

#### **Topic 4 Simulation: Weak instruments and finite samples. How bad is the issue?**

In the main simulation notes this week we looked at issues with weak instruments and we showed the typically weak instruments do not cause an issue, unless they are correlated in some way with the error term. I recently read a very interesting piece on how in small samples (what econometricians often call finite samples) weak instruments do not perform well. I included this note at the end of the other notes but I never actually show the results. In these notes today I aim to use a Monte Carlo approach to show how badly IV can perform if there are weak instruments and the sample size is small. *Please do note, the coding in this piece is rather advanced and I do not expect you to understand it so please do not worry about this. I will explain the general idea of what is going on and then talk you through the results so you can see the key findings.*

The plan today is to run a few thousand regressions! We will use what is known as a Monte Carlo approach (named after the casinos where this topic was first investigated) where we run the same approach a large number of times on simulated data to get a real check of if our result is the norm, or was it a function of the random variables we defined. Due to following this approach I do not “set the seed” as in previous weeks as we want each of the 10,000 samples we use to have slight differences.

The first block of code you will see creates what is known as a programme for STATA, it is basically a list of instructions for STATA to follow if we type a certain keyword, exactly the same as how there is a programme for the command “regress”. Within this programme I first create a dataset with a few variables, an independent (x) variable, an error term (u), a variable that influences both x and y but is not included in the regression (w) to create the endogenous effect we want to model, the dependent variable (y) and finally our instrument (z) that is not related to y but is related to x. This should look familiar from the other set of notes. We then run both an OLS regression and an IV regression and store both the x coefficients and the standard errors. We can then simulate this command 10,000 times and view what happened to our coefficients and our standard errors.

To start with we will look at what happens in small samples with a strong instrument. Thus our first programme is called “strongreg”. Here our x variable has a correlation of over 0.8 with z, and thus this is a strong instrument. In the second section of the work I then reduce this relationship between the two is much small, a tenth of the size in fact with z and x only sharing a 0.08 correlation. This is the programme called “weakreg”.

#### **Finite Samples and a strong instrument**

```
cap program drop strongreg
program strongreg, rclass
clear
set obs `1'
gen z = rnormal()
gen w = rnormal()
gen x = z*2 + rnormal() + w
gen u = rnormal()
gen y = x + w + u*5
reg y x
return scalar reg_x = _b[x]
return scalar reg_se_x = _se[x]
ivreg y (x=z)
return scalar ivreg_x = _b[x]
return scalar iv_se_x = _se[x]
```

end

```
simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///
ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x) ///
,rep(10000): strongreg 100
sum
```

What do we see below? We see the average of the x coefficient (reg\_x) for a regular OLS estimator, the average of the standard error for the OLS estimate of x (reg\_se\_x), the average of the x coefficient on the IV regression (ivreg\_x) and the average of the standard error for the IV estimate of x (iv\_se\_x), when run on 10,000 different samples of 100 observations. So this is a very small sample, with a strong instrument. What do we see? The average estimate of the OLS regression is 1.17, not very close to the value of 1 that the true value of x is, but the average standard error is not too large (0.21) and thus we would still conclude that x is significantly different to zero. The IV results are however much better! The average estimate of x is 1, it's true value and the standard errors are not much larger than those for the OLS. We would still be able to say that x is significantly different to zero. So a small sample size such as 100 does not matter for IV regressions if there is a strong instrument, IV will still be better than the OLS model that we know is biased.

