MEI MT Problem Set 8[1]

Timothee Carayol

January 21, 2010

[1] Available on http://personal.lse.ac.uk/carayolt/ec402.htm

Question 1

Question

- $y = X\beta + \epsilon$ with X: $T \times k$ such that X full column rank and $E(\epsilon|X) = 0$.
- We consider four different specifications for the variance of the error term:
  - (a) $E(\epsilon_t^2|X) = \theta_0 + \theta_1 x_{3t}^2 + \frac{\theta_2}{x_{5t}^4}$, and $\forall s \neq t, E(\epsilon_s \epsilon_t|X = 0)$.
  - (b) $E(\epsilon_t^2|X) = \sigma_\nu^2 n_t$, with $n_1, .., n_T$ known, and $\forall s \neq t, E(\epsilon_s \epsilon_t|X = 0)$.
  - (c) $\epsilon_t = \rho\epsilon_{t-1} + \nu_t$, with $|\rho| < 1$, ν_t i.i.d. with (conditional) expected value 0 and finite (conditional) variance σ_ν^2. (The $\{\epsilon_t\}$ sequence follows an $AR(1)$ process).
  - (d) $\epsilon_t = \nu_t + \lambda\nu_{t-1}$, with $|\rho| < 1$, ν_t i.i.d. with (conditional) expected value 0 and finite (conditional) variance σ_ν^2. (The $\{\epsilon_t\}$ sequence follows a $MA(1)$ process).
- For each specification, we will first write down the variance-covariance matrix of ϵ, and second explain how we would implement the FGLS estimator.

Question 1

Answer

- Reminder: GLS = General Least squares; FGLS = Feasible General Least Squares.
- The BLUE estimator under $A4\Omega$ is $\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$; but Ω is typically unknown so in general GLS has in general only a theoretical appeal.
- It is sometimes possibly to estimate Ω consistently, though (say by $\Omega(\hat{\theta})$), which allows us to use instead $\hat{\beta}_{FGLS} = (X'\Omega(\hat{\theta})^{-1}X)^{-1}X'\Omega(\hat{\theta})^{-1}Y$.
- The FGLS estimator has messy finite sample properties. It may be biased and non-linear. However it can be shown that as long as $\Omega(\hat{\theta})$ is consistent for Ω, FGLS and GLS are *asymptotically* equivalent, which implies that it is consistent and asymptotically efficient for β.

Question 1

Answer

- (a) $E(\epsilon_t^2|X) = \theta_0 + \theta_1 x_{3t}^2 + \frac{\theta_2}{x_{5t}^4}$, and $\forall s \neq t, E(\epsilon_s \epsilon_t|X) = 0$.
- Off-diagonal terms will be zeros (no autocorrelation); diagonal terms are given by $E(\epsilon_t^2|X)$ above.
- Hence $V(\epsilon) = c^2\Omega =$

$$\begin{pmatrix} \theta_0 + \theta_1 x_{31}^2 + \frac{\theta_2}{x_{51}^4} & 0 & \cdots & 0 \\ 0 & \theta_0 + \theta_1 x_{32}^2 + \frac{\theta_2}{x_{52}^4} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_0 + \theta_1 x_{3T}^2 + \frac{\theta_2}{x_{5T}^4} \end{pmatrix}.$$

Question 1

Answer

- To implement FGLS:
  - The parameters in Ω are $\theta = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix}$. If we find a consistent estimator for θ, we are done.
  - To obtain this $\hat{\theta}$: first find $\hat{\beta}_{OLS}$ from the original equation. We know it is not an efficient estimator, but it is consistent. Denote the residual as $\hat{\epsilon}$.
  - Now: perform the OLS regression of $\hat{\epsilon_t}^2$ on a constant, x_{3t} and $\frac{1}{x_{5t}^4}$. *That the resulting $\hat{\theta}$ is then consistent for θ is true, but not trivial: to be entirely thorough this would need a short proof. I prove it IN THIS CASE in the appendix; for (c) and (d) the proof would be similar in spirit.*.
  - Define $\Omega(\hat{\theta})$ as the expression for Ω from previous slide, where θ is replaced by its consistent estimator $\hat{\theta}$.
  - Define $\hat{\beta}_{FGLS} = (X'\Omega(\hat{\theta})^{-1}X)^{-1}X'\Omega(\hat{\theta})^{-1}Y$.

Question 1

Answer

- (b) $E(\epsilon_t^2|X) = \sigma_\nu^2 n_t$, with $n_1, .., n_T$ known, and $\forall s \neq t, E(\epsilon_s \epsilon_t|X) = 0$.
- Off-diagonal terms will be zeros (no autocorrelation); diagonal terms are given by $E(\epsilon_t^2|X)$ above.
- Hence $V(\epsilon) = \sigma_\nu^2 \begin{pmatrix} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & n_T \end{pmatrix} = \sigma_\nu^2 \Omega.$

Question 1

Answer

- To implement FGLS:
  - The parameters $n_1, .., n_T$ are known, so the only unknown parameter here is σ_ν. Since it multiplies every element of the Ω matrix, we can apply GLS rather than FGLS here: σ_ν does *not* appear in the GLS formula.

