Advanced Econometrics I: Principles of Econometrics
Second Retake Exam

23 March 2017

- This exam has three pages and parts.
- Don't copy the instructions.
- Start each part on a new sheet of paper (one sheet comprises four A4-sized pages)
- Good luck!

1 Gauß-Markov Theorem (12 points)

Prove the following version of the Gauß-Markov Theorem: The estimator $Lb = L\left(X'X\right)^{-1}X'y$, where $L \in \mathbb{R}^{s\times K}$ is the best linear unbiased estimator for $L\beta$ in the linear model $y = X\beta + \varepsilon$, where $X \in \mathbb{R}^{N\times K}$, $N \geq K$, is non-stochastic and of full column rank, $\mathbb{E}\left(\varepsilon\right) = 0$, and the error terms are homoskedastic and uncorrelated.

1.1 Preliminary questions Statement and definitions (4,5 points)

1. (1 point) Let $\hat{\theta}$ be an estimator, i.e. a random variable from the sample space to the parameter space, for the population parameter θ. Give the formal definition of an unbiased estimator.
2. (1 point) We consider the class of all unbiased estimators $\tilde{\theta}$ for the population parameter θ. Give the formal definition of the efficient estimator in this class of estimators.
3. (0,5 points) Under the assumptions above, $\mathbb{E}\left(\varepsilon\right) = 0$ implies $\mathbb{E}\left(\varepsilon|X\right) = 0$. True or false?
4. (0,5 points) The statement "$\mathbb{E}\left(\varepsilon\right) = 0$ implies $\mathbb{E}\left(\varepsilon|X\right) = 0$" is true in general (for stochastic X). True or false?
5. (0,5 points) In the case of a stochastic matrix X, the condition $\mathbb{E}\left(\varepsilon_i x_i'\right) = 0$ implies that $\mathbb{E}\left(\varepsilon|X\right) = 0$. True or false?
6. (1 point) Which of the following are valid for stating that the error terms are homoskedastic and uncorrelated? (The question only counts if ALL correct possibilities are indicated.)
   (a) $\mathbb{V}\left(\varepsilon\right) = \sigma^2 I_N$
   (b) $\mathbb{V}\left(\varepsilon\right) = \sigma^2 I_K$
   (c) $\mathbb{V}\left(\varepsilon_i\right) = \sigma^2$ for all i and $\mathbb{C}\left(\varepsilon_i, \varepsilon_j\right) = 0$ for $i \neq j$
   (d) $\mathbb{V}\left(\varepsilon_i\right) = \sigma_i^2$ for all i and $\mathbb{C}\left(\varepsilon_i, \varepsilon_j\right) = 0$ for $i \neq j$

1.2 Part 1 of the proof (3,5 points)

1. (1 point) Calculate the expectation of Lb.
2. (2 points) Consider the linear estimator $L\check{b} = Dy$. Derive the conditions on the matrix D such that a linear estimator $L\check{b} = Dy$ is unbiased for $L\beta$.
3. (0,5 points) What dimensions does D have?

1.3 Part 2 of the proof (4 points)

1. (2 points) Calculate the variance of Lb.
2. (2 points) Show that for matrices D, $X \in \mathbb{R}^{N\times K}$, $L \in \mathbb{R}^{s\times K}$satisfying $DX = L$, the decompositition

$$DD' = \left(L\left(X'X\right)^{-1}X'\right)\left(L\left(X'X\right)^{-1}X'\right)' + \left(D - \left[L\left(X'X\right)^{-1}X'\right]\right)\left(D - \left[L\left(X'X\right)^{-1}X'\right]\right)'$$

holds.

