Slides for the course

Statistics and econometrics

Part 13: An introduction to Instrumental Variables estimation

European University Institute

Andrea Ichino

September 7, 2019

Outline

Section 1

The traditional Interpretation of IV estimation

The setup

Let's continue to assume that

- $U_i(1) = U_i(0)$: no idiosyncratic gain from treatment;
- $\Delta = \mu(1) - \mu(0)$

so that the model in compact form is

$$Y_i = \mu(0) + \Delta D_i + U_i \tag{1}$$

$$D_i^* = \alpha + \beta Z_i + V_i \tag{2}$$

$$D_i = \left\{ \begin{array}{ll} 1 & \text{if } D_i^* \geq 0 \\ 0 & \text{if } D_i^* < 0 \end{array} \right\} \tag{3}$$

$$E\{U_i\} = E\{V_i\} = 0 \tag{4}$$

The IV estimator

If subjects are not randomly selected into treatment:

$$COV\{U,V\} = E(UV) \neq 0 \tag{5}$$

and OLS gives an inconsistent estimate of Δ.

$$\text{plim}\{\hat{\Delta}_{OLS}\} = \frac{COV\{Y,D\}}{V\{D\}} = \Delta + \frac{COV\{U,D\}}{V\{D\}} \neq \Delta \tag{6}$$

But under the assumptions

$$COV(Z,D) \neq 0 \tag{7}$$

$$COV(U,Z) = 0. \tag{8}$$

satisfied by our compact model, we have that:

$$\frac{COV\{Y,Z\}}{COV\{D,Z\}} = \Delta + \frac{COV\{U,Z\}}{COV\{D,Z\}} = \Delta = \text{plim}\{\hat{\Delta}_{IV}\} \tag{9}$$

Substituting the appropriate sample covariances on the LHS of 9 we get the well known IV estimator $\hat{\Delta}_{IV}$.

We will now explore a more recent and inspiring interpretation of this estimator.

Section 2

Instrumental variables as “quasi-experiments”

Setup and notation

Consider the following notation:

- N units denoted by i.
- They are exposed to two possible levels of treatment: $D_i = 0$ and $D_i = 1$.
- Y_i is a measure of the outcome.
- Z_i is a binary indicator that denotes the assignment to treatment.

Three crucial issues:

1. assignment to treatment may or may not be random;
2. assignment to treatment may or may not affect the outcome for given treatment status;
3. the correspondence between assignment and treatment may be imperfect.

Examples: Willis and Rosen (1979), Angrist (1990), Angrist and Krueger (1991), Card (1995), Ichino and Winter-Ebmer (2004).

Participation into treatment

Participation into treatment depends on the vector of assignments **Z**

$$D_i = D_i(\mathbf{Z}) \tag{10}$$

The outcome depends on the vector of assignments **Z** and treatments **D**:

$$Y_i = Y_i(\mathbf{Z}, \mathbf{D}) \tag{11}$$

Note that in this framework we can define three (main) causal effects:

- the effect of assignment Z_i on treatment D_i;
- the effect of assignment Z_i on outcome Y_i;
- the effect of treatment D_i on outcome Y_i.

The first two of these effects are called *intention-to-treat* effects.

The Angrist-Imbens-Rubin Causal model (see Angrist et. al. 1996) defines the minimum set of assumptions that ensures the identification of these effects for a relevant subgroup in the population.

Section 3

Assumptions of the Angrist-Imbens-Rubin causal model

Assumption 1: Stable Unit Treatment Value Assumption

Assumption

The potential outcomes and treatments of unit i are independent of the potential assignments, treatments and outcomes of unit $j \neq i$:

1. $D_i(\mathbf{Z}) = D_i(Z_i)$
2. $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(Z_i, D_i)$

Given this assumption we can write the *intention-to-treat* effects as:

Definition

The Causal Effect of Z on D for unit i is

$$D_i(1) - D_i(0)$$

Definition

The Causal Effect of Z on Y for unit i is

$$Y_i(1, D_i(1)) - Y_i(0, D_i(0))$$

Potential outcomes and SUTVA

Counterfactual reasoning requires to imagine that for each subject the sets of

- potential outcomes $[Y_i(0,0),\ Y_i(1,0),\ Y_i(0,1),\ Y_i(1,1)]$
- potential treatments $[D_i(0) = 0, D_i(0) = 1, D_i(1) = 0, D_i(1) = 1]$
- potential assignments $[Z_i = 0, Z_i = 1]$

exist, although only one item for each set is actually observed.

