Alastair Hall **ECON61001: Econometric Methods** Semester 1

Study Group Questions # 8

In this study group assignment, you explore how the sampling properties of the Instrumental Variables estimators depend on the "strength" of the instruments. Our equation of interest is the following linear regression model,

$$y_i \;=\; x_i\beta_0 \,+\, u_i, \qquad i = 1, 2, \ldots, N, \tag{1}$$

where x_i is a scalar variable that is generated via the "first-stage" regression equation,

$$x_i \;=\; z_i'\gamma_0 \,+\, w_i, \tag{2}$$

where z_i is a $q \times 1$ vector for some $q > 2$ and γ_0 is the vector of finite coefficients. It is convenient to set $v_i = (u_i, w_i)'$. Assume the following conditions hold.

Assumption 1 *(i) $\{z_i\}_{i=1}^N$ is fixed in repeated samples with $\|z_i\| < \infty$ for all i and $lim_{N\to\infty} N^{-1}\sum_{i=1}^N z_i z_i' = Q_{zz}$, a finite, pd matrix of constants; (ii) $\{v_i\}_{i=1}^N$ is a sequence of i.i.d. random vectors with a normal distribution with mean equal to 0_2 and variance-covariance matrix equal to Σ, with*

$$\Sigma \;=\; \begin{bmatrix} \sigma_u^2 & \sigma_{u,w} \\ \sigma_{w,u} & \sigma_w^2 \end{bmatrix}.$$

While our primary focus is on the properties of the 2SLS estimator, part of the analysis uses the OLS estimator as a comparator. To present the formulae for the two estimators in the context of our model above, we introduce the following notation. Let y, x, u and w to be $N \times 1$ vectors whose i^{th} elements are respectively y_i, x_i, u_i and w_i; let Z be the $N \times q$ matrix with i^{th} row z_i', and $P_z = Z(Z'Z)^{-1}Z'$. The generic formula for the OLS and 2SLS estimator from lectures specialize in the context of our model here to the following:

$$\hat{\beta}_{OLS} \;=\; \frac{x'y}{x'x}, \tag{3}$$

$$\hat{\beta}_{2SLS} \;=\; \frac{x'P_z y}{x'P_z x}. \tag{4}$$

As discussed in the Tutorial session, the approximate bias of the OLS and 2SLS estimators is given by the following expressions:

$$bias(\hat{\beta}_{OLS}) \;\approx\; \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{1}{\mu/N \,+\, 1}\right), \tag{5}$$

$$bias(\hat{\beta}_{2SLS}) \;\approx\; \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{q-2}{\mu}\right), \tag{6}$$

where μ is the *concentration parameter* discussed in the Tutorial session that is,

$$\mu = \frac{\gamma_0' Z' Z \gamma_0}{\sigma_w^2}.$$

Please enter the details of the group:

Group name	Group Matrix
Student ID 1	11335127
Student ID 2	10710007
Student ID 3	10704589
Student ID 4	11465531
Student ID 5	

1. Show that $\mu \to \infty$ and $\mu/N \to c$ as $N \to \infty$ where c is a finite positive constant that you must specify as part of your answer. Note: you may quote results established in the Tutorial session but explain clearly how you apply them.

Examine $\mu = \frac{\gamma_0' Z' Z \gamma_0}{\sigma_w^2}$ with Z is a $N \times q$ matrix and γ_0 is $q \times 1$.

We can rewrite $\mu = \frac{\gamma_0' Z'}{\sigma_w} \times \frac{Z \gamma_0}{\sigma_w} = b'b$ with $b = \frac{Z\gamma_0}{\sigma_w}$ is a $N \times 1$ matrix,

$\Rightarrow \mu = \sum_{i=1}^{N} b_i^2$ with b_i is the i^{th} row of matrix b; μ is non-negative.
(μ is positive if $\gamma \neq 0$)

Therefore, it follows that as $N \to \infty$; $\sum_{i=1}^{N} b_i^2 \to \infty$ or $\mu \to \infty$ when $\gamma \neq 0$.

Examine $\frac{\mu}{N} = \frac{\gamma_0' Z' Z \gamma_0}{\sigma_w^2 N} = \frac{\gamma_0'}{\sigma_w} \times \frac{Z'Z}{N} \times \frac{\gamma_0}{\sigma_w} = A' \frac{Z'Z}{N} A$; $A = \frac{\gamma_0}{\sigma_w}$, a $q \times 1$ matrix of constant.

Then, $\lim_{N \to \infty} \frac{\mu}{N} = A' Q_{ZZ} A$ because $Z'Z = \sum_{i=1}^{N} z_i z_i'$ and $\lim_{N \to \infty} \sum_{i=1}^{N} z_i z_i' = Q_{ZZ}$

We can rewrite the quadratic form $A' Q_{ZZ} A$ as $\sum_{j=1}^{q} \sum_{i=1}^{q} Q_{ij} A_i A_j$ in which:

- Q_{ij} is the i^{th} - j^{th} element of Q_{ZZ}
- A_i is the i^{th} element of $A = \frac{\gamma_0}{\sigma_w}$
- A_j is the j^{th} element of A.

$\Rightarrow \lim_{N \to \infty} \frac{\mu}{N} = \sum_{j=1}^{q} \sum_{i=1}^{q} Q_{ij} A_i A_j$ which is a non-negative constant.
(it's positive if $\gamma_0 \neq 0$).

2. Under what condition is the OLS estimator approximately unbiased that is, its approximate bias is zero. Interpret this condition in the context of the model described above.

