

Mostly Harmless Econometrics

Reading Group Discussion Guide

Chapter 4: Instrumental Variables

Chapter 4.1-4.2: Recap of 2SLS

Notation

- Structural Equation: $y = \alpha + \beta w + \gamma x + u$
- Reduced Form: $y = \theta_0 + \theta_1 z + \theta_2 x + v$
- First Stage: $w = \pi_0 + \pi_1 z + \pi_2 x + e$
- Dependent / Outcome / LHS Variable: y
- Endogenous Variable: w
- Included Exogenous (Control) Variables: x
- Excluded Exogenous Variables / Instrument: z

Basic IV Model

- Structural Equation: $y = \alpha + \beta w + \gamma x + u$
- First Stage: $w = \pi_0 + \pi_1 z + \pi_2 x + e$
- Endogeneity: $\text{Cov}(w, u | x) = \rho > 0$
- Instrument Relevance: $\text{Cov}(w, z | x) \neq 0$
- Instrument Exogeneity: $\text{Cov}(z, u | x) = 0$

Note, I am not sure what the name of this set up of assumptions is called. I have seen it called a 'Linear Triangular System' but I am not exactly sure why. I think it suffices to call this a Linear IV model – most people will know what you are talking about.

Why OLS is Wrong

Assume that $\gamma = 0$ (so no x variable for ease of writing):

$$\beta_{OLS} = \frac{\text{Cov}(w, y)}{\text{Var}(w)} = \frac{\text{Cov}(w, \alpha + \beta w + u)}{\text{Var}(w)} \quad (1)$$

$$= \frac{\text{Cov}(w, \alpha) + \beta \text{Cov}(w, w) + \text{Cov}(w, u)}{\text{Var}(w)} \quad (2)$$

$$= \frac{\text{Cov}(w, \alpha)}{\text{Var}(w)} + \beta \frac{\text{Cov}(w, w)}{\text{Var}(w)} + \frac{\text{Cov}(w, u)}{\text{Var}(w)} \quad (3)$$

$$= 0 + \beta \frac{\text{Var}(w)}{\text{Var}(w)} + \frac{\text{Cov}(w, u)}{\text{Var}(w)} \quad (4)$$

$$= \beta + \frac{\rho}{\text{Var}(w)}. \quad (5)$$

The term $\frac{\rho}{\text{Var}(w)}$ is not zero unless $\rho = 0$.

Instrumental Variables

Continue to assume that $\gamma = 0$ (so no x variable for ease of writing).

This is a simple derivation of the IV estimator:

$$y = \alpha + \beta w + \gamma x + u \quad (6)$$

$$\rightarrow \text{Cov}(z, y) = \text{Cov}(z, \alpha + \beta w + u) \quad (7)$$

$$= \text{Cov}(z, \alpha) + \beta \text{Cov}(z, w) + \text{Cov}(z, u) \quad (8)$$

$$= \beta \text{Cov}(z, w) + \text{Cov}(z, u) \quad (9)$$

$$\rightarrow \frac{\text{Cov}(z, y)}{\text{Cov}(z, w)} = \beta \frac{\text{Cov}(z, w)}{\text{Cov}(z, w)} + \frac{\text{Cov}(z, u)}{\text{Cov}(z, w)} \quad (10)$$

$$= \beta + \frac{\text{Cov}(z, u)}{\text{Cov}(z, w)} \quad (11)$$

$$= \beta \quad (12)$$

Under the assumptions we have introduced:

$$\beta_{OLS} = \frac{\text{Cov}(w, y)}{\text{Var}(w)} = \beta + \frac{\rho}{\text{Var}(w)} \quad (13)$$

$$\beta_{IV} = \frac{\text{Cov}(z, y)}{\text{Cov}(z, w)} = \beta \quad (14)$$

Now, how do we estimate $\frac{\text{Cov}(z, y)}{\text{Cov}(z, w)}$?

