_i, we can use OLS to estimate the slopes in the long regression,

$$Y_i = \alpha + \rho S_i + A_i' \gamma + \nu_i,$$

which yields a consistent estimate of ρ, the causal effect of going from $S_i = s$ to $s+1$.

But what if we do not observe A_i? Then we cannot use such a control strategy to estimate ρ.

The problem

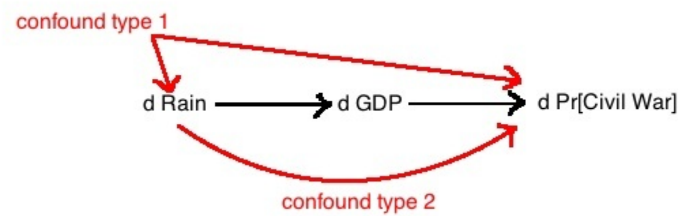


Identification with an instrument

- Suppose we observe some pre-treatment variable Z_i that is correlated with the treatment S_i but not with $\eta_i = A_i'\gamma + \nu_i$:

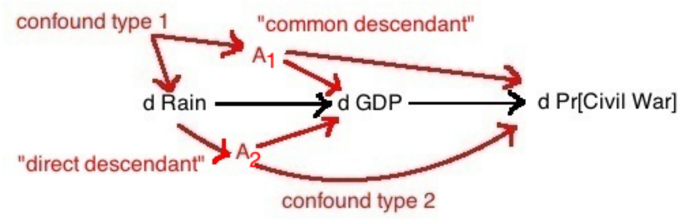
$$\text{Cov}[S_i, Z_i] \neq 0, \text{ but Cov}[\eta_i, Z_i] = 0$$
$$\Rightarrow \text{Cov}[A_i, Z_i] = \text{Cov}[\nu_i, Z_i] = 0.$$

What we cannot have



(Types of confounding that could ruin an instrument, in terms of correlation with v_i.)

What we cannot have



(Types of confounding that could ruin an instrument, in terms of correlation with A_i.)

Identification with an instrument

- So we have Z_i, such that

$$\text{Cov}\left[S_i, Z_i\right] \neq 0, \text{ but } \text{Cov}\left[\eta_i, Z_i\right] = 0$$
$$\Rightarrow \text{Cov}\left[A_i, Z_i\right] = \text{Cov}\left[\nu_i, Z_i\right] = 0.$$

Identification with an instrument

- So we have Z_i, such that

$$\text{Cov}\left[S_i, Z_i\right] \neq 0, \text{ but } \text{Cov}\left[\eta_i, Z_i\right] = 0 \Rightarrow \text{Cov}\left[A_i, Z_i\right] = \text{Cov}\left[\nu_i, Z_i\right] = 0.$$

- Then Z_i is an instrument for S_i and,

$$\text{Cov}\left[Y_i, Z_i\right] = \text{Cov}\left[\alpha + \rho S_i + \eta_i, Z_i\right] = \rho \text{Cov}\left[S_i, Z_i\right]$$

$$\Rightarrow \rho = \frac{\text{Cov}\left[Y_i, Z_i\right]}{\text{Cov}\left[S_i, Z_i\right]} = \frac{\frac{\text{Cov}\left[Y_i, Z_i\right]}{\text{Var}\left[Z_i\right]}}{\frac{\text{Cov}\left[S_i, Z_i\right]}{\text{Var}\left[Z_i\right]}} = \frac{\text{Reduced form}}{\text{First stage}}.$$

Identification with an instrument

- So we have Z_i, such that

$$\text{Cov}[S_i, Z_i] \neq 0, \text{ but } \text{Cov}[\eta_i, Z_i] = 0$$
$$\Rightarrow \text{Cov}[A_i, Z_i] = \text{Cov}[\nu_i, Z_i] = 0.$$

