Assignment 1

Advanced Econometrics 1

Stepan Svoboda, Jan Hynek

October 5, 2017

1 Introduction

In this assignment we investigate the behavior of the Hausman endogeneity test using Monte Carlo simulation. We try to simulate the impact of the true underlying parameters on our $\hat{\beta}_{OLS}$ and $\hat{\beta}_{2SLS}$ estimators and their variance.

2 Simulation

In our simulation we vary the parameters ρ, γ and π_1 and investigate their impact on our estimators and thus also on the Hausman test. We also vary the number of instruments and observations to investigate the impact of these parameters. Lastly we keep σ_z^2 and number of iterations per Monte Carlo simulation constant at 2 and 10000, respectively. The instrument strength is mainly driven by the π_1, σ_z^2 and number of observations. Keeping the σ_z^2 constant means the instrument strength is determined only by the π_1 and the number of observations.

The significance level of our Hausman test is throughout the paper set to 0.05 and the Hausman statistics is asymptotically χ^2 distributed. The H_0 is that $\hat{\beta}_{OLS}$ is consistent and more efficient than $\hat{\beta}_{2SLS}$ and H_1 is that $\hat{\beta}_{OLS}$ is inconsistent while $\hat{\beta}_{2SLS}$ is consistent.

The Hausman test statistic gets higher for two reasons. The first is difference between our estimators and second is the difference between their variances. If $\hat{\beta}_{2SLS}$ has very high variance then it is likely that even for quite biased $\hat{\beta}_{OLS}$ we might not have enough evidence to reject the H_0.

	H_0	H_1
$\hat{\beta}_{OLS}$	consistent & more efficient	inconsistent
$\hat{\beta}_{2SLS}$	consistent	consistent

2.1 Part 1 – γ and $\rho = 0$

In the first part of our simulations we fix both γ and ρ to be zero and thus ensure that both $\hat{\beta}_{OLS}$ and $\hat{\beta}_{2SLS}$ are consistent. We look at the results of these tests and find that in case of a high concentration parameter (i.e. good instrument) we are close to the 5% test *size*. This is not true in the case of a weak instrument as the variance of $\hat{\beta}_{2SLS}$ is greater than in the case of strong instrument and our test statistic is smaller, meaning we reject the null hypothesis in far less cases than we should, given our chosen significance level. We also see that changing the number of instruments does not seem to affect the test results significantly. Theoretically lower r should lead to lower variance of $\hat{\beta}_{2SLS}$ and lower *size* of the Hausman test, which does not fully hold in our simulation but the outcomes are within a reasonable range.

The number of observations, N, is important so that we have valid asymptotic variance and consistency of our $\hat{\beta}_{2SLS}$ estimator. During our simulations in the case of $N = 50$ the variance of the

$\hat{\beta}_{2SLS}$ estimator was lower on several occasions than the variance of $\hat{\beta}_{OLS}$ and the corresponding Hausman statistics were thus negative.

2.2 Part 2 – $\gamma = 0$ and $\rho \neq 0$

In this part we have set the parameters driving the consistency of our estimators in a way that $\hat{\beta}_{2SLS}$ is consistent and $\hat{\beta}_{OLS}$ is inconsistent. Thus we are investigating the *power* of the Hausman test for each set of parameters, i.e. the instrument strength and degree of endogeneity.

We investigate the behavior of Hausman test for two degrees of endogeneity – $\rho = 0.25$ & 0.5. We try to find out the *power* of the test for several different sets of parameters defining different instrument strength (i.e. concentration parameter). This is done mainly by changing π_1 but we have looked also at the impact of changes in number of observations and number of instruments.

We have run three MC simulations (6 – 8) for $\rho = 0.25$ and we can see that in the case of not very strong endogeneity the *power* of the test is not particularily high. When our concentration parameter is low – approximately 4 the test detects endogeneity only in 1.2% of cases. In the cases when it is stronger – approximately 25 and 100 – the *power* of the test is still not large – 16.5 and 53% approximately.

