Michele Pellizzari
(University of Geneva)

Local Average Treatment Effect (LATE)

[This section is based on section 1 and 2 of the following famous paper: Imbens W.I. and J.D. Angrist "Identification and Estimation of Local Average Treatment Effects." *Econometrica, vol.62(2), 467-475.]*

Recent development in the identification and estimation of IV models has radically changed our understanding of this approach. In particular, in this section we will show that only under very specific assumptions IV and OLS estimate the same parameter and, even worse, that the parameter that is identified by IV varies depending on the instrument that is chosen.

Before we enter the discussion of the Local Average Treatment Effect (LATE) and its implications for instrumental variable, it is useful to introduce the notion of causality and how it relates to the simple OLS regression framework. To make this simple, consider a situation in which you are interested in the causal effect of a simple dummy variable x_i on some outcome y_i. A useful working example that we will use throughout this section is about the returns to education with y_i being earnings and x_i an indicator for college graduates. Following the impact evaluation literature define the following two potential outcomes:

- $y_i(0)$ = outcome of i conditional on $x_i = 0$;
- $y_i(1)$ = outcome of i conditional on $x_i = 1$.

The potential outcomes $y_i(0)$ or $y_i(1)$ allows defining formally what is the causal effect of x on y:

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

namely, the difference between the outcome that i would experience with $x_i = 1$ and the outcome that the same subject i would experience with $x_i = 0$. In terms of our working example, the effect of college on earnings for individual i is the difference between the earnings that i would make with a college degree and the earning that she would make without a college degree. Importantly the treatment effect Δ_i has a subscript i indicating that it can be heterogeneous across subjects: for some individuals the returns to college are high, for others they are low.

For each i either $y_i(0)$ or $y_i(1)$ is observed depending on whether x_i is equal to either 0 or 1 but for none of the is both potential outcomes are observed. Hence, the individual causal effect of Δ_i cannot be computed, for none of the observations. Nevertheless, a simple regression of the observable y_i on the observable x_i may identify the *average treatment effect (ATE)*, $E(\Delta_i) = E(y_i(1) - y_i(0))$. To see this, define the observable y_i as

$$y_i = x_i y_i(1) + (1 - x_i) y_i(0) \tag{2}$$

and rewrite each potential outcome as the sum of its mean and a random term:

$$y_i(0) = \mu(0) + \epsilon_i(0) \tag{3}$$
$$y_i(1) = \mu(1) + \epsilon_i(1) \tag{4}$$

where $\mu(0) = E(y_i(0))$ and $\mu(1) = E(y_i(1))$. By definition of the mean, $E(\epsilon_i(0)) = E(\epsilon_i(1)) = 0$. Under these assumptions, the ATE can be rewritten as $E(\Delta_i) = E(y_i(1) - y_i(0)) = E(y_i(1)) - E(y_i(0)) = \mu(1) - \mu(0)$.

Now replace equations 3 and 4 into equation 2 to derive the following linear model:

$$y_i = \mu(0) + [\mu(1) - \mu(0)]x_i + v_i \tag{5}$$

where $v_i = \epsilon_i(0) + [\epsilon_i(1) - \epsilon_i(0)]x_i$. Under exogeneity of x_i in equation 5, the OLS estimator of the slope parameter identifies the average treatment effect.[1]

Now suppose that, for some reason, you are concerned about the exogeneity of x_i in equation 5 and you find a valid instrument z_i that can be used to produce the IV estimator of the ATE. For simplicity, assume that the instrument is a dummy indicator which can only take values 0 or 1. Continuing with the returns to education example, imagine that z_i is the famous college proximity indicator used by David Card.[2] Hence, $z_i = 1$ if individual i lives close to a college and $z_i = 0$ if individual i lives far from college. The idea underlying the instrument is that the cost of attending college increases

[1]Exogeneity is guaranteed by the *mean independence assumption (MI)*: $E(y_i(k)|x_i = 1) = E(y_i(k)|x_i = 0)$ for $k = \{0, 1\}$. If MI only holds for $k = 0$ the linear regression of y_i on x_i identifies the *average treatment effect on the treated (ATT)* $E(y_i(1) - y_i(0)|x_i = 1)$.

[2]David Card. 1993. Using Geographic Variation in College Proximity to Estimate the Return to Schooling. NBER Working Paper No. 4483.

with geographical distance so that those who live far are less likely to attend. Exogeneity requires that living close to a college has no direct effect on earnings other than through college attendance.