- Then Z_i is an instrument for S_i and,

$$\text{Cov}[Y_i, Z_i] = \text{Cov}[\alpha + \rho S_i + \eta_i, Z_i] = \rho \text{Cov}[S_i, Z_i]$$
$$\Rightarrow \rho = \frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \frac{\frac{\text{Cov}[Y_i, Z_i]}{\text{Var}[Z_i]}}{\frac{\text{Cov}[S_i, Z_i]}{\text{Var}[Z_i]}} = \frac{\text{Reduced form}}{\text{First stage}}.$$

- If both S_i and Z_i are binary, then we get the "Wald estimator," $\rho = \frac{\text{E}[Y_i|Z_i=1] - \text{E}[Y_i|Z_i=0]}{\text{E}[S_i|Z_i=1] - \text{E}[S_i|Z_i=0]}$, which will be important later on.

Identification with an instrument

$$\rho = \frac{\text{Cov}\left[Y_i, Z_i\right]}{\text{Cov}\left[S_i, Z_i\right]} = \frac{\frac{\text{Cov}\left[Y_i, Z_i\right]}{\text{Var}\left[Z_i\right]}}{\frac{\text{Cov}\left[S_i, Z_i\right]}{\text{Var}\left[Z_i\right]}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Note how $\text{Cov}\left[S_i, Z_i\right] \approx 0$ would imply volatility ("weak instrument" or "weak first stage").

Identification with an instrument

$$\rho = \frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \frac{\frac{\text{Cov}[Y_i, Z_i]}{\text{Var}[Z_i]}}{\frac{\text{Cov}[S_i, Z_i]}{\text{Var}[Z_i]}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Note how $\text{Cov}[S_i, Z_i] \approx 0$ would imply volatility ("weak instrument" or "weak first stage").
- If the instrument is "imperfect" $\text{Cov}[\eta_i, Z_i] \neq 0$.

Identification with an instrument

$$\rho = \frac{\mathrm{Cov}[Y_i, Z_i]}{\mathrm{Cov}[S_i, Z_i]} = \frac{\frac{\mathrm{Cov}[Y_i, Z_i]}{\mathrm{Var}[Z_i]}}{\frac{\mathrm{Cov}[S_i, Z_i]}{\mathrm{Var}[Z_i]}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Note how $\mathrm{Cov}[S_i, Z_i] \approx 0$ would imply volatility ("weak instrument" or "weak first stage").
- If the instrument is "imperfect" $\mathrm{Cov}[\eta_i, Z_i] \neq 0$.
- Then $\mathrm{Cov}[Y_i, Z_i] = \rho \mathrm{Cov}[S_i, Z_i] + \mathrm{Cov}[\eta_i, Z_i]$.

Identification with an instrument

$$\rho = \frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \frac{\frac{\text{Cov}[Y_i, Z_i]}{\text{Var}[Z_i]}}{\frac{\text{Cov}[S_i, Z_i]}{\text{Var}[Z_i]}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Note how $\text{Cov}[S_i, Z_i] \approx 0$ would imply volatility ("weak instrument" or "weak first stage").
- If the instrument is "imperfect" $\text{Cov}[\eta_i, Z_i] \neq 0$.
- Then $\text{Cov}[Y_i, Z_i] = \rho \text{Cov}[S_i, Z_i] + \text{Cov}[\eta_i, Z_i]$.
- If we construct $\frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]}$, we won't recover ρ but rather

$$\frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \rho + \frac{\text{Cov}[\eta_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \rho + \frac{\text{Cor}[\eta_i, Z_i]}{\text{Cor}[S_i, Z_i]} \frac{\sigma_\eta}{\sigma_S},$$

which is biased for ρ, where bias is large if $\text{Cor}[S_i, Z_i]$ is small.

