Instrumental Variables

Annie Chen

February 26, 2020


Yeah, I'm using lagged variables as instruments
It's fine, I've tested the exclusion restriction
The IV correlates significantly with the treatment but not with the outcome

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z, d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z, d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction
  - $Y_i(z=1,d) = Y_i(z=0,d)$ for $d = 0,1$

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction
  - $Y_i(z=1,d) = Y_i(z=0,d)$ for $d = 0, 1$
- Relevance

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction
  - $Y_i(z=1,d) = Y_i(z=0,d)$ for $d = 0,1$
- Relevance
  - $Cov(D_i, Z_i) \neq 0$

IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction
  - $Y_i(z=1,d) = Y_i(z=0,d)$ for $d = 0,1$
- Relevance
  - $Cov(D_i, Z_i) \neq 0$
- Monotonicity

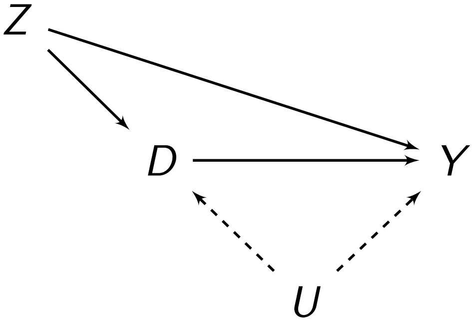
IV Assumptions

Which are directly testable from data?

- Ignorability (Exogeneity)
  - $\{Y_i(z,d), D_i(1), D_i(0)\} \perp\!\!\!\perp Z_i$
- SUTVA
  - $D(z) = D(z')$ if $z = z'$
  - $Y(z,d) = Y(z',d')$
- Exclusion Restriction
  - $Y_i(z=1,d) = Y_i(z=0,d)$ for $d = 0,1$
- Relevance
  - $Cov(D_i, Z_i) \neq 0$
- Monotonicity
  - $D_i(1) \geq D_i(0)$ for all i

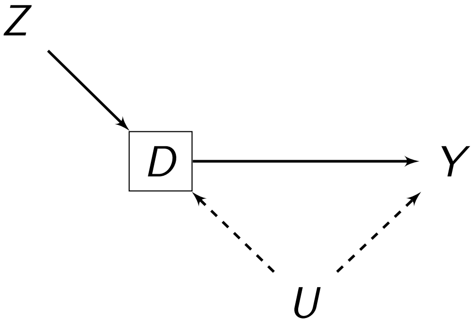
Valid Instruments

- Let's play a game: is Z a valid instrument that identifies a causal effect of D on Y?



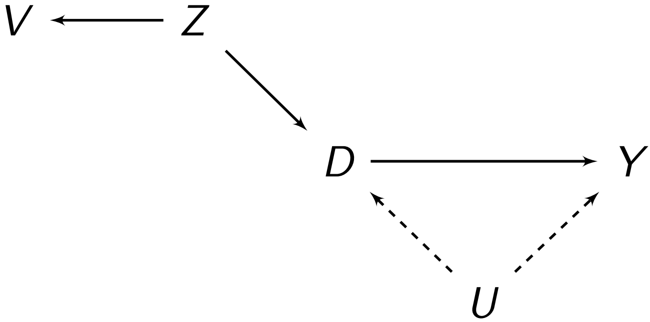
Valid Instruments

- Can we test the exclusion restriction by checking the association between Z and Y after conditioning on D?



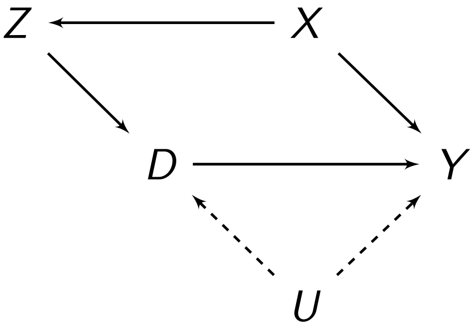
Valid Instruments

- and here? Suppose the instrument Z is unobserved, but we have a proxy V.

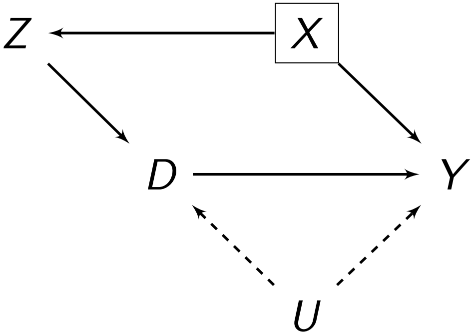


Valid Instruments

- Can we make this one work?