Implications of SUTVA for general equilibrium analysis and external validity.

If SUTVA holds, we can classify subjects according to the following useful typology.

A useful classification

Table: Classification of units according to assignment and treatment status

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

Examples: Willis and Rosen (1979), Angrist (1990), Angrist and Krueger (1991), Card (1995), Ichino and Winter-Ebmer (2004).

Assumption 2: Random Assignment (ignorability)

Assumption

All units have the same probability of assignment to treatment:

$$Pr\{Z_i = 1\} = Pr\{Z_j = 1\}$$

Given SUTVA and random assignment we can identify and estimate the two *intention to treat* causal effects:

$$E\{D_i \mid Z_i = 1\} - E\{D_i \mid Z_i = 0\} = \frac{COV\{D_i Z_i\}}{VAR\{Z_i\}} \tag{12}$$

$$E\{Y_i \mid Z_i = 1\} - E\{Y_i \mid Z_i = 0\} = \frac{COV\{Y_i Z_i\}}{VAR\{Z_i\}} \tag{13}$$

Note that the ratio between these effects is the IV estimand

$$\frac{COV\{Y, Z\}}{COV\{D, Z\}} \tag{14}$$

Is this the causal effect of D_i on Y_i?

Assumption 3: Non-zero average causal effect of Z on D

Assumption

The probability of treatment must be different in the two assignment groups:

$$Pr\{D_i(1) = 1\} \neq Pr\{D_i(0) = 1\}$$

or equivalently

$$E\{D_i(1) - D_i(0)\} \neq 0$$

This assumption requires that the assignment to treatment is correlated with the treatment indicator.

It is easy to test.

It is the equivalent of the“first stage" in the conventional IV approach.

Assumption 4: Exclusion Restrictions

Assumption

The assignment affects the outcome only through the treatment and we can write

$$Y_i(0, D_i) = Y_i(1, D_i) = Y_i(D_i).$$

It cannot be tested because it relates quantities that can never be observed jointly:

$$Y_i(0, D_i) = Y_i(1, D_i)$$

It says that given treatment, assignment does not affect the outcome. So we can define the causal effect of D_i on Y_i with the following simpler notation:

Definition

The Causal Effect of D on Y for unit i is

$$Y_i(1) - Y_i(0)$$

Are the first four assumptions sufficient for identification?

We can now establish the relationship *at the unit level* between the *intention to treat* effects of Z on D and Y and the causal effect of D on Y.

$$
\begin{aligned}
Y_i(1, D_i(1)) - Y_i(0, D_i(0)) & \\
&= Y_i(D_i(1)) - Y_i(D_i(0)) \\
&= [Y_i(1)D_i(1) + Y_i(0)(1 - D_i(1))] - \\
&\quad [Y_i(1)D_i(0) + Y_i(0)(1 - D_i(0))] \\
&= (D_i(1) - D_i(0))(Y_i(1) - Y_i(0)) \qquad (15)
\end{aligned}
$$

At the unit level the causal effect of Z on Y is equal to the product of the the causal effect of Z on D times the causal effect of D on Y.

Can we take the expectation of both sides of 15 and identify the average causal effect of D on Y:

$$E(Y_i(1) - Y_i(0))?$$

The answer is no ...