For $\rho = 0.5$ we ran six MC simulations (9 – 14) and tested the *power* of the Hausman test based on different values of π_1 and we also investigated what happens when the number of observations or instruments changes.

The first three simulations study the behavior when only π_1 changes and our results behave similarly to the case when $\rho = 0.25$. We get for $\pi_1 = \{0.1, 0.25, 0.5\}$ *power* of the Hausman test equal to approximately $3.5, 66.5$ & 99.5%. The difference being that the greater degree of endogeneity is more easily detected by the same instruments with same strength.

The last three MC simulations are aimed at investigating the situation when ρ and π_1 are kept constant and we change the number of observations and number of instruments. The results are in line with our expectations – half the number of observations leads to much lower power of the test (due to higher variance of $\hat{\beta}_{2SLS}$) and the significantly larger number of instruments leads to slightly worse performance as the instruments enlarge the variance of our $\hat{\beta}_{2SLS}$.

2.3 Part 3 – γ and $\rho \neq 0$

In this section we fix both γ and ρ to be greater than zero. The Hausman test is invalid in this case as neither of our estimators is consistent. Results of MC simulations 15 and 16 are quite interesting. The rejection rates are somewhat similar to the results in **Part 1** and if one were to make an assumption that instruments are exogenous in this case (even though they are not) these results would be easily mistaken for a conclusion when both $\hat{\beta}_{OLS}$ and $\hat{\beta}_{2SLS}$ are consistent. This mistake might be understandable, especially in the case 15 where the correlation of instrument and the disturbances is not that high.

3 Conclusion

We would like to sum up the influence of each parameter on the Hausman test:

- The parameter γ determines whether we can use the test or not. If $\gamma \neq 0$ then our test statistic is invalid because the $\hat{\beta}_{2SLS}$ is inconsistent and we do not have an estimator to test the $\hat{\beta}_{OLS}$ against.
- ρ determines if our $\hat{\beta}_{OLS}$ is consistent or not. The Hausman test is used to test whether $\rho \neq 0$.

- The number of observations, N, determines two things. We need it to be sufficiently large for our $\hat{\beta}_{2SLS}$ to be asymptotically valid and consistent and it also lowers the variance of our estimators and makes our statistic more precise.

- The number of instruments, r, does not change the results significantly as it has only a modest impact on the variance of $\hat{\beta}_{2SLS}$ and the overall Hausman statistic is not significantly impacted. Larger r leads however to slightly lower *power* of the Hausman test.

- π_1 is the last parameter of importance to us. The greater the π_1 the greater the concentration parameter (i.e. instrument strength). Higher concentration parameter leads to higher *power* of the Hausman test. That is due to lower variance of $\hat{\beta}_{2SLS}$ when the concentration parameter is high.

4 Table

The table below describes the *power* or *size* of the test, with the exception of the last two cases. In these cases the test is invalid and we have it to illustrate the possible confusion with the valid instances of the test and the dangers of using an invalid instrument.

simulation	r	N	ρ	γ	π_1	**rejection rate**	$\overline{\tau^2}$
1	3	50	0	0	0.5	5.75%	24.98
2	3	200	0	0	0.5	5.14%	99.88
3	1	200	0	0	0.5	5.11%	100.00
4	5	200	0	0	0.5	5.08%	100.10
5	3	200	0	0	0.1	0.98%	4.00
6	3	200	0.25	0	0.1	1.19%	4.00
7	3	200	0.25	0	0.25	16.55%	24.98
8	3	200	0.25	0	0.5	52.99%	99.97
9	3	200	0.5	0	0.1	3.30%	4.00
10	3	200	0.5	0	0.25	66.64%	24.97
11	3	200	0.5	0	0.5	99.5%	100.03
12	3	200	0.5	0	0.38	95.07%	57.78
13	3	100	0.5	0	0.38	65.31%	28.8
14	10	200	0.5	0	0.38	91.5%	57.72
15	3	200	0.1	0.1	0.5	4.90%	99.87
16	3	200	0.2	0.2	0.5	5.86%	100.05