

The Rubin Causal Model and Instrumental Variables

Charlie Gibbons
ARE 212

Spring 2011

Notes

Outline

- 1 Potential outcomes framework
- 2 Instrumental variables
 - Framework
 - Assumption 1: SUTVA
 - Assumption 2: Ignorable treatment assignment
 - Assumption 3: Exclusion restriction
 - Assumption 4: Inclusion restriction
 - Assumption 5: Monotonicity
 - LATE
 - Summary
- 3 Estimation

Notes

Potential outcomes framework

There are two possible outcomes: $Y_i(1)$ if individual i undergoes treatment T and $Y_i(0)$ if he does not. T_i is an indicator of treatment status. The treatment effect for i is $\tau_i = Y_i(1) - Y_i(0)$. Hence,

$$Y_i = (1 - T_i)Y_i(0) + T_iY_i(1) = Y_i(0) + T_i\tau_i.$$

Notice that we are not assuming any error in Y_i . But there is a distribution of τ_i .

One way to connect to linear model: LPM for $Y_i(0)$ plus a treatment dummy.

Notes

Fundamental problem of causal inference

Why not just estimate this directly?

We only observe *either* $Y_i(0)$ or $Y_i(1)$, not both. This is the *fundamental problem of causal inference*.

Notes

SUTVA

We have already made an important assumption: observation i 's outcome only depends upon his treatment status—not anyone else's. This rules out:

- General equilibrium effects (doubling one person's income while keeping everyone else's the same versus doubling everyone's)
- Interaction effects/network effects/spill-overs/externalities (not vaccinating one person when everyone else is versus the opposite)

This is known as the *stable unit treatment value assumption* (SUTVA).

Notes

Treatment effects

The two main effects of interest are the *average treatment effect* (ATE) and the *average treatment effect for the treated* (ATT).

$$\begin{aligned} \text{ATE} &= \mathbb{E}(\tau_i) \\ \text{ATT} &= \mathbb{E}(\tau_i \mid \underbrace{T_i = 1}_{\text{Distribution of treated}}) \\ &= \mathbb{E}_1(\tau_i) \end{aligned}$$

Why are these different? Selection into treatment. These effects are often conditioned on X :

$$\begin{aligned} \text{ATE}(x) &= \mathbb{E}(\tau_i \mid X = x) \\ \text{ATT}(x) &= \mathbb{E}_1(\tau_i \mid X = x) \end{aligned}$$

Notes

Under treatment independence

Of course, τ_i is not observable since both outcomes cannot be realized for each individual.

Suppose that treatment is *randomly assigned*:

$$(Y_i(0), Y_i(1)) \perp T_i,$$

then the distributions of outcomes in both the entire population and among the treated are the same. Hence, the ATE and ATT will be equal. This result stems from random treatment assignment.

Notes

Conditional expectations

By the law of iterated expectations and independence,

$$\begin{aligned} \text{ATT} &= \mathbb{E}_1 [Y_i(1) - Y_i(0)] \\ &= \mathbb{E} [Y_i(1) - Y_i(0)] = \text{ATE} \\ \mathbb{E}_0(Y) &= \mathbb{E}_0 [(1 - T_i)Y_i(0) + T_iY_i(1)] \\ &= \mathbb{E}_0 [Y_i(0)] = \mathbb{E} [Y_i(0)] \\ \mathbb{E}_1(Y) &= \mathbb{E}_1 [(1 - T_i)Y_i(0) + T_iY_i(1)] \\ &= \mathbb{E}_1 [Y_i(1)] = \mathbb{E} [Y_i(1)] \end{aligned}$$

Hence,

$$\begin{aligned} \text{ATE} &= \mathbb{E} [Y_i(1) - Y_i(0)] = \mathbb{E} [Y_i(1)] - \mathbb{E} [Y_i(0)] \\ &= \mathbb{E}_1 [Y_i(1)] - \mathbb{E}_0 [Y_i(0)] \end{aligned}$$

Notes

Estimation

Therefore,

$$\text{ATE} = \text{ATT} = \mathbb{E}_1 [Y_i(1)] - \mathbb{E}_0 [Y_i(0)]$$

These parameters can be estimated by

$$\frac{1}{\#\{i : T_i = 1\}} \sum_{\{i : T_i = 1\}} Y_i - \frac{1}{\#\{i : T_i = 0\}} \sum_{\{i : T_i = 0\}} Y_i$$

This is just a simple difference in means between the two groups.

Notes

Under conditional independence

How can this problem be solved in the absence of random treatment assignment?