Identification with an instrument

$$\rho = \frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \frac{\frac{\text{Cov}[Y_i, Z_i]}{\text{Var}[Z_i]}}{\frac{\text{Cov}[S_i, Z_i]}{\text{Var}[Z_i]}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Note how $\text{Cov}[S_i, Z_i] \approx 0$ would imply volatility ("weak instrument" or "weak first stage").
- If the instrument is "imperfect" $\text{Cov}[\eta_i, Z_i] \neq 0$.
- Then $\text{Cov}[Y_i, Z_i] = \rho \text{Cov}[S_i, Z_i] + \text{Cov}[\eta_i, Z_i]$.
- If we construct $\frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]}$, we won't recover ρ but rather

  $$\frac{\text{Cov}[Y_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \rho + \frac{\text{Cov}[\eta_i, Z_i]}{\text{Cov}[S_i, Z_i]} = \rho + \frac{\text{Cor}[\eta_i, Z_i]}{\text{Cor}[S_i, Z_i]} \frac{\sigma_\eta}{\sigma_S},$$

  which is biased for ρ, where bias is large if $\text{Cor}[S_i, Z_i]$ is small.
- Thus, "strong first stage" is important—"first stage F-stat larger than 10" on (excluded) instruments (Stock & Watson, 2007, based on bias bounds).

Identification with an instrument

We can express this in terms of a set of population regression equations. Let's add some covariates, X_i, to the model, such that,

$$Y_i \equiv X_i'\alpha + \rho S_i + \eta_i.$$

Then define the population regression equations,

$$\begin{aligned} S_i &= X_i'\pi_{10} + \pi_{11} Z_i + \xi_{1i} && \text{first stage} \\ Y_i &= X_i'\pi_{20} + \pi_{21} Z_i + \xi_{2i} && \text{reduced form} \end{aligned}$$

Here, S_i and Y_i are endogenous, while Z_i and X_i are exogenous. Assuming that Z_i is an instrument as defined above,

$$\rho = \frac{\pi_{21}}{\pi_{11}} = \frac{\text{Cov}\left[\tilde{Y}_i, \tilde{Z}_i\right]}{\text{Cov}\left[\tilde{S}_i, \tilde{Z}_i\right]} \quad \text{(by FWL)}$$

First stage:

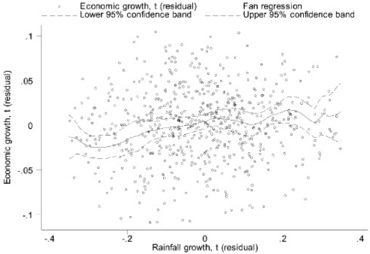


FIG. 1.—Current economic growth rate on current rainfall growth. Nonparametric Fan regression, conditional on country fixed effects and country-specific time trends.

Reduced form:

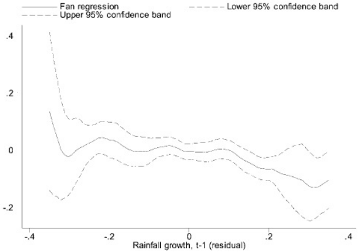


FIG. 2.—Current likelihood of civil conflict (≥25 battle deaths) on lagged rainfall growth. Nonparametric Fan regression, conditional on current rainfall growth, country fixed effects, and country-specific time trends.

Deriving the 2SLS expression

We can substitute first stage into the causal model in order to derive a useful expression:

$$\begin{aligned} Y_i &= X_i'\alpha + \rho(X_i'\pi_{10} + \pi_{11}Z_i + \xi_{1i}) + \eta_i \\ &= X_i'(\alpha + \rho\pi_{10}) + \rho\pi_{11}Z_i + +(\rho\xi_{1i} + \eta_i) \\ &= X_i'\pi_{20} + \pi_{21}Z_i + \xi_{2i} \qquad \text{(from above)} \end{aligned}$$

So $\xi_{2i} = \rho\xi_{1i} + \eta_i$. Plugging that back into the first line yields,

$$\begin{aligned} Y_i &= X_i'\alpha + \rho(X_i'\pi_{10} + \pi_{11}Z_i) + \xi_{2i} \\ &= X_i'\alpha + \rho \mathrm{E}[S_i|X_i, Z_i] + \xi_{2i}, \end{aligned}$$

which shows how we can use Z_i to obtain predicted values of S_i, and then use those predicted values as the regressors in order to estimate ρ.