Question 1

Answer

- (c) $\epsilon_t = \rho\epsilon_{t-1} + \nu_t$, with $|\rho| < 1$, ν_t i.i.d. with (conditional) expected value 0 and finite (conditional) variance σ^2_ν. (The $\{\epsilon_t\}$ sequence follows an $AR(1)$ process).
- This time we will have autocorrelation. In particular, $cov(\epsilon_t, \epsilon_{t-1}) = cov(\rho\epsilon_{t-1} + \nu_t, \epsilon_{t-1}) = \rho\sigma^2_\epsilon$. Likewise, $\forall s, t$, $cov(\epsilon_t, \epsilon_s) = \rho^{|t-s|}\sigma^2_\varepsilon$. We could compute σ^2_ϵ too: but the nice thing here is that we do not need it.
- Hence $V(\epsilon) = c^2\Omega = \sigma^2_\epsilon \begin{pmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{T-1} & \cdots & \rho & 1 \end{pmatrix}$.

Question 1

Answer

- To implement FGLS:
  - The only parameter in Ω (besides σ_ϵ, which cancels out in the expression for $\hat{\beta}_{FGLS}$), is ρ.
  - Denote the residual as $\hat{\epsilon}$ the OLS residual from the original equation.
  - Now: perform the OLS regression of $\hat{\epsilon_t}$ on $\hat{\epsilon_{t-1}}$. *That the resulting $\hat{\rho}$ is then consistent for ρ is true, but not trivial: to be entirely thorough this would need a short proof*.
  - Define $\Omega(\hat{\rho})$ as the expression for Ω from previous slide, where ρ is replaced by its consistent estimator $\hat{\rho}$.
  - Define $\hat{\beta}_{FGLS} = (X'\Omega(\hat{\rho})^{-1}X)^{-1}X'\Omega(\hat{\rho})^{-1}Y$.

Question 1

Answer

- (d) $\epsilon_t = \nu_t + \lambda \nu_{t-1}$, with ν_t i.i.d. with (conditional) expected value 0 and finite (conditional) variance σ^2_ν. (The $\{\epsilon_t\}$ sequence follows a $MA(1)$ process).
- $V(\epsilon_t) = \sigma^2_\epsilon = V(\nu_t + \lambda\nu_{t-1}) = \sigma^2_\nu(1+\lambda^2)$, i.e. $\sigma^2_\nu = \frac{\sigma^2_\epsilon}{1+\lambda^2}$. Likewise, this time we will have autocorrelation, but only first order. i.e., $cov(\epsilon_t, \epsilon_{t-1}) = cov(\nu_t + \lambda\nu_{t-1}, \nu_{t-1} + \lambda\nu_{t-2}) = \lambda\sigma^2_\nu = \sigma^2_\epsilon \frac{\lambda}{1+\lambda^2}$; but, $\forall s, t$ such that $|t-s| > 1$, $cov(\epsilon_t, \epsilon_s) = 0$.
- Hence *in fine*,

$$V(\epsilon) = c^2\Omega = \sigma^2_\epsilon \begin{pmatrix} 1 & \frac{\lambda}{1+\lambda^2} & 0 & \cdots & 0 \\ \frac{\lambda}{1+\lambda^2} & 1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \frac{\lambda}{1+\lambda^2} \\ 0 & \cdots & 0 & \frac{\lambda}{1+\lambda^2} & 1 \end{pmatrix}.$$

Question 1

Answer

- To implement FGLS:
  - The only parameter in Ω (besides σ_ν, which cancels out in the expression for $\hat{\beta}_{FGLS}$), is $\frac{\lambda}{1+\lambda^2}$.
  - We know that $cov(\epsilon_t, \epsilon_{t-1}) = \frac{\lambda}{1+\lambda^2}\sigma\epsilon^2$, which implies $corr(\epsilon_t, \epsilon_{t-1}) = \frac{\lambda}{1+\lambda^2}$.
  - This, intuitively, should lead us to consider the sample correlation between $\hat{\epsilon_t}$ and $\hat{\epsilon_{t-1}}$ and use it as an estimator for $\frac{\lambda}{1+\lambda^2}$. *That the resulting estimator is then consistent for* $\frac{\lambda}{1+\lambda^2}$ *is true, but not trivial: to be entirely thorough this would need a short proof*.
  - Define $\hat{\Omega}$ as the expression for Ω from previous slide, where $\frac{\lambda}{1+\lambda^2}$ is replaced by its consistent estimator.
  - Define $\hat{\beta}_{FGLS} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y$.

Question 2

Question

- $y = X\beta + \epsilon$, X: $N \times 5$.
- Want to test $H_0 : \begin{array}{c} \beta_1\beta_2 = 1 \\ \beta_3 = 4\beta_4 - 2 \end{array}$.

Answer

- Problem: we are used to dealing with linear hypotheses. But this time, the first equation is non-linear in β.
- Rewrite $H_0 : g(\beta) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ where $g(\beta) = \begin{pmatrix} \beta_1\beta_2 - 1 \\ \beta_3 - 4\beta_4 + 2 \end{pmatrix}$.