We need the assumption of *unconfoundedness*:

$$(Y_i(0), Y_i(1)) \perp T_i | X_i$$

This assumption intuitively states that, if we observe X , we are able to determine all the ways in which the treatment group differs from the control; the X covariates explain treatment assignment completely. For this reason, this assumption is also known as *selection on observables*.

Notes

Confirmation of previous results

Note that all the results shown in the previous set of slides, namely that the ATT equals the ATE, remain after conditioning on X . Specifically,

$$ATE(x) = ATT(x) = \mathbb{E}_1 [Y | X = x] - \mathbb{E}_0 [Y | X = x]$$

Hence, to find $ATT(x)$, find the average value of Y for those members of the treated group with $X = x$ and subtract it from the average Y for the control population with $X = x$.

We can use the linear model for Y to estimate this quantity (see the opening slide).

Notes

IV setup

Thus far we have only considered treatment status, but now we want to distinguish between treatment assignment and treatment status; the former is prescribed by the researcher and the latter is the actual treatment taken up by the individual.

These two measures would be the same under perfect compliance. But compliance is a choice made by the individual based upon unobservable factors—specifically, we have non-ignorable treatment status.

Notes

Framework

Z_i is an instrument taking values 0 or 1 (*e.g.*, were you drafted into the military?)

\mathbf{Z} is an N -dimensional vector of treatment assignments to all individuals

$D_i(\mathbf{Z})$ is the treatment status of individual i taking values 0 or 1 (*i.e.*, there is no partial compliance; *e.g.*, did you serve?)

$\mathbf{D}(\mathbf{Z})$ is the treatment status of all individuals

If there was perfect compliance, then $\mathbf{D}(\mathbf{Z}) = \mathbf{Z}$.

$Y_i(\mathbf{Z}, \mathbf{D}(\mathbf{Z}))$ is the response of individual i given all treatment assignments and statuses.

$\mathbf{Y}(\mathbf{Z}, \mathbf{D}(\mathbf{Z}))$ is the responses of all individuals

Since compliance isn't perfect, both $Y_i(\mathbf{Z}, \mathbf{D}(\mathbf{Z}))$ and $D_i(\mathbf{Z})$ are potential outcomes.

Notes

SUTVA

Assumption 1: Stable Unit Treatment Value Assumption (SUTVA)

Both potential outcomes for individual i are independent of the treatment assignment to other individuals.

- If an individual has the same value of the instrument in two different treatment assignments, then his treatment status will be the same under both assignments; if $Z_i = Z'_i$, then $D_i(\mathbf{Z}) = D_i(\mathbf{Z}')$.
- If an individual has the same treatment assignment and status under two different treatment assignments, then his response will be the same under both assignments; if $Z_i = Z'_i$ and $D_i(\mathbf{Z}) = D_i(\mathbf{Z}')$, then $Y_i(\mathbf{Z}, \mathbf{D}(\mathbf{Z})) = Y_i(\mathbf{Z}', \mathbf{D}(\mathbf{Z}'))$.

Notes

SUTVA, continued

SUTVA permits us to write $D_i(\mathbf{Z}) = D_i(Z_i)$ and $Y_i(\mathbf{Z}, \mathbf{D}(\mathbf{Z})) = Y_i(Z_i, D_i(Z_i))$.

Notes

Randomization

Assumption 2: Ignorable assignment of \mathbf{Z}

$$0 < \Pr(Z_i = 1 | X_i = x) = \Pr(Z_j = 1 | X_j = x) < 1 \quad \forall i, j.$$

Notes

Exclusion restriction

Assumption 3: Exclusion restriction

The instrument can only effect the outcome by altering treatment status.

$$\mathbf{Y}(\mathbf{Z}, \mathbf{D}) = \mathbf{Y}(\mathbf{Z}', \mathbf{D}) \quad \forall \mathbf{Z}, \mathbf{Z}', \mathbf{D}$$

That is, the effect of Z on Y must be solely through the effect of Z on D and a change in the value of the instrument does not effect the outcome unless it changes treatment status.

The combination of the exclusion restriction, ignorable treatment effects, and the linearity assumption imply that Z_i is uncorrelated with both error terms: $\mathbb{E}[Z_i \nu_i] = \mathbb{E}[Z_i \epsilon_i] = 0$.

Notes

Simplification

This lets us write $\mathbf{Y}(\mathbf{Z}, \mathbf{D})$ as $\mathbf{Y}(\mathbf{D})$.

Assumption 1 lets us write $Y_i(\mathbf{D})$ as $Y_i(D_i)$.