Deriving the 2SLS expression

We do not know $\mathrm{E}[S_i|X_i, Z_i]$. However, OLS consistently estimates the first stage population coefficients, allowing us to compute consistent predicted values,

$$\hat{S}_i = X_i'\hat{\pi}_{10} + \hat{\pi}_{11} Z_i$$

Adding and subtracting $\rho\hat{S}_i$ to the causal model yields,

$$\begin{aligned} Y_i &= X_i'\alpha + \rho S_i + \eta_i + \rho\hat{S}_i - \rho\hat{S}_i \\ &= X_i'\alpha + \rho\hat{S}_i + [\eta_i + \rho(S_i - \hat{S}_i)], \end{aligned}$$

which is the "two-stage least squares" (2SLS) expression for Y_i.

Deriving the 2SLS expression

With H instruments, $Z_i = (Z_{1i} \dots Z_{Hi})'$, we construct the first stage equation as,

$$S_i = X_i' \pi_{10} + Z_i' \pi_Z + \xi_{1i}.$$

Then *if* each instrument captures the same causal effect, we can proceed as above to obtain a more precise estimate of ρ.

We come back to the issue of "which causal effect an instrument captures" when we move back to the potential outcomes setting.

Deriving the 2SLS expression

In practice, we estimate 2SLS in one step. This helps to ensure that we estimate the variance of the coefficients correctly.

Some gratuitous matrix algebra...

Deriving the 2SLS expression

...let $\mathbf{S}$ be endogenous variables, $\hat{\mathbf{S}}$ predicted values, $\mathbf{X}$ regressors, and $\mathbf{Z}$ instruments.

$$\hat{\Gamma}_{2SLS} = (\underbrace{[\hat{\mathbf{S}}\mathbf{X}]'[\hat{\mathbf{S}}\mathbf{X}]}_{A})^{-1} \underbrace{[\hat{\mathbf{S}}\mathbf{X}]'Y}_{B},$$

where $\hat{\Gamma}_{2SLS}$ includes estimate of ρ and other coefficients in regression model.

Deriving the 2SLS expression

...let $\mathbf{S}$ be endogenous variables, $\hat{\mathbf{S}}$ predicted values, $\mathbf{X}$ regressors, and $\mathbf{Z}$ instruments.

$$\hat{\Gamma}_{2SLS} = \underbrace{([\hat{\mathbf{S}}\mathbf{X}]'[\hat{\mathbf{S}}\mathbf{X}])}_{A}^{-1} \underbrace{[\hat{\mathbf{S}}\mathbf{X}]'Y}_{B},$$

where $\hat{\Gamma}_{2SLS}$ includes estimate of ρ and other coefficients in regression model.Consider A:

$$\begin{aligned}
[\hat{\mathbf{S}}\mathbf{X}]'[\hat{\mathbf{S}}\mathbf{X}] &= \left[[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'\mathbf{S} \quad \mathbf{X}\right]' \left[[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'\mathbf{S} \quad \mathbf{X}\right] \\
&= [\mathbf{M}'\mathbf{S} \quad \mathbf{X}]'[\mathbf{M}'\mathbf{S} \quad \mathbf{X}] \qquad (\mathbf{M} \text{ idempotent}) \\
&= \begin{bmatrix} \mathbf{SMM'S} & \mathbf{S'MX} \\ \mathbf{X'M'S} & \mathbf{X'X} \end{bmatrix} = \begin{bmatrix} \mathbf{SMS} & \mathbf{S'MX} \\ \mathbf{X'M'S} & \mathbf{X'MX} \end{bmatrix} \qquad (\text{by } \mathbf{MX} = \mathbf{X}) \\
&= [\mathbf{SX}]'\mathbf{M}[\mathbf{SX}] = [\mathbf{SX}]'[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'[\mathbf{SX}].
\end{aligned}$$

Similar for B.

