E797B - Problem Set 3

Jesús Lara

October 2020

Instrumental Variables Estimation

In this exercise we simulate two Data Generating Processes (DGPs) characterized by endogeneity and weak instruments in order to explore the performance of different estimation and hypothesis testing methods regarding bias and type one error. The DGPs that we simulate are:

$$y_i = \beta x_i + \eta_i$$
$$x_i = \sum_{j=1}^{Q} \pi_j z_{ij} + \xi_i$$

With $\beta = 1$, $\pi = 0.1$, $\pi_j = 0 \ \forall j > 1$, $\begin{pmatrix} \eta_i \\ \xi_i \end{pmatrix} |Z \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \right)$ and $z_j \sim N(0,1) \ \forall j$

Table 1: Different estimations of the effect of x (endogenous) on y

	$Q = 1$		$Q = 10$		$Q = 20$	
	Bias	Type 1 Error (%)	Bias	Type 1 Error (%)	Bias	Type 1 Error (%)
OLS	0.792	100.000	0.792	100.000	0.793	100.000
2SLS	-0.012	6.600	0.390	59.300	0.526	93.300
CLR	-0.012	0.000	0.390	4.300	0.526	4.500
LIML	-0.012	6.600	-0.006	6.000	0.007	8.400
Lasso[a]	0.089	9.635	0.192	20.216	0.205	21.455

Note. In this table we report the median bias $\hat{\beta} - 1$ and the percentage times that the correct null hypothesis that $\beta_1 = 1$ was rejected at the 5% significance level
a. Lasso regression did not pick any instrument 31.50 , 62.90 and 72.50 % of the times for $Q = 1, 10$ and 20 respectively. The reported outcomes for Lasso correspond to the cases when it selected at least one instrument

We simulate 1000 random samples of these DGDPs and in each one we estimate β and test the correct null hypothesis that $\beta = 1$ with 5 different methods: OLS, 2SLS, CLR, LIML, and Lasso, considering 1, 10 and 20 possible instruments (Q). In Table 1 we report the median bias and the percentage times that each estimation method incorrectly rejected H_0 (type one error, called reject rate herein after).

According to the specifications above, x is endogenous because it is correlated with the error term η. We know that because:

$$cov(x_i, \eta_i) = cov(0.1z_1 + \xi_i, \eta_i) = cov(\xi_i, \eta_i) = 0.8 \neq 0$$

Thus, the OLS estimator can be shown to be $\beta_{OLS} = \beta + \frac{cov(x,\eta)}{var(x)} = \beta + bias$.

Given that $var(x) = var(0.1z_1 + \xi_i) = (0.1)^2 var(z_1) + var(\xi) = 1.01$, we get that $\beta_{OLS} \approx 1 + 0.7921 \approx 1.7921$.

In contrast, β can be obtained by using the instrument z_1, which is relevant ($x_i = 0.1z_{i1} + \xi_i$) and excludable ($cov(z_{i1}, \eta_i)$=0). In particular, when we only use the relevant instrument z_1, $\beta_{IV} = \frac{cov(y_i, z_{i1})/var(x_i)}{cov(x_i, z_{i1})/var(x_i)}$ which can be estimated by 2-Stage Least Squares. The bias of the $2SLS$ estimator is given by

$$E[\hat{\beta}_{IV} - \beta] \approx \frac{\sigma_{\eta\xi}}{\sigma_{xi}^2} \frac{1}{F(Q) + 1}$$

F is the statistic associated to the 1st stage test that all instruments are weak (not jointly significant). F is inversely proportional to the number of instruments ($\frac{\Delta F}{\Delta Q} < 0$). Hence, many weak instruments increase the bias, while few strong instruments reduce it. We confirm this by comparing 2SLS with OLS. OLS has a constant median bias equal to 0.792 (which means that $\hat{\beta} \approx 1.792$) since the inclusion of instruments is irrelevant for this estimation. We get exactly the results that we expected from our theoretical analysis. Thus, OLS unambiguously rejects the correct null hypothesis that $\beta = 1$.

2SLS, in contrast, gives a very low median bias and type-1 error reject rate when including only the relevant instrument z_1, just as we would expect given the intuition provided above. However, we observe that as we incorporate irrelevant instruments, both the median bias and the reject rate increase, just as expected. When we include 20 instruments, one relevant and 19 irrelevant, the bias is 0.526. This means that, as we include irrelevant instruments, the IV estimator approaches to OLS as bias increases.

In the case of 2SLS, the reliability of the hypothesis test that $\beta = 1$ is affected by the fact that it uses the estimates $\hat{\pi}$, which given the construction of x, are not statistically significant for $j > 1$. The Conditional Likelihood Ratio (CLR) test corrects for the problems associated to a weak 1st stage. We can observe that the reject rate is much lower than with a simple $H_0 : \beta = 1$ test. The reject rate is never above 5%, even when we include 19 irrelevant instruments, which illustrates that it is a better test than a simple t-test.

However, the method that performs better is Limited-Maximum Likelihood information (LIML), which basically estimates β without bias and with a slightly higher reject rate than CLR. According to our exercise, LIML is the method that performs better when there are weak instruments. Hence, when we are doing real research, comparing the LIML with the 2SLS estimates is a good strategy to know the reliability of our instruments

Finally, we estimate β with Lasso. This method selects some variables from the pool of instruments to perform 2SLS. The bias and reject rate barely increase when we move from 9 to 19 irrelevant instruments. Although it does better than 2SLS when we include irrelevant instruments, LIML still is better regarding bias and reject rate. Also, I find it interesting that Lasso regression did not pick any instrument 31.50 , 62.90 and 72.50 % of the times for $Q = 1, 10$ and 20 respectively, despite z_1 being a relevant instrument. This illustrates the point that was made in the lecture that the application of Lasso to the selection of instruments is still in a very preliminary stage.