2 Lagrange Multiplier Test (12 points)

We consider a model given as a set of conditional densities $f\left(y_1,\ldots,y_N|x_1,\ldots,x_N;\theta\right)$ which are parameterized by a K-dimensional parameter vector $\theta = \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}$ where $\theta_1 \in \mathbb{R}^{K-J}$ and $\theta_2 \in \mathbb{R}^J$, $K, J > 0$, θ and where $y_i|x_i$ are independently and identically distributed. The **unconstrained ML estimator** $\hat{\theta}$ is obtained as maximizing argument of $\log\left(L\left(\theta\right)\right) = \sum_{i=1}^N \log\left(L_i(\theta)\right)$ where $\log\left(L\left(\theta\right)\right)$ is the log-likelihood function corresponding to $f\left(y_1,\ldots,y_N|x_1,\ldots,x_N;\theta\right) = \prod_{i=1}^N f\left(y_i|x_i;\theta\right)$, while the **restricted ML estimator** $\tilde{\theta}$ is obtained as maximizing argument of the constrained maximization problem $\sum_{i=1}^N \log\left(L_i(\theta)\right)$ subject to $\theta_2 = q$. The constrained maximization problem can also be solved by maximizing the Lagrangian $H(\theta,\lambda) = \sum_{i=1}^N \log\left(L_i(\theta)\right) - \lambda'\left(\theta_2 - q\right)$ with respect to θ and λ from which we obtain maximizing arguments $\tilde{\theta} = \begin{pmatrix}\tilde{\theta}_1\\ q\end{pmatrix}$ and $\tilde{\lambda}$.

The Lagrangian multiplier is asymptotically normally distributed, i.e. $\sqrt{N}\left(\frac{1}{N}\tilde{\lambda}_N\right) \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0, I^{22}\left(\tilde{\theta}\right)^{-1}\right)$ where $I^{22}\left(\tilde{\theta}\right)$ is obtained as a submatrix of an estimator of the inverse of the covariance matrix

$$I(\theta) = \begin{pmatrix} I_{11}(\theta) & I_{12}(\theta)\\ I_{21}(\theta) & I_{22}(\theta)\end{pmatrix} = \lim_{N\to\infty} -\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \frac{\partial^2 \log\left(L_i\left(\theta\right)\right)}{\partial\theta\partial\theta'}\right)$$

of the asymptotic distribution of the scores. In particular,

$$I(\theta)^{-1} = \begin{pmatrix} I^{11}(\theta) & I^{12}(\theta)\\ I^{21}(\theta) & I^{22}(\theta)\end{pmatrix} = \begin{pmatrix} * & *\\ * & \left(I_{22}(\theta) - I_{21}(\theta)I_{11}(\theta)^{-1}I_{12}(\theta)\right)^{-1}\end{pmatrix}$$

Assume in the following that the information matrix is estimated from the first derivatives, i.e. $\hat{I}_G\left(\tilde{\theta}\right) = \frac{1}{N}\sum_{i=1}^N s_i\left(\tilde{\theta}\right)s_i\left(\tilde{\theta}\right)'$ where $s_i\left(\tilde{\theta}\right) = \left.\frac{\partial\log(L_i(\theta))}{\partial\theta}\right|_{\theta=\tilde{\theta}}$ are the individual score contributions evaluated at the constrained MLE.

1. (4 points) Show that the LM test statistic $\xi_{LM} = \frac{1}{N}\tilde{\lambda}_N' I^{22}\left(\tilde{\theta}\right)\tilde{\lambda}_N$ can be written as

$$\xi_{LM} = \frac{1}{N}\begin{pmatrix}\sum_{i=1}^N s_i^{(1)}\left(\tilde{\theta}\right)' & \sum_{i=1}^N s_i^{(2)}\left(\tilde{\theta}\right)'\end{pmatrix}\begin{pmatrix} I^{11}\left(\tilde{\theta}\right) & I^{12}\left(\tilde{\theta}\right)\\ I^{21}\left(\tilde{\theta}\right) & I^{22}\left(\tilde{\theta}\right)\end{pmatrix}\begin{pmatrix}\sum_{i=1}^N s_i^{(1)}\left(\tilde{\theta}\right)\\ \sum_{i=1}^N s_i^{(2)}\left(\tilde{\theta}\right)\end{pmatrix}, \tag{1}$$

where the column vector $s_i^{(1)}\left(\tilde{\theta}\right) = \left.\frac{\partial\log(L_i(\theta))}{\partial\theta_1}\right|_{\theta=\tilde{\theta}}$ denotes the first derivative of $\log\left(L_i\left(\theta\right)\right)$ with respect to the unrestricted parameters θ_1, and the column vector $s_i^{(2)}\left(\tilde{\theta}\right) = \left.\frac{\partial\log(L_i(\theta))}{\partial\theta_2}\right|_{\theta=\tilde{\theta}}$ are the scores pertaining to the restricted parameters θ_2, and explain your calculations.