Because:

$$
\begin{aligned}
E & \quad \{Y_i(1, D_i(1)) - Y_i(0, D_i(0))\} \\
= & \quad E\{(D_i(1) - D_i(0))(Y_i(1) - Y_i(0))\} \\
= & \quad E\{Y_i(1) - Y_i(0) \mid D_i(1) - D_i(0) = 1\}Pr\{D_i(1) - D_i(0) = 1\} - \\
& \quad E\{Y_i(1) - Y_i(0) \mid D_i(1) - D_i(0) = -1\}Pr\{D_i(1) - D_i(0) = -1\}
\end{aligned}
\tag{16}
$$

Equation 16 shows that even with the four assumptions that were made so far we still have an identification problem.

What we observe (the left hand side), is equal to the weighted difference between the average effect for *compliers* and the average effect for *defiers*.

To solve this problem we need a further and last assumption.

Intuition for the final assumption

Table: Causal effect of Z on Y according to assignment and treatment status

		$Z_i = 0$	
		$D_i(0) = 0$	$D_i(0) = 1$
$Z_i = 1$	$D_i(1) = 0$	*Never-taker* $Y_i(1,0) - Y_i(0,0) = 0$	*Defier* $Y_i(1,0) - Y_i(0,1) = -(Y_i(1) - Y_i(0))$
	$D_i(1) = 1$	*Complier* $Y_i(1,1) - Y_i(0,0) = Y_i(1) - Y_i(0)$	*Always-taker* $Y_i(1,1) - Y_i(0,1) = 0$

Interpretation of the previous table

- Each cell contains the causal effect of Z on Y (the numerator of LATE).
- The SUTVA assumption allows us to write this causal effect for each unit independently of the others.
- The random assignment assumption allows us to identify the causal effect for each group.
- Exclusion restrictions ensure that the causal effect is zero for the *always-* and *never-takers*; it is non-zero only for *compliers* and *defiers* (via D).
- The assumption of strong monotonicity ensures that there are no *defiers* and that *compliers* exist.

All this ensures that the numerator of the LATE estimator is the average effect of Z on Y for the group of *compliers* (absent general equilibrium considerations).

Assumption 5: Monotonicity

Assumption

No one does the opposite of his/her assignment, no matter what the assignment is:

$$D_i(1) \geq D_i(0) \qquad \forall i \tag{17}$$

This assumption amounts to excluding the possibility of *defiers*.

The combination of Assumptions 3 and 5 is called *Strong Monotonicity*

$$D_i(1) \geq D_i(0) \qquad \forall i \;\; \text{with strong inequality for at least some } i \tag{18}$$

and ensures that:

- there is no defier and
- there exists at least one complier.

Since now *defiers* do not exist by assumption, we can use equation 16 to identify the average treatment effect for *compliers*.

Section 4

The Local Average Treatment Effect

The LATE

Equation 16 now is:

$$\begin{aligned} E & \{Y_i(1, D_i(1)) - Y_i(0, D_i(0))\} \\ = & \; E\{Y_i(1) - Y_i(0) \mid D_i(1) - D_i(0) = 1\} Pr\{D_i(1) - D_i(0) = 1\} \end{aligned} \tag{19}$$

Rearranging this equation, the Local Average Treatment Effect is defined as:

$$E\{Y_i(1) - Y_i(0) \mid D_i(1) - D_i(0) = 1\} = \frac{E\{Y_i(1, D_i(1)) - Y_i(0, D_i(0))\}}{Pr\{D_i(1) - D_i(0) = 1\}}$$

Definition

The Local Average Treatment Effect is the average effect of treatment for those who change treatment status because of a change of the instrument; i.e. the average effect of treatment for compliers.

Equivalent expressions for the LATE estimator

There are different ways to write the LATE

$$
\begin{aligned}
E\{Y_i(1) - Y_i(0) \quad &| \quad D_i(1) = 1, D_i(0) = 0\} \\
&= \frac{E\{Y_i \mid Z_i = 1\} - E\{Y_i \mid Z_i = 0\}}{Pr\{D_i(1) = 1\} - Pr\{D_i(0) = 1\}} \qquad (20) \\
&= \frac{E\{Y_i \mid Z_i = 1\} - E\{Y_i \mid Z_i = 0\}}{E\{D_i \mid Z_i = 1\} - E\{D_i \mid Z_i = 0\}} \qquad (21) \\
&= \frac{COV\{Y, Z\}}{COV\{D, Z\}} \qquad (22)
\end{aligned}
$$

- The IV estimand is the LATE.
- The LATE is the only treatment effect that can be estimated by IV, unless we are willing to make further assumptions.