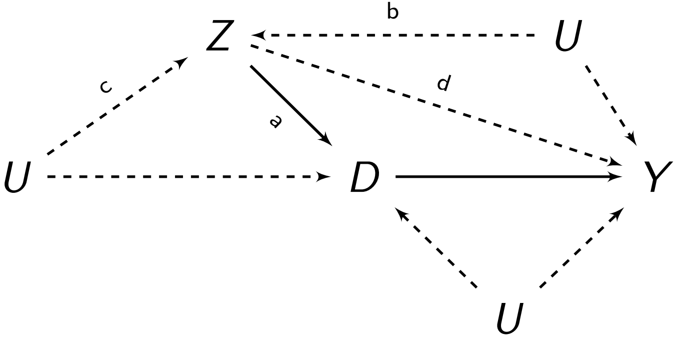


- Can we make this one work?[1]



[1] Note that this is now a conditional instrument and we adjust our assumptions accordingly. I.e., $\{Y(z,d), D(1), D(0)\} \perp Z|X$

Violating IV Assumptions



Review LATE Framework

- For unit i, if $Z_i \in \{0, 1\}$ is the instrument (encouragement) and $D_i \in \{0, 1\}$ is treatment, then we have 4 principal strata (latent class variable C).

C_i	$D_i(z)$	
	$Z_i = 1$	$Z_i = 0$
Compliers	1	0
Always-takers	1	1
Never-takers	0	0
Defiers	0	1

Review LATE Framework

- For unit i, if $Z_i \in \{0, 1\}$ is the instrument (encouragement) and $D_i \in \{0, 1\}$ is treatment, then we have 4 principal strata (latent class variable C).

C_i	$D_i(z)$	
	$Z_i = 1$	$Z_i = 0$
Compliers	1	0
Always-takers	1	1
Never-takers	0	0
Defiers	0	1

- Can we determine which C unit i falls under from our observed data?

Review LATE Framework

- For unit i, if $Z_i \in \{0, 1\}$ is the instrument (encouragement) and $D_i \in \{0, 1\}$ is treatment, then we have 4 principal strata (latent class variable C).

C_i	$D_i(z)$	
	$Z_i = 1$	$Z_i = 0$
Compliers	1	0
Always-takers	1	1
Never-takers	0	0
Defiers	0	1

- Can we determine which C unit i falls under from our observed data?
- Say $Z_i = 1$ and $D_i = 0$. What type(s) could unit i be?

IV Estimand and 2SLS Estimator

- Intent-to-treat (ITT) effect: $\mathbb{E}[Y_i(z=1) - Y_i(z=0)]$

IV Estimand and 2SLS Estimator

- Intent-to-treat (ITT) effect: $\mathbb{E}[Y_i(z=1) - Y_i(z=0)]$
- ITT_d: $\mathbb{E}[D_i(z=1) - D_i(z=0)]$

IV Estimand and 2SLS Estimator

- Intent-to-treat (ITT) effect: $\mathbb{E}[Y_i(z=1) - Y_i(z=0)]$
- ITT_d: $\mathbb{E}[D_i(z=1) - D_i(z=0)]$
- Local Average Treatment Effect (LATE):[2] $= \frac{ITT}{ITT_d}$

[2] also known as Complier Average Causal Effect (CACE) ref. GG p.144 and p.180

IV Estimand and 2SLS Estimator

- Intent-to-treat (ITT) effect: $\mathbb{E}[Y_i(z=1) - Y_i(z=0)]$
- ITT_d: $\mathbb{E}[D_i(z=1) - D_i(z=0)]$
- Local Average Treatment Effect (LATE):[2] $= \frac{ITT}{ITT_d}$
- Can be estimated with 2SLS assuming monotonicity and excludability.

[2] also known as Complier Average Causal Effect (CACE) ref. GG p.144 and p.180

Weak Instruments

- If instrument(s) are so weak that there is no first stage, the 2SLS distribution is centered on the plim of OLS.

Weak Instruments

- If instrument(s) are so weak that there is no first stage, the 2SLS distribution is centered on the plim of OLS.
- Run a joint-significance test (f-test) under the null that your instrument(s) is/are weak.

Weak Instruments

- If instrument(s) are so weak that there is no first stage, the 2SLS distribution is centered on the plim of OLS.
- Run a joint-significance test (f-test) under the null that your instrument(s) is/are weak.
- The *rule of thumb* given by Staiger and Stock (1997) is a first-stage F-statistic > 10 (in which case, we may reject the null and conclude the instrument is relevant).[3]

[3] Stock and Yogo (2005) tabulate the critical values needed to reject null for some tolerable level of relative bias (between OLS and 2SLS estimator).