   - Hint: In order to get started, you may use the fact that $\tilde{\lambda}_N = \sum_{i=1}^N s_{i2}\left(\tilde{\theta}\right)$. Subsequently, extend this vector of scores and do some algebraic manipulations.

2. (2 points) Show that equation (1) can be written as

$$\xi_{LM} = \iota' S \left[S'S\right]^{-1} S\iota \tag{2}$$

where

$$S = \begin{pmatrix} s_1\left(\tilde{\theta}\right)' \\ \vdots \\ s_i\left(\tilde{\theta}\right)' \\ \vdots \\ s_N\left(\tilde{\theta}\right)' \end{pmatrix}$$

is the matrix of individual score contributions and $\iota = (1, \ldots, 1)' \in \mathbb{R}^{N\times 1}$ and explain your calculations (2 points).

3. (2 points) Write equation (2) as an uncentered R-squared of an auxiliary regression and describe the auxiliary regression.

Now, consider the model $y_i = x_i\beta + z_i\gamma + \varepsilon_i$ where $\varepsilon_i \sim NID\left(0, \sigma^2\right)$, (x_i, z_i) and ε_i are independent, and $\beta, \gamma \in \mathbb{R}$. We want to test $H_0:\ \gamma = 0$. The associated log-likelihood function is $\log\left(L\left(\beta,\gamma,\sigma^2\right)\right) = -\frac{N}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\sum_{i=1}^{N}\left(\frac{y_i - x_i\beta - z_i\gamma}{\sigma}\right)^2$ and its derivatives are

$$\begin{pmatrix} \frac{\partial}{\partial\beta} \\ \frac{\partial}{\partial\gamma} \\ \frac{\partial}{\partial\sigma^2} \end{pmatrix} \log\left(L\left(\beta,\gamma,\sigma^2\right)\right) = \begin{pmatrix} \sum_{i=1}^{N}\left(\frac{y_i - x_i\beta - z_i\gamma}{\sigma^2}\right) x_i \\ \sum_{i=1}^{N}\left(\frac{y_i - x_i\beta - z_i\gamma}{\sigma^2}\right) z_i \\ -\frac{N}{2\sigma^2} + \frac{1}{2}\sum_{i=1}^{N}\left(\frac{y_i - x_i\beta - z_i\gamma}{\sigma^2}\right)^2 \end{pmatrix}.$$

4. (2 points) Give expressions for the residuals $\tilde{\varepsilon}_t$ and $\hat{\varepsilon}_t$ implied by the restricted ML and unrestricted MLE respectively.

5. (0.5 points) Which residuals correspond to H_0?

6. (1.5 points) What are the FOC for $\tilde{\beta}$, $\tilde{\gamma}$, and $\tilde{\sigma}^2$ of the associated constrained optimization problem?

3 Short Questions (12 points)

We consider the linear regression model

$$y = X\beta + \varepsilon$$

where $X \in \mathbb{R}^{N\times K}$ is non-stochastic, $\mathbb{E}(\varepsilon) = 0$ and $\mathbb{V}(\varepsilon) = \sigma^2\Psi$, where $\sigma^2 > 0$ and Ψ is positive definite.

1. (3 points) Calculate the covariance matrix of the OLS estimator $b = \left(X'X\right)^{-1}X'y$ in the model above .
2. (1 point) The OLS estimator is unbiased in this model. True or false?
3. (1 point) The OLS estimator is efficient in this model. True or false?
4. (1 point) The weighted least squares (WLS) estimator gives a different weight to each row of X. True or false?
5. (1.5 point) Define the size of a test.
6. (1.5 point) Define the power of a test.
7. Assume that you want to test the hypothesis $H_0:\ \beta = \beta_0$ and that the p-value is 0.075.
   (a) (1 point) H_0 is rejected for a test with size 10%. True of false?
   (b) (1 point) H_0 is rejected for a test with size 5%. True or false?
8. (1 point) Assume that you are using the OLS estimator b for β in a model with heteroskedastic errors. If one uses heteroskedasticity-consistent standard errors, the resulting (Wald) test statistic is asymptotically appropriate. True or false?