How to estimate this?

If one stares at the above long enough, then one realizes it is equal to the ratio of the first stage and the reduced form:

1. Regress w on z , get $\pi_{1,OLS} = \text{Cov}(z, w)/\text{Var}(z)$

2. Regress y on z , get $\theta_{1,OLS} = \text{Cov}(z, y)/\text{Var}(z)$
3. Divide (2) by (1):

$$\frac{\text{Cov}(z, y)/\text{Var}(z)}{\text{Cov}(z, w)/\text{Var}(z)} = \frac{\text{Cov}(z, y)}{\text{Cov}(z, w)} = \beta \quad (15)$$

While this has two parts, this is not technically Two Stage Least Squares.

Two Stage Least Squares

Two stage least squares is technically a different two step approach that uses the first stage and the structural equation:

1. Regress w on z , get $\hat{w} = \pi_0 + \pi_1 z$
2. Regress y on \hat{w} , get β_{2SLS}

What is β_{2SLS} :

$$\beta_{2SLS} = \frac{\text{Cov}(\hat{w}, y)}{\text{Var}(\hat{w})} \quad (16)$$

$$= \frac{\text{Cov}(\pi_0 + \pi_1 z, \alpha + \beta w + u)}{\text{Var}(\pi_0 + \pi_1 z)} \quad (17)$$

$$= \frac{\text{Cov}(\pi_1 z, \beta w + u)}{\text{Var}(\pi_1 z)} \quad (18)$$

$$= \frac{\text{Cov}(\pi_1 z, \beta w)}{\text{Var}(\pi_1 z)} + \frac{\text{Cov}(\pi_1 z, u)}{\text{Var}(\pi_1 z)} \quad (19)$$

$$= \frac{\pi_1 \beta \text{Cov}(z, w)}{\pi_1^2 \text{Var}(z)} + \frac{\pi_1 \text{Cov}(z, u)}{\pi_1^2 \text{Var}(z)} \quad (20)$$

$$= \frac{\beta \text{Cov}(z, w)}{\pi_1 \text{Var}(z)} \quad (21)$$

$$= \frac{\beta}{\pi_1} \pi_1 \quad (22)$$

$$= \beta \quad (23)$$

Both estimate the same underlying population parameter. 2SLS is used because it is a little easier to deal with additional variables and get standard errors (because it is still based on an underlying regression).

Ch4.4: Local Average Treatment Effects

Introduce Heterogeneity

Did you notice that all the equations above had fixed parameters? While I did not write it down, the above all implicitly assumed (based on how the equations were written) that there was no heterogeneity in the effects of the variables. However, using our own intuition and experience, we know that often people respond differently to the same stimulus or policy.

To be clear, heterogeneity in treatment effects just means (1) different people are affected by the policy different, (2) there is no single treatment effect, (3) there is a distribution of treatment effects. These three statements are essentially the same just phrased different.

Internal vs External Validity

Once we open up the doors to heterogeneity, a new question arises: when we estimate a treatment effect, for whom is the estimate relevant? One way this manifests is: can we use our estimate to predict outcomes for people who were not in the study? This is called internal versus external validity.

Internal validity concerns whether the estimates of the treatment effect really tell us about the studied population. External validity concerns whether the estimates of the treatment effect can tell us about others in similar situations but different context.

MHE uses the example of use the Vietnam War draft lottery to assess how military service affected labor earning after the war. Internal validity: do the estimates tell us how those drafted into the Vietnam war were affected? The authors argue 'yes' – the authors believe their instrumental variable strategy can tell us about the average treatment effect on the treated who were affected by the draft lottery. External validity: do the estimates tell us how those drafted into World War 2 were affected; do the estimates tell us about those who volunteered to fight in Vietnam? The authors say for the first question 'yes' but the second question 'no' – another study found similar labor market effects on WW2 drafted veterans, but the authors themselves found that voluntary enlisted veterans have very different outcomes compared to drafted.