- **Theorem (Delta Method)**

  If x_n is a sequence of random variables such that $\sqrt{n}(x_n - x) \to_d \mathcal{N}(0, \Sigma)$ and g is continuous and differentiable, then $\sqrt{n}(g(x_n) - g(x)) \to_d \mathcal{N}(0, \frac{\partial g(x)}{\partial x'} \Sigma \frac{\partial g'(x)}{\partial x})$.

Question 2

Answer

- Here: we know (e.g. from PS7 question 1) that, under $A1, A2, A3Rsru$, $\sqrt{N}(\hat{\beta}_{OLS} - \beta) \to_d \mathcal{N}(0, aVar(\hat{\beta}))$. (Where $aVar$ stands for "asymptotic variance".
  $aVar(\hat{\beta}) = \sigma_\varepsilon^2 E(x_t x_t')^{-1}$ under A4GM; and we have
  $Var(\hat{\beta}) = \frac{aVar(\hat{\beta})}{N}$).
- Hence it is also the case that
  $\sqrt{N}(g(\hat{\beta}) - g(\beta)) \to_d \mathcal{N}(0, g_\beta(\hat{\beta}) aVar(\hat{\beta}) g_\beta(\hat{\beta})')$. Note that under H_0, $g(\beta) = 0_2$.
- This in turn implies that
  $Ng(\hat{\beta})' \left(g_\beta(\hat{\beta}) aVar(\hat{\beta}) g_\beta(\hat{\beta})' \right)^{-1} g(\hat{\beta}) \to_d \chi^2(2)$. (Remember that if $x \sim \mathcal{N}(0, \Sigma)$, then $x'\Sigma^{-1}x \sim \chi^2(rank(\Sigma))$
- Equivalently $g(\hat{\beta})' \left(g_\beta(\hat{\beta}) Var(\hat{\beta}) g_\beta(\hat{\beta})' \right)^{-1} g(\hat{\beta}) \to_d \chi^2(2)$.

Question 2

Answer

- This asymptotic distribution will also hold (via Slutsky) if we replace $Var(\hat{\beta})$ by a consistent estimator, e.g. $\hat{Var}(\hat{\beta}) = s^2(X'X)^{-1}$ if A4GM holds. We therefore recognize $Q = g(\hat{\beta})' \left(g_\beta(\hat{\beta})\hat{V}(\hat{\beta})g_\beta(\hat{\beta})'\right)^{-1} g(\hat{\beta})$ as the Wald test statistic corresponding to hypothesis H_0.
- Hence, under H_0, $Q \rightarrow_d \chi^2(2)$.

Question 2

Answer

- Alternative approach: likelihood ratio test, based on the likelihood ratio: $LR = \frac{L(\tilde{\beta})}{L(\hat{\beta})}$ where I define $\hat{\beta}$ as the maximum likelihood estimator under the unconstrained model, and $\tilde{\beta}$ as the maximum likelihood estimator under the (non-linear) constrained model. Then it can be shown that $-2ln(LR) \rightarrow_d \chi^2(2)$, and this test is asymptotically equivalent to the Wald test.

Appendix

Proof of consistency of θ in (a)

- Let us consider the difference between $\hat{\theta}$, estimated using the $\hat{\varepsilon}$ as outlined in the slides, and the (hypothetical) $\tilde{\theta}$, estimated using the (unobserved) ε. Let us denote Z the 3×1 matrix of regressors containing an intercept, x_3^2 and $\frac{1}{x_5^4}$.
- Then:

$$\begin{aligned}
\hat{\theta} - \tilde{\theta} &= (Z'Z)^{-1}Z'(\hat{\varepsilon}^2) - (Z'Z)^{-1}Z'(\varepsilon^2) = (Z'Z)^{-1}Z'(\hat{\varepsilon}^2 - \varepsilon^2) \\
&= (\frac{Z'Z}{T})^{-1}(\frac{\sum_t z_t'(\hat{\varepsilon}_t^2 - \varepsilon_t^2)}{T}) \\
&= (\frac{Z'Z}{T})^{-1}(\frac{\sum_t z_t'(2\varepsilon_t x_t'(\beta - \hat{\beta}) + [x_t'(\beta - \hat{\beta})]^2}{T}) \qquad (1)
\end{aligned}$$

Appendix

Proof of consistency of θ in (a)

- In the last equation, I used the fact that:

$$\hat{\varepsilon}_t = y_t - x_t'\hat{\beta} = y_t - x_t'\beta + x_t'(\beta - \hat{\beta}) = \varepsilon_t + x_t'(\beta - \hat{\beta})$$

  implying $\hat{\varepsilon}_t^2 = \varepsilon_t^2 + [x_t'(\beta - \hat{\beta}]^2 + 2\varepsilon_t x_t'(\beta - \hat{\beta})$.
- From (1), it is straightforward to see that the consistency of $\hat{\beta}$ for β implies that $\hat{\theta}$ converges in probability to $\tilde{\theta}$. Since the latter is consistent for θ, this means the former is also consistent, which is what we wanted to prove.