MEI MT Problem Set 4[1]

Timothee Carayol

January 15, 2010

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

Question 1 (a)

Question

$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$; $A1, A2$ hold. (a) Obtain $\hat{\beta}_{1OLS}$ and $\hat{\beta}_{2OLS}$.

Answer

- (a) From last week, $(X'X)^{-1} = \frac{1}{T(\overline{x^2} - \bar{x}^2)} \begin{pmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}$; also have $X'y = \begin{pmatrix} \sum_t y_t \\ \sum_t x_t y_t \end{pmatrix}$.
- Could do (somewhat) tedious algebra to find an expression for $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'y$.

Question 1 (a) (cont)

Answer

- Simplest perhaps is to first find $\hat{\beta}_2 = \frac{1}{T(\overline{x^2}-\bar{x}^2)} \begin{pmatrix} -\bar{x} & 1 \end{pmatrix} \begin{pmatrix} \sum_t y_t \\ \sum_t x_t y_t \end{pmatrix} = \frac{1}{\overline{x^2}-\bar{x}^2}(-\bar{x}\bar{y} + \overline{xy}) = \frac{\hat{cov}(x,y)}{\hat{Var}(x)}$[2].
- From here, a "trick" allows us to get $\hat{\beta}_1$ with no further algebra. Note that $\hat{y} = \hat{\beta}_1 + x\hat{\beta}_2$, so that, taking the average over the sample, $\bar{y} = \hat{\beta}_1 + \bar{x}\hat{\beta}_2 \Rightarrow \hat{\beta}_1 = \bar{y} - \bar{x}\frac{\hat{cov}(x,y)}{\hat{Var}(x)}$.

[2] A formula which holds with one regressor and one intercept only and which you should learn by heart

Question 1 (b)

Question

(b) Obtain $Var(\hat{\beta}_1)$, $Var(\hat{\beta}_2)$, $cov(\hat{\beta}_1, \hat{\beta}_2)$.

Answer

- Already found $Var(\hat{\beta}_1)$, $cov(\hat{\beta}_1, \hat{\beta}_2)$. last week.
- With the same method we would find: $Var(\hat{\beta}_2) = \frac{\sigma_\varepsilon^2}{T(\bar{x^2} - \bar{x}^2)}$

Question 2

Question

In each of five seemingly nonlinear models, explain whether and how we can apply linear regression methods to estimate the unknown parameters.

Answer

- (a) $y_t = \beta_1 x_{2t}^{\beta_2} x_{3t}^{\beta_3} \varepsilon_t$. By taking the logs, we get $log(y_t) = log(\beta_1) + \beta_2 log(x_{2t}) + \beta_3 log(x_{3t}) + log(\varepsilon_t)$. Hence OLS of $y^* = log(y)$ on an intercept, $x_2^* = log(x_2)$ and $x_3^* = log(x_3)$ can estimate $\beta_1^* = log(\beta_1)$, $\beta_2^* = \beta_2$ and $\beta_3^* = \beta_3$ without bias if $E(log(\varepsilon_t)) = 0$ (which is not the same as $E(\varepsilon_t) = 1$, because of Jensen's inequality.)
- (b) $y_t = \beta_1 + \frac{\beta_2}{x_{2t}} + \beta_3 x_{3t}^2 + \varepsilon_t$. OLS of $y^* = y$ on an intercept, $x_2^* = \frac{1}{x_2}$ and $x_3^* = x_3^2$ can estimate $\beta_1^* = \beta_1$, $\beta_2^* = \beta_2$ and $\beta_3^* = \beta_3$ without bias.

Question 2

Answer

- (c) $y_t = \beta_1 \beta_2^{x_{2t}} + \varepsilon_t$. No way to make this linear.
- (d) $log(y_t) = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta 4 x_{2t} x_{3t} + \beta_5 x_{2t}^2 + \beta_6 x_{3t}^2 + \varepsilon_t$. Regress $log(y_t)$ on intercept, x_{2t}, x_{3t}, $x_{2t}x_{3t}$, x_{2t}^2, x_{3t}^2. Note: some students think it's not ok to include a variable and its square in the same model, because they may be highly correlated. Actually makes sense if you suspect the relationship between y and the variable is nonlinear. In most cases, correlation is not much of a problem. $x_{2t}x_{3t}$ is also fine. It is an interaction term, telling us whether, e.g., there is a "complementarity" between x_{2t} and x_{3t}.
- (e) $y_t = (x_{1t}^{\beta_1} + x_{2t}^{\beta_2} + \beta_3)^{\beta_4} + \varepsilon_t$. No way to make this linear.

