

Practice Test

2SLS with Multiple Instruments

Framework: Consider the linear structural model

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + u_i, \quad i = 1, \dots, n,$$

where $\mathbf{x}_i \in \mathbb{R}^K$ is a vector of endogenous regressors, $\boldsymbol{\beta} \in \mathbb{R}^K$ is the parameter vector of interest, and u_i is the error term. Let $\mathbf{z}_i \in \mathbb{R}^L$ be a vector of instruments. Assume that there are no additional exogenous covariates.

Question 1. Assume $L = K$. State the conditions required for $\mathbf{z}_i$ to be a valid instrument for $\mathbf{x}_i$. Derive the population moment restrictions that identify $\boldsymbol{\beta}$ in vector form.

Let $\mathbf{Y} \in \mathbb{R}^{n \times 1}$, $\mathbf{X} \in \mathbb{R}^{n \times K}$, and $\mathbf{Z} \in \mathbb{R}^{n \times L}$ and $\boldsymbol{\Sigma} := E[\mathbf{Z}'\mathbf{X}]$. Derive a closed-form expression for the population parameter $\boldsymbol{\beta}$ in terms of the population moments. then derive the Method-of-Moments estimator $\hat{\boldsymbol{\beta}}_{\text{MoM}}$, all in matrix notation.

Question 2. Assume now that $L > K$ (the overidentified case). Explain mathematically why the population parameter $\boldsymbol{\beta}$ can no longer be obtained by directly solving the population moments.

Let $\mathbf{g}_n(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i (y_i - \mathbf{x}_i' \boldsymbol{\beta})$ denote the sample moment function. Suppose one researcher proposes to estimate $\boldsymbol{\beta}$ by minimizing

$$\hat{\boldsymbol{\beta}} = \arg \min_{\mathbf{b} \in \boldsymbol{\Theta}} \mathbf{g}_n(\mathbf{b})' \mathbf{g}_n(\mathbf{b}),$$

while another proposes to minimize instead

$$\hat{\boldsymbol{\beta}} = \arg \min_{\mathbf{b} \in \boldsymbol{\Theta}} \mathbf{g}_n(\mathbf{b})' \hat{\mathbf{W}} \mathbf{g}_n(\mathbf{b}),$$

for some positive definite matrix $\hat{\mathbf{W}}$. Discuss what is the rationale for introducing a weighting matrix.

Question 3. Consider the estimator proposed by the second researcher. What is the corresponding population criterion function, $\mathbf{Q}(\mathbf{b})$, that she has in mind? Define it in terms of the population moment function $\mathbf{g}(\mathbf{b})$ and a symmetric positive definite matrix $\mathbf{W}$. Show that $\mathbf{Q}(\mathbf{b}) \geq 0$ for all $\mathbf{b} \in \boldsymbol{\Theta}$ and that $\mathbf{Q}(\mathbf{b}) = 0$ if and only if $\mathbf{b} = \boldsymbol{\beta}$.

Then, conclude by arguing that if there exist two minimizers $\mathbf{b}_1$ and $\mathbf{b}_2$, then it must be $\mathbf{b}_1 = \mathbf{b}_2$. (*hint: show that $\mathbf{g}(\mathbf{b}) = \mathbb{E}[\mathbf{z}_i \mathbf{x}_i'](\boldsymbol{\beta} - \mathbf{b})$, and impose a rank condition to prove that $\boldsymbol{\beta}$ is the unique minimizer of the population criterion function.*)

Question 4. Derive the first-order condition for the sample criterion function proposed by the second researcher. Let

$$\mathbf{g}_n(\mathbf{b}) = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i (y_i - \mathbf{x}'_i \mathbf{b}) \quad \text{and} \quad \mathbf{Q}_n(\mathbf{b}) = \mathbf{g}_n(\mathbf{b})' \hat{\mathbf{W}} \mathbf{g}_n(\mathbf{b}),$$

where $\hat{\mathbf{W}}$ is symmetric positive definite. Compute $\nabla_{\mathbf{b}} \mathbf{Q}_n(\mathbf{b})$ step by step, being explicit about the use of the chain rule for a quadratic form in a vector-valued function.

In particular, let $\mathbf{Q}(\boldsymbol{\beta}) = \mathbf{x}(\boldsymbol{\beta})' \mathbf{A} \mathbf{x}(\boldsymbol{\beta})$ where $\mathbf{x}(\boldsymbol{\beta}) \in \mathbb{R}^L$ and $\mathbf{A} = \mathbf{A}'$ use the following results:

$$\nabla_{\boldsymbol{\beta}} \mathbf{Q}_n(\boldsymbol{\beta}) = \left(\frac{\partial \mathbf{x}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} \right)' \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}' \mathbf{A} \mathbf{x}) \quad \frac{\partial}{\partial \mathbf{x}} \mathbf{x}' \mathbf{A} \mathbf{x} = 2 \mathbf{A} \mathbf{x} \quad \frac{\partial}{\partial \boldsymbol{\beta}'} \mathbf{C} \boldsymbol{\beta} = \mathbf{C}.$$

Derive a closed-form solution for $\hat{\boldsymbol{\beta}}$.

Question 5. We now establish consistency of the estimator using the General Consistency Theorem for extremum estimators. Argue first why the identification condition has already been verified in the previous question.