Frequency of types in the population

Table: Frequency of types in the population

		$Z_i = 0$	
		$D_i(0) = 0$	$D_i(0) = 1$
$Z_i = 1$	$D_i(1) = 0$	*Never-taker* $Pr\{D_i(1) = 0, D_i(0) = 0\}$	*Defier* $Pr\{D_i(1) = 0, D_i(0) = 1\}$
	$D_i(1) = 1$	*Complier* $Pr\{D_i(1) = 1, D_i(0) = 0\}$	*Always-taker* $Pr\{D_i(1) = 1, D_i(0) = 1\}$

Interpretation of the previous table

In the previous table:

- The denominator of the Local Average Treatment Effect is the frequency of *compliers*.
- Note that the frequency of compliers is also the average causal effect of Z on D (see eq 21):

$$E\{D_i \mid Z_i = 1\} - E\{D_i \mid Z_i = 0\} =$$

$$Pr\{D_i = 1 \mid Z_i = 1\} - Pr\{D_i = 1 \mid Z_i = 0\}.$$

- Indeed the LATE-IV estimator is the ratio of the two average *intention-to-treat* effects: the effect of Z on Y divided by the effect of Z on D.

Comments on the LATE interpretation of IV

- The AIR approach clarifies the set of assumptions under which the IV estimand is an average causal effect, but shows that this is not the ATT.

- To identify the ATT the conventional approach implicitly assumes that the causal effect is the same for all treated independently of assignment.

- Translated in the AIR framework this conventional assumption is (see the debate Heckman-AIR in Angrist et al., 1996):

$$E\{Y_i(1) - Y_i(0) \mid Z_i, D_i(Z_i) = 1\} = E\{Y_i(1) - Y_i(0) \mid D_i(Z_i) = 1\} \quad (23)$$

$$\begin{aligned} E\{Y_i(1) - Y_i(0) \quad & \mid \quad D_i(1) = 1; D_i(0) = 1\} \quad (24) \\ & = \quad E\{Y_i(1) - Y_i(0) \mid D_i(1) = 1; D_i(0) = 0\} \end{aligned}$$

  i.e., the causal effect of D on Y must be the same for *compliers* and *always-taker*. Typically this assumption cannot be tested and is unlikely to hold in many applications.

Comments on the LATE interpretation of IV (cont.)

- The conventional approach hides also the assumption of strong monotonicity.

- The AIR approach concludes that the only causal effect that IV can identify with a minimum set of assumptions is the causal effect for *compliers*, i.e. the LATE: the effect of treatment for those who change treatment status because of a different assignment.

- Intuitively this makes sense because *compliers* are the only group on which the data can be informative :
  - *compliers* are the only group with units observed in both treatments (given that *defiers* have been ruled out).
  - *always takers* and *never-takers* are observed only in one treatment.
  - The LATE is analogous to a regression coefficient estimated in linear models with unit effects using panel data. The data can only be informative about the effect of regressors on units for whom the regressor changes over the period of observation.

Comments on the LATE interpretation of IV (cont.)

- The conventional approach to IV, however, argues that the LATE is a controversial parameter because it is defined for an unobservable sub-population and because it is instrument dependent. And therefore it is no longer clear which interesting policy question it can answer.
- Furthermore it is difficult to think about the LATE in a general equilibrium context
- Hence, the conventional approach concludes that it is preferable to make additional assumptions, in order to answer more interesting and well posed policy questions.
- Yet there are many relevant positive and normative questions for which the LATE seems to be an interesting parameter in addition to being the only one we can identify without making unlikely assumptions.