We know that the IV estimator of equation 5 where we use z_i as an instrument converges in probability to the ratio of the covariance between z and y over the covariance between z and x:

$$\widehat{ATE}_{IV} = [\widehat{\mu(1) - \mu(0)}]_{IV} \xrightarrow{p} \frac{Cov(z_i, y_i)}{Cov(z_i, x_i)} \tag{6}$$

Now we need to do a bit of algebra to rewrite equation 6 in a more interesting format. Let us start from the numerator:

$$\begin{aligned} Cov(z_i, y_i) &= E(z_i y_i) - E(z_i)E(y_i) = \\ &= Pr(z_i = 1)E(y_i|z_i = 1) - Pr(z_i = 1)E(y_i) = \\ &= Pr(z_i = 1)E(y_i|z_i = 1) - \\ &\quad -Pr(z_i = 1)[Pr(z_i = 1)E(y_i|z_i = 1) + Pr(z_i = 0)E(y_i|z_i = 0)] = \\ &= Pr(z_i = 1)\{E(y_i|z_i = 1)[1 - Pr(z_i = 1)] - Pr(z_i = 0)E(y_i|z_i = 0)\} = \\ &= Pr(z_i = 1)Pr(z_i = 0)[E(y_i|z_i = 1) - E(y_i|z_i = 0)] \end{aligned} \tag{7}$$

Now, the exact same calculations on the denominator of equation 6 yield:

$$Cov(z_i, x_i) = Pr(z_i = 1)Pr(z_i = 0)[E(x_i|z_i = 1) - E(x_i|z_i = 0)] \tag{8}$$

Finally, replace equations 7 and 8 into 6:

$$[\widehat{\mu(1) - \mu(0)}]_{IV} \xrightarrow{p} \frac{E(y_i|z_i = 1) - E(y_i|z_i = 0)}{E(x_i|z_i = 1) - E(x_i|z_i = 0)} \tag{9}$$

Now, we need to define another set of potential variables similar to $y_i(0)$ and $y_i(1)$ but in this case it is not the potential y but rather the potential x:

- $x_i(0)$ = value of x_i conditional on $z_i = 0$;
- $x_i(1)$ = value of x_i conditional on $z_i = 1$.

In words, $x_i(0)$ is the value that x_i takes when the instrument z_i equals zero and, similarly, $x_i(1)$ is the value that x_i takes when the instrument z_i equals 1. The returns to schooling example may help clarify the meaning of these potential treatment indicators: $x_i(0)$ is equal to zero if the behaviour of individual i is such that she does not go to college (i.e. $x_i = 0$) if she lives far from one (i.e. when $x_i = 0$). Likewise for $x_i(1)$. The potential treatment indicators allow defining the following four groups of individuals:

1. the *compliers* are those is such that $x_i(0) = 0$ and $x_i(1) = 1$. They are individuals who go to college if they live close to one and do not go to college if they live far from one;

2. the *defiers* are those is such that $x_i(0) = 1$ and $x_i(1) = 0$. They are individuals who go to college if they live far from one and do not go to college if they live close to one;

3. the *always-takers* are those is such that $x_i(0) = 1$ and $x_i(1) = 1$. They are individuals who go to college regardless of whether they live close to one or not;

4. the *never-takers* are those is such that $x_i(0) = 0$ and $x_i(1) = 0$. They are individuals who don't go to college regardless of whether they live close to one or not.

Now we can define exogeneity of the instrument in a more subtle way as:

$$\{y_i(0), y_i(1), x_i(k)\} \perp z_i \quad \forall k \in \Gamma_z \tag{10}$$

where Γ_z is the support of z_i, which, in our specific example is simply $\{0, 1\}$. In order to understand this special version of exogeneity of the instrument, consider first the orthogonality of z_i and $x_i(k)$. What it implies is that the distribution of potential college goers and potential non-college goers is orthogonal to whether individuals live close to college or not. In other words, for exogeneity to hold it should not be the case that all those who are more likely to go to college live close to one. If it happens, for example, that all or most always-takers live close to college and all or most never-takers live far from a college, then exogeneity of the instrument would fail. Obviously, random assignment of z_i, i.e. randomly assigning individuals to residential locations, guarantees orthogonality with $x_i(k)$. However, random assignment of z_i does not necessarily imply orthogonality with the potential outcomes. In order for the instrument to be orthogonal to outcomes it must be that the returns to college are not particularly high (or particularly low) in areas close to colleges. In this case, it is the random assignment of colleges to locations that would guarantee orthogonality of $\{y_i(0), y_i(1)\}$ and z_i.

We now use the definition of exogeneity in 10 to further develop the numerator of equation 9. First of all replace y_i with equation 2 of both expected values to obtain the following:

$$\begin{aligned}
& E\left[x_i(1)y_i(1) + (1 - x_i(1))\, y_i(0) | z_i = 1\right] - E\left[x_i(0)y_i(1) + (1 - x_i(0))\, y_i(0) | z_i = 0\right] \\
= \; & E\left[x_i(1)y_i(1) + (1 - x_i(1))\, y_i(0) - x_i(0)y_i(1) + (1 - x_i(0))\, y_i(0)\right] \\
= \; & E\left[(x_i(1) - x_i(0))\,(y_i(1) - y_i(0))\right] = \\
= \; & Pr\left(x_i(1) - x_i(0) = 1\right) \cdot E\left[y_i(1) - y_i(0) | x_i(1) - x_i(0) = 1\right] - \\
& -Pr\left(x_i(1) - x_i(0) = -1\right) \cdot E\left[y_i(1) - y_i(0) | x_i(1) - x_i(0) = -1\right] \qquad (11)
\end{aligned}$$

where the first equality comes from direct application of the exogeneity assumption in 10.