Hence, the causal effect of D on Y is $Y_i(1) - Y_i(0)$, the standard result of the Rubin causal model framework.

Notes

Inclusion restriction

Assumption 4: Inclusion restriction

The instrument has a non-zero average effect on treatment status.

$$\mathbb{E}[D_i(1) - D_i(0)|X] \neq 0$$

This can also be expressed as $\text{Cov}(Z, D | X) \neq 0$ —the instrument must change the behavior of at least some people.

The inclusion restriction considers the effect of the assignment mechanism on treatment status, while the exclusion restriction considers the effect of the assignment mechanism on the outcome.

This assumption can be verified by testing the coefficient on the instrument in the so-called *first stage regression* of the treatment on the instrument.

Notes

Monotonicity

Assumption 5: Monotonicity

The instrument effects all individuals' treatment statuses in the same direction.

$$D_i(1) \geq D_i(0) \text{ or } D_i(1) \leq D_i(0) \quad \forall i$$

Combined with Assumption 4, this equality must be strict for some individual i .

Notes

Respondent types

		$D_i(0)$	
		0	1
$D_i(1)$	0	Never-taker	Defier
	1	Complier	Always-taker

The monotonicity assumption rules out defiers.

The behavior of never-takers and always-takers does not change in response to a change in the instrument Z and treatment status is always the same. Because only one status is ever observed, we can't estimate the effect of a change in status for these groups.

Thus, the effect of Z on Y is identified solely by this effect for compliers.

Notes

LATE

Using the definition of conditional expectations, the IV estimator can be written as

$$\frac{\mathbb{E}[Y_i(1, D_i(1))|X] - \mathbb{E}[Y_i(0, D_i(0))|X]}{\mathbb{E}[D_i(1) - D_i(0)|X]} \\ = \mathbb{E}[Y_i(1, D_i(1)) - Y_i(0, D_i(0)) | D_i(1) - D_i(0) = 1, X]$$

This is called the *local average treatment effect*, where “local” refers to the fact that the estimate is the average treatment effect for the group of compliers only.

Note that this group is *unobservable*.

Notes

Definition of an instrument

A covariate Z is an instrument for the causal effect of D on Y if:

- SUTVA holds
- It is randomly (or ignorably) assigned
- It satisfies the exclusion restriction
- Its average effect on D is non-zero (*i.e.*, it satisfies the inclusion restriction)
- It satisfies the monotonicity assumption

Notes

Estimating equations

The following two equations define the instrumental variables framework:

$$X = Z\gamma + \nu \\ Y = X\beta + \epsilon$$

Note that the Z matrix contains the instruments and the non-endogenous X regressors. The instruments do not appear in the outcome regression.

Notes

Just-identified case

The simplest case is when there are the same number of endogenous regressors as there are instruments (the just-identified case). Then we can solve the problem as follows:

$$\begin{aligned} Y &= X\hat{\beta} + \epsilon \\ Z'Y &= Z'X\hat{\beta} + Z'\epsilon \\ Z'Y &= Z'X\hat{\beta} \\ \hat{\beta} &= (Z'X)^{-1}Z'Y \end{aligned}$$

Notes

Over-identified case

If there are more instruments than there are endogenous regressors (the over-identified case), then we can follow a two-stage least squares process:

- 1 Calculate $\hat{\gamma} = (Z'Z)^{-1}Z'X$.
- 2 Calculate $\hat{X} = Z\hat{\gamma} = Z(Z'Z)^{-1}Z'X$.
- 3 Use \hat{X} instead of X in the outcome equation:
 $Y = \hat{X}\beta = Z(Z'Z)^{-1}Z'X\beta$.
- 4 The estimate of β becomes:

$$\begin{aligned} \hat{\beta} &= (\hat{X}'\hat{X})^{-1}\hat{X}'Y \\ &= [X'Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'X]^{-1}X'Z(Z'Z)^{-1}Z'Y \\ &= [X'Z(Z'Z)^{-1}Z'X]^{-1}X'Z(Z'Z)^{-1}Z'Y \end{aligned}$$

Notes

Consistency

Instrumental variables use only part of the variability in the endogenous regressor—the part that is correlated with the instrument and uncorrelated with omitted variables—to identify the effect of that regressor.

IV estimates are built around estimates of \hat{X} , rather than a fixed, “true” value X . As a result, IV estimates have finite sample bias. This bias increases with the number of instruments employed, but decreases when these instruments are highly correlated with the endogenous regressors and as the sample size increases. The estimates are consistent, however.

Notes