We have $\text{bias}(\hat{\beta}_{OLS}) \approx \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{1}{\mu/N + 1}\right)$

From Question 1, we know that μ/N converges to a finite non-negative constant.
Therefore, $\text{bias}(\hat{\beta}_{OLS}) = 0 \Leftrightarrow \sigma_{u,w} = 0$, which means u and w are uncorrelated.

In our setting, u is the error term from the structural equation (1) and w is the error term from the first stage regression (2).
Combined with the notion from Tutorial Question 1 that $E[x_i u_i] = \sigma_{u,w}$ we can interpret that the OLS estimator is approximately unbiased if x_i is exogenous.

3. Under what conditions is the 2SLS estimator approximately unbiased. Interpret these conditions in the context of the model described above.

We have $\text{bias}(\hat{\beta}_{2SLS}) \approx \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{q-2}{\mu}\right)$ with $q > 2$.

There are three cases for which $\text{bias}(\hat{\beta}_{2SLS}) \simeq 0$

Case 1: $N \longrightarrow \infty$ and $\sigma_{u,w} \neq 0$

As $N \to \infty$, $\mu \to \infty$ (Question 1)

$$\lim_{N \to \infty} \text{bias}(\hat{\beta}_{2SLS}) \approx \lim_{N \to \infty} \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{q-2}{\mu}\right) = \left(\frac{\sigma_{u,w}}{\sigma_w^2}\right) \times 0 = 0$$

Case 2: N is not sufficiently large and $\sigma_{u,w} = 0$.

Then, similar to Question 2, $\text{bias}(\hat{\beta}_{2SLS}) \approx 0$.

Case 3: Both N is large and $\sigma_{u,w} = 0$ can exist at the same time.

In our setting, the conditions under which the 2SLS estimator is approximately unbiased includes:

1) x_i is endogenous but the sample size is very large, toward ∞.
2) x_i is exogenous in small samples.
3) x_i is exogenous and the sample size is large.

4. Define the relative approximate bias of 2SLS to OLS as:

$$rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS}) = \frac{bias(\hat{\beta}_{2SLS})}{bias(\hat{\beta}_{OLS})}.$$

What is the differential of $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS})$ with respect to a change in μ holding all else constant? Interpret your result.

Examine $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS}) = \dfrac{bias(\hat{\beta}_{2SLS})}{bias(\hat{\beta}_{OLS})} = \dfrac{\left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{q-2}{\mu}\right)}{\left(\frac{\sigma_{u,w}}{\sigma_w^2}\right)\left(\frac{1}{\mu/N+1}\right)}$

$\Rightarrow$ $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS}) = \left(\frac{q-2}{\mu}\right)\left(\frac{\mu}{N}+1\right) = \frac{q-2}{N} + \frac{q-2}{\mu} > 0$

$\Rightarrow$ $d(rbias) = \frac{2-q}{\mu^2}\, d\mu$ is the differential of rbias with respect to μ.

$\Rightarrow$ Holding all else equal, a change in μ associates with $\frac{2-q}{\mu^2}$ change in rbias.

Since $q > 2 \Rightarrow \frac{2-q}{\mu^2} < 0$. Thus, μ and rbias move in opposite directions.

We also have that $\mu = \frac{\gamma_0' Z'Z \gamma_0}{\sigma_w^2}$ meaning μ is increasing with $|\gamma_0|$, keeping all else unchanged

As μ is the non-centrality parameter related to F-statistics to test $H_0: \gamma_0 = 0$ vs. $H_1: \gamma_0 \neq 0$. When γ_0 is substantially large, under H_1, μ is also large.

Moreover, when the sample size $N \to \infty$, $\mu \to \infty$ as well if $\gamma_0 \neq 0$.

Therefore, we can interpret that the relative approximate bias is reduced when z_i is highly correlated with x_i. This effect will enhance when the sample size is large.

In conclusion, if the variation of x_i is well-explained by z_i, which means the relevance condition holds, relative bias between 2SLS and OLS estimator is reduced.

In this case, rbias will tend towards 0 if N goes towards ∞.

5. What happens to $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS})$ as μ becomes close to zero holding all else constant?

We have $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS}) = \frac{q-2}{N} + \frac{q-2}{\mu}$

As $\mu \to 0 \Rightarrow \frac{q-2}{\mu} \to \infty \Rightarrow rbias \longrightarrow \infty$.

Under $H_0: \gamma_0 = 0$, μ tends towards 0 (even when the sample size is large).
Which means that if z_i is not relevant to x_i, $rbias(\hat{\beta}_{2SLS}, \hat{\beta}_{OLS})$ tends to be large.
Even in case of weak instrument where z_i is weakly relevant to x_i, μ is small making rbias substantially large.
Therefore, holding all else equal, if the variance in x_i is not well-explained by z_i, relative approximate bias between OLS and 2SLS estimator is large.

6. Use your answers to Questions 2-5 to evaluate the following statement: "In a linear regression model with an endogenous regressor it is always preferable to base inferences on an IV estimator than to base inferences on the OLS estimator."

In the presence of endogeneity, we know (from Question 2) that OLS is not a good estimator since the approximate bias is not zero.
On the other hand, 2SLS estimator may be able to reduce the bias in large samples (Question 3). This is because when the correlation between Z and X significantly differs 0, the non-centrality parameter, μ, will grow as sample size grows, leading to smaller bias. In that case, when Z well-explains X and the sample size is large, we can expect the relative bias between 2SLS estimator and OLS estimator to tend towards 0 (Question 4) and the 2SLS estimator is less biased than the OLS estimator.
However, when Z and X are not strongly relevant, holding all else equal, the relative bias will become large (Question 5).
Therefore, only when we have Z that is strongly relevant to X and a large sample size that the use of 2SLS is preferable to OLS in case of endogeneity given all else unchanged.