Instrumental Variables

Martin J. Conyon

October 2015

Lancaster University

Potential outcomes framework

- Let D_i is the treatment
  $D_i = 1$ if individual i has been exposed to treatment;
  $D_i = 0$ if individual i has not been exposed to treatment.

- $Y_i(D_i)$ is the outcome
  $Y_i(1)$ is the outcome in case of treatment;
  $Y_i(0)$ is the outcome in case of no treatment.

Potential outcomes framework

- If i participates [does not participate] in the program, $Y_i(1)$ [$Y_i(0)$] will be realized and $Y_i(0)$ [$Y_i(1)$] will ex post be a counterfactual outcome.
- Defining the realized outcome by Y_i, with Y the N-vector with i-th element equal to Y_i. The observed outcome for each unit can be written as:

$$Y_i(D_i) = D_i Y_i(1) + (1 - D_i) Y_i(0) \tag{1}$$

Causal effect

Definition 1. Causal effect.

For a individual i, the treatment D_i has a causal effect on the outcome Y_i if the event $D_i = 1$ instead of $D_i = 0$ implies that $Y_i = Y_i(1)$ instead of $Y_i = Y_i(0)$. In this case the causal effect of D_i on Y_i is

$$\Delta_i = Y_i(1) - Y_i(0) \tag{2}$$

The identification and the measurement of this effect is logically impossible.

Fundamental Problem of Causal Inference

Proposition 1. Fundamental Problem of Causal Inference.

It is impossible to observe for the same individual i the values $D_i = 1$ *and* $D_i = 0$ *as well as the values* $Y_i(1)$ *and* $Y_i(0)$ *and, therefore, it is impossible to observe the effect of D on Y for individual i*

Another way to express this problem is to say that we cannot infer the effect of treatment because we do not have the *counterfactual* evidence i.e. what would have happened in the absence of treatment.

Average Treatment Effects

The most popular treatment effect is the Average Treatment Effect (ATE), the population expectation of the unit-level causal effect $Y_i(1) - Y_i(0)$.

Suppose you pick a person at random in the population and you expose him/her to treatment. The expected effect on the outcome for this person is given by:

$$\tau_{ATE} = E[\Delta_i] = E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)] \tag{3}$$

We cannot observe the outcome in both counterfactual situations for all the individuals and therefore we cannot compute the expectations on the right-hand side.

Average Treatment Effects

The Average Treatment effect on the Treated (ATT) is the average over the subpopulation of those who are actually treated.

It is the difference between the average outcome in case of treatment (which we observe) minus the average outcome in the counterfactual situation of no-treatment (which we do not observe).

$$\begin{aligned}\tau_{ATT} =& E[\Delta_i|D_i = 1] = E[Y_i(1) - Y_i(0)|D_i = 1]\\ =& E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1].\end{aligned} \tag{4}$$

Problem: both these average treatment effects cannot be easily identified and estimated with observational data.

Difference in means estimator

Average treatment effect can be decomposed into ATT plus a sample selection bias

$$\tau_{ATE} = E[\Delta_i] = E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$$

$$\begin{aligned} E[Y_i(1)] - E[Y_i(0)] &= E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0] \\ &\Rightarrow E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] \\ &\quad + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0] \end{aligned}$$

$$\tau_{ATE} = \tau_{ATT} + E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0). \tag{5}$$

where $\tau_{ATT} = E[\Delta_i|D_i = 1]$ and a sample selection bias is equal to $E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0)$

Difference in means estimator

So,

$$\tau_{ATE} = \tau_{ATT} + E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0).$$

- The difference between the left hand side (which we can estimate) and τ_{ATT}, is the *sample selection bias*
- The problem is that the outcome of the treated and the outcome of the control subjects are not identical in the no-treatment situation.
- The goal of most empirical economic research is to overcome selection bias, and to say something about the causal effect of a variable like D_i.

Random Assignment

Random assignment of D_i solves the selection problem because it makes D_i independent of potential outcomes:

$$Y(1), Y(0) \perp D.$$

Randomization solves the Fundamental Problem of Causal Inference because it allows to use the control units (C) as an image of what would happen to the treated units (T) in the counterfactual situation of no treatment, and vice-versa.

Thus randomized treatment guarantees that the difference-in-means estimator from basic statistics is unbiased, consistent and asymptotically normal.

Causality in a Regression Framework

Specify the following potential outcomes:

$$Y_i(1) = \mu(1) + U_i(1)$$
$$Y_i(0) = \mu(0) + U_i(0)$$

where $E[U_i(1)] = E[U_i(0)] = 0$. Recall from earlier:

$$Y_i(D_i) = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

and on substitution:

$$Y_i = \mu(0) + (\mu(1) - \mu(0)) D_i + U_i(0) + ((U_i(1) - U_i(0)) D_i$$

Causality in a Regression Framework

The causal effect of treatment for an individual is

$$\begin{aligned}\Delta_i =& Y_i(1) - Y_i(0) = \rho_i \\ =&[\mu(1) - \mu(0] + [U_i(1) - U_i(0)] \\ =&E[\Delta_i] + [U_i(1) - U_i(0)].\end{aligned}$$

where:

- $E[\Delta_i] = \mu(1) - \mu(0)$
  is the common gain from treatment equal for every individual i.
- $[U_i(1) - U_i(0)]$
  is the idiosyncratic gain from treatment that differs for each individual i and that may or may not be observed by the individual.