**Figure 1: Monte Carlo Repeat Estimation of IV with strong instruments and a sample of 100,  $y = x$**

| Variable | Obs    | Mean     | Std. Dev. | Min       | Max      |
|----------|--------|----------|-----------|-----------|----------|
| reg_x    | 10,000 | 1.16961  | .2110978  | .4058651  | 1.957093 |
| reg_se_x | 10,000 | .2099269 | .0215022  | .1375987  | .3160827 |
| ivreg_x  | 10,000 | 1.000649 | .2622388  | -.1446195 | 1.98582  |
| iv_se_x  | 10,000 | .260052  | .0334994  | .1586964  | .456611  |

### **Finite Samples and a weak instrument**

```
cap program drop weakreg
program weakreg, rclass
clear
set obs `1'
gen z = rnormal()
gen w = rnormal()
gen x = z*.2 + rnormal() + w
gen u = rnormal()
gen y = x + w + u*5
reg y x
return scalar reg_x = _b[x]
return scalar reg_se_x = _se[x]
ivreg y (x=z)
return scalar ivreg_x = _b[x]
return scalar iv_se_x = _se[x]
end
```

```

simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///      ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x)
///
,rep(10000): weakreg 100
sum

```

Figure 2 shows the same set up as our last results, but this time with a much weaker instrument. So we are again taking 10,000 randomly created datasets, running the regressions on a sample size of 100 in each dataset and seeing what the average results are. Here OLS is again biased, with IV being biased by a much large amount (IV is giving a biased estimate of the coefficient now the instrument is weak and the sample is small). However the even more important thing to note here is the standard error in the IV regression.... It is 74350. So with a weak instrument and a small sample IV is biased and the standard errors are terrible. OLS would have been much much better.

**Figure 2: Monte Carlo Repeat Estimation of IV with weak instruments and a sample of 100,  $y = x$**

| Variable | Obs    | Mean     | Std. Dev. | Min       | Max      |
|----------|--------|----------|-----------|-----------|----------|
| reg_x    | 10,000 | 1.490666 | .3591648  | -.0411166 | 2.860064 |
| reg_se_x | 10,000 | .3574395 | .0363166  | .2364828  | .511851  |
| ivreg_x  | 10,000 | 9.408526 | 734.3879  | -1243.433 | 72797.4  |
| iv_se_x  | 10,000 | 74350.4  | 7188877   | .7055197  | 7.19e+08 |

What happens if we try and increase the sample size a little in each of the 10,000 repetitions? This is what the following figures show. We first move from 100 observations per simulation to 300, with the results in Figure 3.

```

simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///
ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x) ///
,rep(10000): weakreg 300
sum

```

Figure 3 shows that IV is still pulling out a biased estimate of x, but downwards biased this time, and the standard error is still huge at 342. OLS is still much better in terms of the standard error and is not too much more biased.

**Figure 3: Monte Carlo Repeat Estimation of IV with weak instruments and a sample of 300,  $y = x$**

| Variable | Obs    | Mean     | Std. Dev. | Min      | Max      |
|----------|--------|----------|-----------|----------|----------|
| reg_x    | 10,000 | 1.488467 | .2068948  | .692089  | 2.39728  |
| reg_se_x | 10,000 | .2051238 | .0118226  | .1598531 | .2590476 |
| ivreg_x  | 10,000 | .8600619 | 34.82768  | -2315.27 | 1525.686 |
| iv_se_x  | 10,000 | 342.2435 | 19401.34  | .5512501 | 1530097  |

Increasing the sample size again up to 500 now we can see that the coefficient estimate from the IV regressions is starting to get much closer to the true value and is better than the OLS estimates. The IV standard error is still very large however and we would obviously not be able to comment on if there was a significant relationship between x any y at this point.

```
simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///
ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x) ///
, rep(10000): weakreg 500
sum
```

**Figure 4: Monte Carlo Repeat Estimation of IV with weak instruments and a sample of 500,  $y = x$**

| Variable | Obs           | Mean            | Std. Dev.       | Min             | Max             |
|----------|---------------|-----------------|-----------------|-----------------|-----------------|
| reg_x    | <b>10,000</b> | <b>1.491677</b> | <b>.1587471</b> | <b>.8390567</b> | <b>2.13764</b>  |
| reg_se_x | <b>10,000</b> | <b>.1584343</b> | <b>.0070924</b> | <b>.1341993</b> | <b>.1898102</b> |
| ivreg_x  | <b>10,000</b> | <b>1.137308</b> | <b>35.87997</b> | <b>-1378.52</b> | <b>3307.648</b> |
| iv_se_x  | <b>10,000</b> | <b>163.0325</b> | <b>14190.25</b> | <b>.5106497</b> | <b>1403402</b>  |