The second requirement is uniform convergence of the criterion function, that is, $\sup_{\mathbf{b} \in \Theta} |\mathbf{Q}_n(\mathbf{b}) - \mathbf{Q}(\mathbf{b})| = o_p(1)$. Suppose that (i) the parameter space Θ is compact, (ii) $\hat{\mathbf{W}} \xrightarrow{p} \mathbf{W}$, (iii) $\mathbf{g}_n(\mathbf{b})$ is continuous in $\mathbf{b}$, and (iv) $\mathbb{E}[\sup_{\mathbf{b} \in \Theta} \|\mathbf{z}_i (y_i - \mathbf{x}'_i \mathbf{b})\|] < \infty$. Prove uniform convergence by proceeding through the following steps:

1. Using compactness of Θ and continuity of $\mathbf{g}_n(\mathbf{b})$, prove that $\sup_{\mathbf{b} \in \Theta} \|\mathbf{g}(\mathbf{b})\| < \infty$.
2. Using the ULLN Theorem, prove that $\sup_{\mathbf{b} \in \Theta} \|\mathbf{g}_n(\mathbf{b}) - \mathbf{g}(\mathbf{b})\| \xrightarrow{p} 0$.
3. Show that $\sup_{\mathbf{b} \in \Theta} \|\mathbf{g}_n(\mathbf{b})\| \xrightarrow{p} \sup_{\mathbf{b} \in \Theta} \|\mathbf{g}(\mathbf{b})\|$, by applying the triangle inequality and the inequality $|||a| - |c||| \leq \|a - c\|$.
4. Finally, prove that $\sup_{\mathbf{b} \in \Theta} |\mathbf{Q}_n(\mathbf{b}) - \mathbf{Q}(\mathbf{b})| = o_p(1)$, by adding and subtracting $\mathbf{g}_n(\mathbf{b})' \mathbf{W} \mathbf{g}_n(\mathbf{b})$ and using the matrix operator norm together with the consistency of $\hat{\mathbf{W}}$.

Question 6. Was there another way to prove consistency?

Question 7. Derive the empirical Hessian Matrix defined as $\hat{\mathbf{H}} = \frac{\partial^2}{\partial \mathbf{b} \partial \mathbf{b}'} \mathbf{Q}_n(\mathbf{b})$, use the following:

$$\hat{\mathbf{H}} = \frac{\partial}{\partial \mathbf{b}'} (\nabla_{\mathbf{b}} \mathbf{Q}_n(\mathbf{b})) \quad \frac{\partial}{\partial \boldsymbol{\beta}'} \mathbf{C} \boldsymbol{\beta} = \mathbf{C}$$

Question 8. Derive the asymptotic linear representation of the estimator obtained in the previous questions. Assume that $\hat{\mathbf{W}} \xrightarrow{u} \mathbf{W}$ and simplify notation by using $\boldsymbol{\Sigma} = \mathbb{E}[\mathbf{Z}' \mathbf{X}]$ and its sample counterpart $\hat{\boldsymbol{\Sigma}}$.

Starting from the closed-form expression of the estimator in matrix notation, derive an explicit expression for $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi(\mathbf{z}_i, u_i) + o_p(1)$ in terms of population moments.

State sufficient conditions under which a multivariate Central Limit Theorem applies to $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Then apply Slutsky's theorem and continuous mapping arguments to derive the asymptotic

distribution of $\hat{\beta}$, and provide the explicit expression for the asymptotic variance.

Question 9. Prove that the choice $\mathbf{W}^* = \mathbf{\Omega}^{-1} = E[\mathbf{Z}'\mathbf{U}\mathbf{U}'\mathbf{Z}]$ yields the most efficient GMM estimator. Recall that the asymptotic variance of the GMM estimator with weighting matrix $\mathbf{W} \succ 0$ is

$$\mathbf{V}(\mathbf{W}) = (\mathbf{\Sigma}'\mathbf{W}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}'\mathbf{W}\mathbf{\Omega}\mathbf{W}\mathbf{\Sigma}(\mathbf{\Sigma}'\mathbf{W}\mathbf{\Sigma})^{-1},$$

Let $\mathbf{W}^* = \mathbf{\Omega}^{-1}$ and denote the corresponding asymptotic variance by $\mathbf{V}^*$. First, show that the efficient asymptotic variance simplifies to $\mathbf{V}^* = (\mathbf{\Sigma}'\mathbf{\Omega}^{-1}\mathbf{\Sigma})^{-1}$, then define:

$$\mathbf{g} = \mathbf{A}\mathbf{Z}'\mathbf{U} \quad \mathbf{g}^* = \mathbf{B}\mathbf{Z}'\mathbf{U},$$

where $\mathbf{A} = (\mathbf{\Sigma}'\mathbf{W}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}'\mathbf{W}$ and $\mathbf{B} = (\mathbf{\Sigma}'\mathbf{\Omega}^{-1}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}'\mathbf{\Omega}^{-1}$. Write $\mathbf{g} = \mathbf{g}^* + (\mathbf{g} - \mathbf{g}^*)$ and expand $\mathbb{V}[\mathbf{g}]$. Show that the cross terms vanish and conclude the proof.