Causality in a Regression Framework

The effect of treatment on a random individual (ATE)

$$
\begin{aligned}
E[\Delta_i] =& E[Y_i(1) - Y_i(0)] \\
=& E[Y_i(1)] - E[Y_i(0)] \\
=& \mu(1) - \mu(0) = \tau_{ATE}
\end{aligned}
$$

The effect of treatment on the treated (ATT)

$$
\begin{aligned}
E[\Delta_i | D_i = 1] =& E[Y_i(1) - Y_i(0) | D_i = 1] \\
=& E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 1] \\
=& \mu(1) - \mu(0) + E[U_i(1) - U_i(0) | D_i = 1] \\
=& \tau_{ATE} + E[U_i(1) - U_i(0) | D_i = 1] = \tau_{ATT}.
\end{aligned}
$$

Causality in a Regression Framework

The two effects differ because of the idiosyncratic gain for the treated

$$E[U_i(1) - U_i(0)|D_i = 1].$$

This is the average gain that those who are treated obtain on top of the average gain for a random person in the population.

Regression with Random Coefficients

Let D_i indicate treatment so the outcome can be written as:

$$
\begin{aligned}
Y_i =& D_i Y_i(1) + (1 - D_i) Y_i(0) \\
=& \mu(0) + [\mu(1) - \mu(0) + U_i(1) - U_i(0)] D_i + U_i(0) \\
=& \mu(0) + \Delta_i D_i + U_i(0) \\
=& \mu(0) + \rho_i D_i + U_i(0).
\end{aligned}
$$

where $D_i = 1$ in case of treatment and $D_i = 0$ otherwise. This is a linear regression with a random coefficient, ρ_i on the RHS variable D_i. This is the case of heterogenous treatment effects.

Regression with Constant Coefficients

Suppose now that the treatment effect is the same for everyone, this implies equal idiosyncratic gains for each individuals:

$$
\begin{aligned}
[U_i(1) = & U_i(0)] \\
Y_i(1) - Y_i(0) = & \mu(1) - \mu(0) = \rho
\end{aligned}
$$

then the equation reduces to

$$
\begin{aligned}
Y_i = & D_i Y_i(1) + (1 - D_i) Y_i(0) \\
= & \mu(0) + [\mu(1) - \mu(0) + U_i(1) - U_i(0)] D_i + U_i(0) \\
= & \mu(0) + \rho D_i + U_i(0).
\end{aligned}
$$

Regression with Constant Coefficients

From before: $Y_i = \mu(0) + \rho D_i + U_i(0)$

and so the treatment effects are equal:

$$\begin{aligned} E[\Delta_i] =& E[Y_i(1) - Y_i(0)] = \tau_{ATE} = \rho \\ E[\Delta_i | D_i = 1] =& E[Y_i(1) - Y_i(0) | D_i = 1] \\ =& \mu(1) - \mu(0) + E[U_i(1) - U_i(0) | D_i = 1] \\ =& \tau_{ATE} = \tau_{ATT} = \rho. \end{aligned}$$

Conditional Independence Assumption

When D_i and $(Y_i(1), Y_i(0))$ are allowed to be correlated, we need an assumption in order to identify the treatment effects.

The Conditional Independence Assumption (CIA) allows causal interpretation of regression estimates. It asserts that conditional on observed covariates, X_i there are no unobserved factors that are associated both with the assignment, D_i and with the potential outcomes, Y_i.

Conditional Independence Assumption

Conditional independence:

Assumption CIA

$Y_i(1), Y_i(0) \perp D_i | X_i$.

Assume, then, ignorability in conditional mean independence sense:

Assumption CIA

- $E[Y_i(0)|X_i, D_i] = E[Y_i(0)|X_i]$
- $E[Y_i(1)|X_i, D_i] = E[Y_i(1)|X_i]$

Conditional Independence Assumption

Implications: Given the CIA, conditional-on-X_i comparison of averages by treatment status (difference-in-means estimator) has a causal interpretation

$$E(Y_i(1)|X_i, D_i = 1) - E(Y_i(0)|X_i, D_i = 0) = E[Y_i(1) - Y_i(0)|X_i]$$

This kind of assumption is used routinely in multiple regression analysis.

Conditional Independence Assumption

Let's consider model:

$$Y_i(1) = \mu(1) + U_i(1)$$
$$Y_i(0) = \mu(0) + U_i(0)$$

under CIA Assumption 2 and assume in addition

$$E[U_i(1)|X_i] = E[U_i(0)|X_i]$$

then the treatment effect, ρ, is constant, so that, for each random draw i,

$$\rho = Y_i(1) - Y_i(0) = \tau_{ATE} = \tau_{ATT}$$

Conditional Independence Assumption

Further, assume that

$$Y_i(0) = \alpha + \beta X_i + \nu_i$$

where X_i are **observable** covariates, uncorrelated with ν_i

$$E[\nu_i | X_i] = E[\nu_i] = 0.$$

Then, with the observed outcome defined as in the potential outcomes model we can write

$$Y_i = \alpha + \rho D_i + \beta X_i + \epsilon_i$$

where

$$\epsilon_i = \nu_i + [U_i(1) - U_i(0)] D_i$$

and

$$E[\epsilon_i | X_i] = 0$$

In this case unconfoundedness is equivalent to independence of ϵ_i