Increasing the sample size once again up to 750 observations per sample gives the results we see in Figure 5. Here we can see the standard error for the IV estimator has dropped significantly and the point estimate is starting to get tighter. The model fits much better at a sample size of 750 than it does at 500. The only problem would be the IV estimator still has such large variation that both the OLS estimates and 0 would be included in most confidence intervals.

```
simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///
ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x) ///
, rep(10000): weakreg 750
sum
```

**Figure 5: Monte Carlo Repeat Estimation of IV with weak instruments and a sample of 750,  $y = x$**

| Variable | Obs           | Mean            | Std. Dev.       | Min              | Max             |
|----------|---------------|-----------------|-----------------|------------------|-----------------|
| reg_x    | <b>10,000</b> | <b>1.489861</b> | <b>.1305232</b> | <b>.9236135</b>  | <b>1.973313</b> |
| reg_se_x | <b>10,000</b> | <b>.1292407</b> | <b>.0046574</b> | <b>.1126836</b>  | <b>.1522837</b> |
| ivreg_x  | <b>10,000</b> | <b>.9463559</b> | <b>1.111938</b> | <b>-18.49012</b> | <b>7.89286</b>  |
| iv_se_x  | <b>10,000</b> | <b>1.094023</b> | <b>1.437748</b> | <b>.465523</b>   | <b>95.62006</b> |

Moving up to 1000 observations per sample gives the results seen in Figure 6 below. The IV estimator is now clearly stronger than the OLS estimator with a non-biased coefficient estimate and while it still has a larger standard error this will always be the case. 750 and 1000 both seem to be adequate sample sizes to run IV on with a weak instrument in this case, with small sample sizes showing that OLS would be better performing even if it is biased.

```

simulate reg_x=r(reg_x) reg_se_x=r(reg_se_x) ///
ivreg_x=r(ivreg_x) iv_se_x=r(iv_se_x) ///
, rep(10000): weakreg 1000
sum

```

**Figure 6: Monte Carlo Repeat Estimation of IV with weak instruments and a sample of 1000,  $y = x$**

| Variable | Obs           | Mean            | Std. Dev.       | Min             | Max             |
|----------|---------------|-----------------|-----------------|-----------------|-----------------|
| reg_x    | <b>10,000</b> | <b>1.48817</b>  | <b>.1116647</b> | <b>1.065875</b> | <b>1.876478</b> |
| reg_se_x | <b>10,000</b> | <b>.1119513</b> | <b>.0035783</b> | <b>.0999686</b> | <b>.1277457</b> |
| ivreg_x  | <b>10,000</b> | <b>.9695334</b> | <b>.8901244</b> | <b>-5.33989</b> | <b>6.038361</b> |
| iv_se_x  | <b>10,000</b> | <b>.8894818</b> | <b>.3106881</b> | <b>.4439676</b> | <b>7.431123</b> |

What is our take away point from today's investigation? While IV is a very useful tool, we do need to be careful when using a weak instrument on a small sample size. The sample size where this becomes an issue will be dependent on how strong the correlation is between your independent variable and your instrument, but we saw today that with sample sizes of 500 or less there are some major issues. Don't be afraid to run IV regressions, but try your best to find either a strong instrument or a large sample size!

Don't worry if you didn't understand all the code today, hopefully the results tell the story for this pretty clearly. As always I'm more than happy to answer any questions on this concept for those interested, and I will see you next week for a look at some panel data. Have a good week!

*David Morris*