International School of Economics at TSU
Econometrics II
Lasha Chochua

Problem Set 2

Instructions: You are encouraged to solve the problems before the recitation. Additionally, you are encouraged to work in groups. It is **not mandatory** to submit solutions unless stated otherwise. However, if you would like to share your solution, I would be happy to review it.

Problem 1 Define the **2SLS residuals**:

$$\hat{\mathbf{e}} = \mathbf{Y}_1 - \mathbf{X}\hat{\beta}_{\text{2sls}}$$

Define the **fitted values** from the first stage:

$$\hat{\mathbf{X}} = \mathbf{P}_Z\mathbf{X} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$$

Show that:

$$\hat{\mathbf{X}}'\hat{\mathbf{e}} = 0$$

Hint: Use the definition of $\hat{\beta}_{\text{2sls}}$ and the fact that $\hat{\mathbf{X}}' = \mathbf{X}'\mathbf{P}_Z$ (since $\mathbf{P}_Z$ is symmetric).

Problem 2: The model is

$$y_i = \mathbf{x}_i'\beta + e_i \quad (1)$$
$$\mathbb{E}(\mathbf{z}_i e_i) = 0 \quad (2)$$

The dimensions are: $\mathbf{x}_i$, $\mathbf{z}_i$, and β are $k \times 1$, $k > 1$, and y_i and e_i are 1×1. Let

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{xx} & \mathbf{Q}_{xz} \\ \mathbf{Q}_{zx} & \mathbf{Q}_{zz} \end{bmatrix} = \begin{bmatrix} \mathbb{E}(\mathbf{x}_i\mathbf{x}_i') & \mathbb{E}(\mathbf{x}_i\mathbf{z}_i') \\ \mathbb{E}(\mathbf{z}_i\mathbf{x}_i') & \mathbb{E}(\mathbf{z}_i\mathbf{z}_i') \end{bmatrix}$$

Assume both $\mathbf{Q}_{xx}$ and $\mathbf{Q}_{zz}$ have full rank k.

Let $\hat{\beta}$ be the least-squares estimate obtained by regressing y_i on $\mathbf{x}_i$, and let $\tilde{\beta}$ be the 2SLS estimator obtained by estimation of (1) using the instrument $\mathbf{z}_i$.

1. Find

$$\delta = \operatorname*{plim}_{n\to\infty} \left(\hat{\beta} - \tilde{\beta}\right)$$

2. Suppose that in addition to (1) and (2),

$$\mathbb{E}(\mathbf{x}_i e_i) = 0 \quad (3)$$

Quite simply, what does this condition mean? What is δ under this assumption?

3. Write the difference $\hat{\beta} - \tilde{\beta}$ as a function of sample moments of $\mathbf{x}_i$, $\mathbf{z}_i$, and e_i.
4. Under (1)-(3), find the asymptotic distribution of

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{pmatrix} \mathbf{x}_i \\ \mathbf{z}_i \end{pmatrix} e_i$$

as $n \to \infty$.

5. Under (1)-(3), find the asymptotic distribution of

$$\sqrt{n}\left(\hat{\beta} - \tilde{\beta}\right) \quad \text{as } n \to \infty.$$

6. Suppose that

$$\mathbb{E}(e_i^2 \mid \mathbf{x}_i, \mathbf{z}_i) = \sigma^2 \tag{4}$$

How does the asymptotic variance from question 5 simplify under (4)?

7. Propose an estimator of the asymptotic variance under (4).
8. Propose a test statistic for (3) under (4) and find its asymptotic distribution under the assumption that $\mathbf{Q} > 0$.
9. Describe how to use this statistic to test the hypothesis that $\mathbf{x}_i$ is exogenous.
10. Show where $\mathbf{Q} > 0$ is used in the answer to question 8.

Problem 3 Let the structural equation be:

$$Y_1 = X'\beta + e, \quad \mathbb{E}[Ze] = 0$$

with 2SLS estimator:

$$\hat{\beta}_{2SLS} = (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1$$

Let $\mathbf{C}$ be an invertible $\ell \times \ell$ matrix and define $\tilde{Z} = \mathbf{C}Z$.

(a) Show that replacing Z with $\tilde{Z} = \mathbf{C}Z$ leaves $\hat{\beta}_{2SLS}$ unchanged. That is, show:

$$(\tilde{\mathbf{X}}'\tilde{\mathbf{Z}}(\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}})^{-1}\tilde{\mathbf{Z}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{Z}}(\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}})^{-1}\tilde{\mathbf{Z}}'\mathbf{Y}_1 = \hat{\beta}_{2SLS}$$

where $\tilde{\mathbf{Z}} = \mathbf{Z}\mathbf{C}'$ is the $n \times \ell$ sample matrix of transformed instruments.