Question 3

Question

Suppose $y = \beta_0 + X\beta + \varepsilon$; $A1, A2, A3Rmi, A4GM$ hold. (a) Find matrix A such that the differenced formulation $\Delta y = \Delta X\beta + \Delta\varepsilon$ can be written as $Ay = AX\beta + A\varepsilon$. (b) Show that the $\hat{\beta}_\Delta$ from the differenced equation cannot have a lower variance than $\hat{\beta}$ from the levels data.

Answer

- (a) First, note that A has to be $T - 1 \times T$. We want
$$A \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_T \end{pmatrix} = \begin{pmatrix} z_2 - z_1 \\ z_3 - z_2 \\ \vdots \\ z_T - z_{T-1} \end{pmatrix}.$$

Question 3

Answer

- Taking $A = \begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & -1 & 1 \end{pmatrix}$

Question 3

Answer

- $Ai = 0$, so we have to drop the constant, otherwise $A(iX)$ will not be full-rank.
- (b) $\hat{\beta}_\Delta = (X'A'AX)^{-1}X'A'Ay$ is linear in y and unbiased (from $A3Rmi$), hence cannot be better than OLS, from the Gauss-Markov theorem.
- Note 1: that we are talking about estimating $\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix}$, i.e. we do not consider β_0 as it can't be identified from the differenced specification. Hence, what we are comparing to $\hat{\beta}_\Delta$ is, from Frisch-Waugh: $\hat{\beta} = (X'M_iX)^{-1}X'M_iy$.
- Note 2: We know from PS2 question 8 than M_i has the effect of taking the deviations from the mean of the vectors it multiplies.

Question 4

Question

Given estimates of $\hat{\beta}$ and its variance / covariance matrix, test at the 5% level (a) $H_0 : \beta_2 = 1$ vs. $H_1 : \beta_2 < 1$; and (b) $\beta_3 + \beta_4 + \beta_5 = 1$ vs. $\beta_3 + \beta_4 + \beta_5 \neq 1$.

Answer

- Let us denote as $\hat{V}$ the estimate variance covariance matrix of $\hat{\beta}$.
- (a) Under H_0, $T = \frac{\hat{\beta}_2 - 1}{\sqrt{\hat{V}_{11}}} \sim t(140)$[3]. Reject H_0 iff $T < -t_{140,.05} = -1.655$. We find here $T = -16$, so we do reject CRS.

[3] $140 = 145 - 5$

Question 4

Answer

- (b) First note: $Var(\hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5) = Var(\hat{\beta}_3) + Var(\hat{\beta}_4) + Var(\hat{\beta}_5) + 2cov(\hat{\beta}_3, \hat{\beta}_4) + 2cov(\hat{\beta}_3, \hat{\beta}_5) + 2cov(\hat{\beta}_4, \hat{\beta}_5)$. So our estimate for this, given our data, is $\hat{V}_{33} + \hat{V}_{44} + \hat{V}_{55} + 2(\hat{V}_{34} + \hat{V}_{35} + \hat{V}_{45}) = .222$.
- Under H_0, $T = \frac{\hat{\beta}_3+\hat{\beta}_4+\hat{\beta}_5-1}{S.E.(\hat{\beta}_3+\hat{\beta}_4+\hat{\beta}_5)} = \frac{\hat{\beta}_3+\hat{\beta}_4+\hat{\beta}_5-1}{\sqrt{\hat{\beta}_3+\hat{\beta}_4+\hat{\beta}_5}} \sim t(140)$.
- Reject H_0 iff $T < -t_{140,.025} = -1.977$ or $T < t_{140,.025} = 1.977$. We find here $T = -.7575$, so we do not reject homogeneity of degree one in prices.