Estimating 2SLS models

Putting it together yields,

$$\hat{\Gamma}_{2SLS} = \left([\mathbf{SX}]'[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'[\mathbf{SX}]\right)^{-1}[\mathbf{SX}]'[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'Y$$
$$\equiv (\mathbf{W}'\mathbf{M}\mathbf{W})^{-1}\mathbf{W}'\mathbf{M}Y,$$

where **W** is the matrix of regressors in the causal model, and **M** is the hat matrix from the first stage. This is how regression software actually computes 2SLS.

Estimating 2SLS models

Putting it together yields,

$$\hat{\Gamma}_{2SLS} = \left([\mathbf{SX}]'[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'[\mathbf{SX}]\right)^{-1}[\mathbf{SX}]'[\mathbf{ZX}]([\mathbf{ZX}]'[\mathbf{ZX}])^{-1}[\mathbf{ZX}]'Y$$
$$\equiv (\mathbf{W}'\mathbf{M}\mathbf{W})^{-1}\mathbf{W}'\mathbf{M}Y,$$

where $\mathbf{W}$ is the matrix of regressors in the causal model, and $\mathbf{M}$ is the hat matrix from the first stage. This is how regression software actually computes 2SLS.

The command in Stata is `ivregress 2sls...` with the usual robust or cluster-robust standard error methods available. In R use `ivreg(...)` in the `AER` package, and then you can use the `sandwich` or `clubSandwich` packages for robust and cluster-robust standard errors.

Standard errors for $\hat{\Gamma}_{2SLS}$

Following the presentation in MHE, let $V_i \equiv [X_i' \;\; \hat{S}_i]'$. Then,

$$\begin{aligned}
\hat{\Gamma}_{2SLS} = (\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'Y &= \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1} \sum_{i=1}^{N} V_i Y_i \\
&= \Gamma + \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1} \sum_{i=1}^{N} V_i[\eta_i + \rho(S_i - \hat{S}_i)] \\
&= \Gamma + \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1} \sum_{i=1}^{N} V_i \eta_i, \qquad (1)
\end{aligned}$$

by the fact that $\rho(S_i - \hat{S}_i)$ are orthogonal to V_i by the OLS solution in the first stage.

Standard errors for $\hat{\Gamma}_{2SLS}$

Asymptotically,

$$\hat{S}_i = X_i'\hat{\pi}_{10} + Z_i'\hat{\pi}_Z \xrightarrow{p} X_i'\pi_{10} + Z_i'\pi_Z,$$

which is fixed. Therefore, a consistent estimator for the variance of $\hat{\Gamma}_{2SLS}$ applies the similar results as we used for OLS.

Standard errors for $\hat{\Gamma}_{2SLS}$

Asymptotically,

$$\hat{S}_i = X_i'\hat{\pi}_{10} + Z_i'\hat{\pi}_Z \xrightarrow{p} X_i'\pi_{10} + Z_i'\pi_Z,$$

which is fixed. Therefore, a consistent estimator for the variance of $\hat{\Gamma}_{2SLS}$ applies the similar results as we used for OLS.With no clustering, this yields the following consistent estimator for the variance,

$$\widehat{\text{Var}}[\hat{\Gamma}_{2SLS}] = \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1} \left(\sum_{i=1}^{N} V_i V_i' \hat{\eta}_i^2\right) \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1},$$

which looks a lot like our heteroskedasticity-robust estimator for OLS. With clustering we would construct the analogue to the cluster-robust standard error, substituting in V_i and $\hat{\eta}_i$.