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate
  - You need one instrument, Z, to *just*-idenitify the IV model.

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate
  - You need one instrument, Z, to *just*-idenitify the IV model.

- Now suppose you have >1 endogenous regressors:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... \beta_k X_k + \gamma W + u$$

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate
  - You need one instrument, Z, to *just*-idenitify the IV model.

- Now suppose you have >1 endogenous regressors:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_k + \gamma W + u$$

  - $X_1, X_2, ..., X_k$ are your k endogenous regressors

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate
  - You need one instrument, Z, to *just*-idenitify the IV model.

- Now suppose you have >1 endogenous regressors:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k + \gamma W + u$$

  - $X_1, X_2, \ldots, X_k$ are your k endogenous regressors
  - And you have $Z_1, \ldots, Z_m$ instruments

Under Identification

- Consider this regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u$$

  - X is your endogenous regressor
  - W is an exogenous covariate
  - You need one instrument, Z, to *just*-idenitify the IV model.

- Now suppose you have >1 endogenous regressors:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_k + \gamma W + u$$

  - $X_1, X_2, ..., X_k$ are your k endogenous regressors
  - And you have $Z_1, ..., Z_m$ instruments
  - Where $m < k$ = underidenfication, $m > k$ = overid, $m = k$ is just-identified.

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.
- Sargan-Hansen Test ($\mathcal{J}$-test): the test statistic is $\mathcal{J} = mF$ where F is the F-statistic for the null that coefficients are jointly zero in $\hat{u} = \beta_0 + \beta_1 Z_1 + ... + \beta_m Z_m + \gamma W + \nu$.

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.
- Sargan-Hansen Test ($\mathcal{J}$-test): the test statistic is $\mathcal{J} = mF$ where F is the F-statistic for the null that coefficients are jointly zero in $\hat{u} = \beta_0 + \beta_1 Z_1 + ... + \beta_m Z_m + \gamma W + \nu$.
- $\mathcal{J} \sim \mathcal{X}^2_{m-k}$ (k is number of endogenous regressors, m is number of instruments)

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.
- Sargan-Hansen Test ($\mathcal{J}$-test): the test statistic is $\mathcal{J} = mF$ where F is the F-statistic for the null that coefficients are jointly zero in $\hat{u} = \beta_0 + \beta_1 Z_1 + ... + \beta_m Z_m + \gamma W + \nu$.
- $\mathcal{J} \sim \mathcal{X}^2_{m-k}$ (k is number of endogenous regressors, m is number of instruments)
- Null hypothesis is that all instruments are uncorrelated with $\hat{u}$

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.
- Sargan-Hansen Test ($\mathcal{J}$-test): the test statistic is $\mathcal{J} = mF$ where F is the F-statistic for the null that coefficients are jointly zero in $\hat{u} = \beta_0 + \beta_1 Z_1 + ... + \beta_m Z_m + \gamma W + \nu$.
- $\mathcal{J} \sim \mathcal{X}^2_{m-k}$ (k is number of endogenous regressors, m is number of instruments)
- Null hypothesis is that all instruments are uncorrelated with $\hat{u}$
- Assumes homoskedasticity[4]

[4] if heteroskedastic, see Kleibergen–Paap rk Wald statistic

Over Identification

- When you have more instruments than endogenous regressors, you can conduct an overidentification test.
- Sargan-Hansen Test ($\mathcal{J}$-test): the test statistic is $\mathcal{J} = mF$ where F is the F-statistic for the null that coefficients are jointly zero in $\hat{u} = \beta_0 + \beta_1 Z_1 + ... + \beta_m Z_m + \gamma W + \nu$.
- $\mathcal{J} \sim \mathcal{X}^2_{m-k}$ (k is number of endogenous regressors, m is number of instruments)
- Null hypothesis is that all instruments are uncorrelated with $\hat{u}$
- Assumes homoskedasticity[4]
- Is this really a test for exogeneity?

[4] if heteroskedastic, see Kleibergen–Paap rk Wald statistic

AER::IVREG()

```r
ivmod <- ivreg(y ~ x1 + w1 | w1 + z1 + z2, data)
summary(ivmod, diagnostics = TRUE)$diagnostics
```

```
##                    df1 df2  statistic      p-value
## Weak instruments     2  60 17.4114496 1.089283e-06
## Wu-Hausman           1  60 34.3521510 2.093826e-07
## Sargan               1  NA  0.1457748 7.026062e-01
```

ADDITIONAL NOTES

Limited Information Maximum Likelihood (LIML)

- Alternative to 2SLS is LIML estimator, which has better small sample properties than 2SLS