Potential Outcomes for IV Model

Let Z be a binary instrument, D be an endogenous binary treatment variable, and Y be our outcome variable. Let $Y_i(D, Z)$ be a function that maps Z, D to Y for each individual.

Let $D_{z,i}$ indicate the status of the instrument, so we have $D_{1,i}$ when $Z_i = 1$ and $D_{0,i}$ when $Z_i = 0$. Note, $D_{z,i}$ can be 0 or 1 for either value of Z_i – this notation simply tells us what the treatment status is given the instrument's value. To see this, suppose we have the following linear model:

$$D_i = D_{0,i} + (D_{1,i} - D_{0,i})Z_i = (\pi_0 + \xi_i) + \pi_{1,i}Z_i, \quad (24)$$

where $\pi_0 = E[D_{0,i}]$.

Using the MHE example, for individual i , $D_{0,i}$ is the enlistment status if having a high draft lottery number (higher numbers mean less likely to be drafted) and $D_{1,i}$ is the enlistment status with a low lottery number.

Assumptions for Estimation

It turns out that we need stronger assumptions in this more heterogeneous world.

- Independence of Instrument on Potential Outcomes:

$$[\{Y_i(d, z); \forall d, z\}, D_{1,i}, D_{0,i}] \perp\!\!\!\perp Z_i \quad (25)$$

- Exclusion Restriction: $Y_i(d, 0) = Y_i(d, 1)$
- Monotonicity: $D_{1,i} \geq D_{0,i}$ or $D_{1,i} \leq D_{0,i}$

Independence is essentially the assumption of random assignment of the instrument. It solves issues like selection bias. The exclusion restriction is a more general version of (or at least a different way of writing) $\text{Cov}(z, u) = 0$ – it requires that the effect of the instrument is only through the endogenous variable of interest. Thus, if we know the treatment status, then the instrument does not give us any additional information about the potential outcomes of the outcome variable.

Monotonicity is a completely new assumption. It means that the effect of the instrument can only go one way. The MHE example is that having a low draft lottery number cannot make some people *less likely* to enlist in the military. This assumption is primarily to help us estimate things – that is, it is for convenience. We want to calculate something like a weighted average of treatment effects, but if we do not have monotonicity some of the ‘weights’ could be negative, which is bad.

Compliers, Always-Takers, Never-Takers, Defiers

This terminology is for the different groups of treatment status and instrument status:

- Compliers: $D_i = 1$ if $Z_i = 1$ and $D_i = 0$ if $Z_i = 0$
- Always Takers: $D_i = 1$ regardless of Z_i
- Never Takers: $D_i = 0$ regardless of Z_i
- Defiers: $D_i = 0$ if $Z_i = 1$ and $D_i = 1$ if $Z_i = 0$

Essentially, monotonicity says you cannot have defiers.

Potential Outcomes of the Outcome Variable

As always, we only ever observe one of the potential outcomes of individuals:

$$Y_i = Y_i(0, Z_i) + [Y_i(1, Z_i) - Y_i(0, Z_i)]D_i \quad (26)$$

$$= \alpha_0 + \rho_i D_i + n_i. \quad (27)$$

Local Average Treatment Effect (LATE)

This name should make you recall the Average Treatment Effect, except we are calling it ‘local.’ What does local mean? Essentially, it highlights that the LATE is going to be based on Compliers, so it is not true for all people (which would be ‘global’) but only for some (so ‘local’).

The LATE:

$$\rho_{LATE} = \frac{E[Y_i | Z_i == 1] - E[Y_i | Z_i == 0]}{E[D_i | Z_i == 1] - E[D_i | Z_i == 0]} = E[(Y_{1,i} - Y_{0,i}) | D_{1,i} > D_{0,i}] = E[\rho_i | \pi_{1,i} > 0]. \quad (28)$$

Note, $D_{1,i} > D_{0,i}$ means that we are only looking at Compliers.