Standard errors for $\hat{\Gamma}_{2SLS}$

Asymptotically,

$$\hat{S}_i = X_i'\hat{\pi}_{10} + Z_i'\hat{\pi}_Z \xrightarrow{p} X_i'\pi_{10} + Z_i'\pi_Z,$$

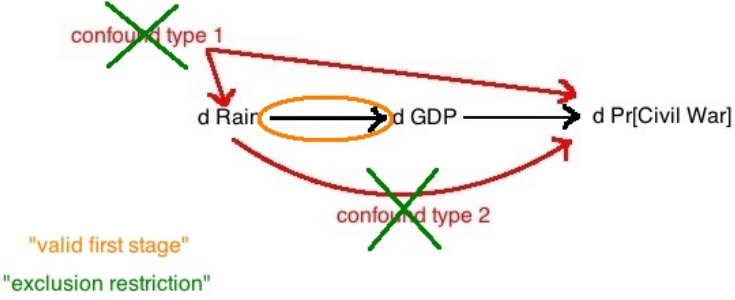
which is fixed. Therefore, a consistent estimator for the variance of $\hat{\Gamma}_{2SLS}$ applies the similar results as we used for OLS.With no clustering, this yields the following consistent estimator for the variance,

$$\widehat{\text{Var}}[\hat{\Gamma}_{2SLS}] = \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1} \left(\sum_{i=1}^{N} V_i V_i' \hat{\eta}_i^2\right) \left(\sum_{i=1}^{N} V_i V_i'\right)^{-1},$$

which looks a lot like our heteroskedasticity-robust estimator for OLS. With clustering we would construct the analogue to the cluster-robust standard error, substituting in V_i and $\hat{\eta}_i$. Note that the $\hat{\eta}_i$ are estimated by plugging the original S_i back into the second stage with the 2SLS coefficients, getting $\hat{y}_{i,2SLS}$ from that, and then getting residuals.

Remarks

- The key assumptions are
  1. exclusion, which embeds both exogeneity and "no-alternative-pathways" assumptions, and
  2. valid first stage.


confound type 1
d Rain
d GDP
d Pr[Civil War]
confound type 2
"valid first stage"
"exclusion restriction"

Remarks

- The key assumptions are
  1. exclusion, which embeds both exogeneity and "no-alternative-pathways" assumptions, and
  2. valid first stage.
- Any IV application needs to make the case for both of these!

Remarks

- The key assumptions are
  1. exclusion, which embeds both exogeneity and "no-alternative-pathways" assumptions, and
  2. valid first stage.
- Any IV application needs to make the case for both of these!
- First stage can be checked directly. Exclusion cannot.
- MHE (p. 131) provide some useful indirect tests, known as "placebo tests", for exclusion:

Remarks

- The key assumptions are
  1. exclusion, which embeds both exogeneity and "no-alternative-pathways" assumptions, and
  2. valid first stage.
- Any IV application needs to make the case for both of these!
- First stage can be checked directly. Exclusion cannot.
- MHE (p. 131) provide some useful indirect tests, known as "placebo tests", for exclusion:
  - Look for associations between Z_i and things that predict Y_i but should not be sensitive to Z_i (e.g., things that were determined prior to Z_i).

Remarks

- The key assumptions are
  1. exclusion, which embeds both exogeneity and "no-alternative-pathways" assumptions, and
  2. valid first stage.
- Any IV application needs to make the case for both of these!
- First stage can be checked directly. Exclusion cannot.
- MHE (p. 131) provide some useful indirect tests, known as "placebo tests", for exclusion:
  - Look for associations between Z_i and things that predict Y_i but should not be sensitive to Z_i (e.g., things that were determined prior to Z_i).
  - Find subpopulations for which there should be no relationship between Z_i and S_i. If Z_i predicts Y_i in this subsample, then you may have an exclusion violation.

Remarks

Next we will move back to the potential outcomes framework which loosens the constant effects assumption.

We will discover that little changes in terms of how we do estimation, but there are important refinements in how we interpret IV estimates.