(b) Using the projection matrix $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, show that $\mathbf{P}_{\tilde{Z}} = \mathbf{P}_Z$. What does this say geometrically?

(c) Using (b), write the 2SLS estimator in the form $\hat{\beta}_{2SLS} = (\mathbf{X}'\mathbf{P}_Z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{Y}_1$ and argue directly that invariance to $\mathbf{C}$ follows from $\mathbf{P}_{\tilde{Z}} = \mathbf{P}_Z$.

(d) Now show that the asymptotic variance $V_\beta^0 = (Q_{XZ}Q_{ZZ}^{-1}Q_{ZX})^{-1}\sigma^2$ is also invariant to replacing Z with $\tilde{Z} = \mathbf{C}Z$. Derive each step explicitly showing how the $\mathbf{C}$ matrices cancel.

(e) Does invariance hold if $\mathbf{C}$ is not invertible? Explain what goes wrong, relating your answer to the rank condition $\text{rank}(\mathbb{E}[ZX']) = k$.

Problem 4 Let $\theta = r(\beta) \in \mathbb{R}^q$ be a differentiable function of the 2SLS estimator, with $\hat{\theta} = r(\hat{\beta}_{2SLS})$. Denote $\mathbf{R} = \frac{\partial}{\partial\beta}r(\beta)'$ $(k \times q)$. We test $H_0 : \theta = \theta_0$ using the Wald statistic:

$$W = n(\hat{\theta} - \theta_0)'\hat{V}_\theta^{-1}(\hat{\theta} - \theta_0)$$

(a) Starting from Theorem 5 in the related slides:

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, V_\theta), \quad V_\theta = \mathbf{R}'V_\beta\mathbf{R}$$

Show step by step that $W \xrightarrow{d} \chi_q^2$ under H_0. You must explicitly use the definition of the χ^2 distribution.

(b) At which step does $\hat{V}_\theta \xrightarrow{p} V_\theta$ enter the argument? Which theorem justifies replacing V_θ with $\hat{V}_\theta$ without changing the limiting distribution?

(c) Now suppose $e_i \mid X_i, Z_i \sim \mathcal{N}(0, \sigma^2)$ (normality). Consider the special case $q = 1$ (scalar θ). Show that the t-statistic:

$$t = \frac{\sqrt{n}(\hat{\theta} - \theta_0)}{\sqrt{\hat{V}_\theta}}$$

satisfies $t^2 = W$ and derive its exact distribution under normality. At which step does normality play a role?

(d) Now consider general $q \geq 1$ under normality. The F-statistic is defined as:

$$F = \frac{W}{q}$$

Show that $F \sim F(q, n - k_1 - 2k_2)$ by expressing W as a ratio of two independent χ^2 random variables, each divided by their degrees of freedom. State clearly what each χ^2 corresponds to.

(e) Why does the F distribution reduce to the asymptotic χ^2_q / q as $n \to \infty$? Show this using the definition $F(q, m) \xrightarrow{d} \chi^2_q / q$ as $m \to \infty$.

Problem 5 (Sargan Test)

Part A – Intuition

- Recall the IV moment condition: $\mathbb{E}[\mathbf{Z}e] = 0$
- We have ℓ instruments but only k parameters to estimate
- Explain intuitively:
    1. Why overidentification ($\ell > k$) creates a **testable restriction**
    2. Why $\hat{\alpha} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\hat{\mathbf{e}}$ is the natural object to test
    3. Why the degrees of freedom of the test should be $\ell - k$, not ℓ

Part B – Derivation

- Under $H_0 : \mathbb{E}[\mathbf{Z}e] = 0$ and homoskedasticity $\mathbb{E}[e^2 \mid \mathbf{Z}] = \sigma^2$, show that:

$$S = \frac{\hat{\mathbf{e}}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\hat{\mathbf{e}}}{\hat{\sigma}^2} \xrightarrow{d} \chi^2_{\ell-k}$$

- **Hints**:
    1. Write $\hat{\mathbf{e}} = \mathbf{e} - \mathbf{X}(\hat{\beta}_{\text{2sls}} - \beta)$ and show $\frac{1}{\sqrt{n}}\mathbf{Z}'\hat{\mathbf{e}} \xrightarrow{d} \frac{1}{\sqrt{n}}\mathbf{Z}'\mathbf{e}$ asymptotically
    2. Apply the CLT to $\frac{1}{\sqrt{n}}\mathbf{Z}'\mathbf{e}$
    3. Count the moment conditions used up by estimation