Question 1: How do we know there are no Always Takers or Never Takers embedded in the estimate?

When is the LATE the ATT?

If there are no Always Takers or Never Takers, then the LATE is the ATT. Page 164 derives this. Loosely, the intuition for the result is that when there are only compliers, the instrument effectively becomes an indicator for treatment status.

Two Instruments

Suppose we have two instruments: z_1 and z_2 . Our first stage regression is: $w = \pi_1 z_1 + \pi_2 z_2 + e$. Our structural equation is the same, and as expected the reduced form has two IVs. I am going to combine the two derivations of the instrumental variable estimator. Note, the book does this using the LATE framework (page 174), but it is easier to write using constant effects.

Note on terminology: if the number of instruments is equal to the number of endogenous variables, then we say this is 'Just Identified' (or 'Exactly Identified'); whereas, if the number of instruments is greater than the number of endogenous variables, then it is 'Overidentified.' If we get to more advanced topics, then we can talk about this more.

The resulting 2SLS estimator is a weighted average of the single IV estimators:

$$\beta_{2SLS} = \frac{\text{Cov}(y, \hat{w})}{\text{Cov}(w, \hat{w})} \quad (29)$$

$$= \frac{\text{Cov}(y, \pi_1 z_1 + \pi_2 z_2)}{\text{Cov}(w, \pi_1 z_1 + \pi_2 z_2)} \quad (30)$$

$$= \frac{\text{Cov}(y, \pi_1 z_1) + \text{Cov}(y, \pi_2 z_2)}{\text{Cov}(w, \pi_1 z_1) + \text{Cov}(w, \pi_2 z_2)} \quad (31)$$

$$= \frac{\pi_1 \text{Cov}(y, z_1)}{\pi_1 \text{Cov}(w, z_1) + \pi_2 \text{Cov}(w, z_2)} + \frac{\pi_2 \text{Cov}(y, z_2)}{\pi_1 \text{Cov}(w, \pi_1 z_1) + \pi_2 \text{Cov}(w, \pi_2 z_2)} \quad (32)$$

$$= \frac{\pi_1 \text{Cov}(y, z_1)}{\pi_1 \text{Cov}(w, z_1) + \pi_2 \text{Cov}(w, z_2)} \frac{\text{Cov}(w, z_1)}{\text{Cov}(w, z_1)} + \frac{\pi_2 \text{Cov}(y, z_2)}{\pi_1 \text{Cov}(w, \pi_1 z_1) + \pi_2 \text{Cov}(w, \pi_2 z_2)} \frac{\text{Cov}(w, z_2)}{\text{Cov}(w, z_2)} \quad (33)$$

$$= \frac{\text{Cov}(y, z_1)}{\text{Cov}(w, z_1)} \frac{\pi_1 \text{Cov}(w, z_1)}{\pi_1 \text{Cov}(w, z_1) + \pi_2 \text{Cov}(w, z_2)} + \frac{\text{Cov}(y, z_2)}{\text{Cov}(w, z_2)} \frac{\pi_2 \text{Cov}(w, z_2)}{\pi_1 \text{Cov}(w, \pi_1 z_1) + \pi_2 \text{Cov}(w, \pi_2 z_2)} \quad (34)$$

$$= \beta_{IV, z_1} \cdot \psi_1 + \beta_{IV, z_2} \cdot \psi_2, \quad (35)$$

where $\psi_j = \frac{\pi_j \text{Cov}(w, z_j)}{\pi_1 \text{Cov}(w, z_1) + \pi_2 \text{Cov}(w, z_2)} \in (0, 1)$.

Thus, 2SLS with two instruments and one endogenous variable ('overidentified') is a weighted average of the two 'just identified' estimators, where the weights ψ are based on the relative covariances of the instrument and the endogenous variable.

Question 2: Instrument relevance was an assumption for IV model. How does this assumption interact with the weights?