Equation 11 is the key to the LATE interpretation of the instrumental variable estimator. It show two crucial points:

1. *never-takers* and *always-takers* do not contribute to the identification of the treatment effect. This implies that in general the IV estimator does not identify the same parameter as OLS, namely the average treatment effect on the entire population of interest, but rather the effect of the treatment over the unknown sub-population of individuals whose endogenous x is responsive to variation in the instrument. In the returns to education example, those who would go to college regardless of where they lived and those who would never go to college regardless of where they lived do not contribute to identification. Those observations are not used in equation 11;

2. the presence of *defiers* can severely bias the estimates. To see this, imagine that the treatment effect is positive and homogeneous, i.e. $\Delta_i = y_i(1) - y_i(0) = \Delta > 0$ for all i, and that there is an equal number of compliers and defiers, i.e. $Pr\left(x_i(1) - x_i(0) = 1\right) = Pr\left(x_i(1) - x_i(0) = -1\right)$, then the IV estimator would be zero, despite the positive average treatment effect (over any sub-population, given homogeneity).

There is no real solution to the first problem other than assuming it away by assuming homogeneous treatment. So the question is rather whether the local effect identified by a specific instrument is interesting or not and this is a question that cannot have a general answer and needs to be discussed on a case by case basis.

The second problem can be addressed by appropriately choosing the instrument, namely by choosing an instrument that does not allow for defiers or, more generally, a *monotonic* instrument. An instrument is monotonic when it influences the endogenous variable in the same direction for all units:

- *(Monotonicity)* For all $k, h \in \Gamma_z$, either $D_i(h) \geq D_i(k)$ for all i or $D_i(h) \leq D_i(k)$ for all i.

In other words monotonicity guarantees that $Pr\left(x_i(1) - x_i(0) = -1\right) = 0$ (or that $Pr\left(x_i(1) - x_i(0) = 1\right) = 0$, depending on how one defines the defiers) and it is an additional important property that a good IV should satisfy. As for exogeneity, it is a property that cannot be tested and one can only argue in favour of it.

To translate all this into the returns to college example, monotonicity implies that nobody is more likely to go to college when she lives far from one than when she lives close to one, which seems pretty reasonable in this context. Despite monotonicity of the instrument appears reasonable, there are likely to be many always-takers and never-takers. Presumably the always-takers are the best students who presumably have the highest treatment effects and who would presumably go to college regardless of where they lived. Likewise, the never-takers are presumably those students whose returns to college are particularly low and who would not go to college regardless of where they lived. Hence, the returns to college estimated via the college proximity instrument identify only the average effect over the sub-population of marginal college goers, namely those who would go to college if they lived closed to one and would not go if they lived far from one and vice-versa (the defiers). And such an effect is local in the sense that it is the average taken over a local sub-population of the original population of interest.

Equation 11 and the discussion above should also have made it clear that the specific local effect identified by an IV estimator depend on the instrument that is chosen. In other words, if we happen to have two instruments for the same endogenous variable, we would generally estimate different local effects when using one or the other. Every instrument identifies a different local effect.

This result has important implications for some of the most classical tests that are still taught in standard econometrics. In the light of LATE, the usual overidentification test (Sargan test) makes no sense. The fact that two IV estimators produced using different instruments differ from one another does not imply anything in terms of the exogeneity of the instruments. They are simply identifying different local effects and, unless we are willing to assume that the two instruments define the same sub-populations of compliers, defiers and never- or always-takers, there is no reason to expect them to produce the same IV estimators.

Likewise for the popular Hausmann test, which is supposed to test the exogeneity of x in linear models such as equation 5 by comparing the OLS and the IV estimator, makes little sense. OLS generally identifies a global effect, i.e. the average treatment over the entire population, whereas IV identifies a local effect and, unless one can provide good reasons to argue that the instrument is such that there are no defiers and no always- and never-takers, there is no particular reason to expect that the OLS and IV estimators would converge in probability to the same parameter. In other words, the difference between the OLS and the IV estimators does not provide any evidence to test the exogeneity of the explanatory variables of the model.

To conclude this section, it is worth mentioning that here we have derived all results under a number of simplifying assumptions but all of them are quite general and can be extended to multivariate models with continuous explanatory variables and instruments.