MEI MT Problem Set 7 Part 2[1]

Timothee Carayol

January 15, 2010

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

Question 3

Question

- True model: $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ with X_1: $N \times k_1$; X_2: $N \times k_2$; $k_1 + k_2 = k$; X_1 and X_2 fixed in repeated samples ($A3f$). Finally, assume that $A4GM$ holds: $V(\varepsilon) = \sigma_\varepsilon^2 I_N$.
- Researcher omits (or cannot observe) X_2, and therefore uses the following estimator for β_1: $\hat{\gamma} = (X_1'X_1)^{-1}X_1'y$. Also uses, as an estimator for the variance of $\hat{\gamma}$: $V_1 = \hat{V}(\hat{\gamma}) = s_1^2(X_1'X_1)^{-1}$ with $s_1^2 = \frac{(y-X_1\hat{\gamma})'(y-X_1\hat{\gamma})}{T-k_1}$
- Note that those are exactly what one would obtain with OLS on the (wrong) specification $y = X_1\beta_1 + \varepsilon$.

Question 3

Question

(a) $\hat{\gamma}$ biased for β_1.

Answer

- $$E(\hat{\gamma}) = E\left((X_1'X_1)^{-1}X_1'y\right) = (X_1'X_1)^{-1}X_1'E\left(X_1\beta_1 + X_2\beta_2 + \varepsilon\right)$$
  $$= (X_1'X_1)^{-1}X_1'\left(X_1\beta_1 + X_2\beta_2 + 0\right) = \beta_1 + \underbrace{(X_1'X_1)^{-1}X_1'X_2\beta_2}_{bias}$$
- Note how the bias relates to the true β_2 on the one hand, and on the OLS estimates of the X_2's regressed on the X_1's on the other hand.

Question 3

Question

(b) $V(\hat{\gamma}) = \sigma_\varepsilon^2 (X_1' X_1)^{-1}$.

Answer

$$\begin{aligned} V(\hat{\gamma}) &= V\left((X_1'X_1)^{-1}X_1'y\right) = (X_1'X_1)^{-1}X_1'V(y)\left((X_1'X_1)^{-1}X_1'\right)' \\ &= (X_1'X_1)^{-1}X_1'E(\varepsilon\varepsilon')X_1(X_1'X_1)^{-1} = \sigma_\varepsilon^2(X_1'X_1)^{-1} \end{aligned}$$

from $A4GM$.

Question 3

Question

(c) s_1^2 biased for σ_ε^2.

Answer

$$
\begin{aligned}
E(s_1^2) &= \frac{1}{T-k_1} E\left((y - X_1\hat{\gamma})'(y - X_1\hat{\gamma})\right) \\
&= \frac{1}{T-k_1} E\left((X_2\beta_2 + \varepsilon)' M_{X_1} (X_2\beta_2 + \varepsilon)\right) \\
&= \frac{1}{T-k_1} E\left[\left((X_2\beta_2)' M_{X_1} (X_2\beta_2)\right) + \left(\varepsilon' M_{X_1} \varepsilon\right) + 2\underbrace{\left(\varepsilon' M_{X_1} X_2\beta_2\right)}_{E(..)=0}\right] \\
&= \frac{1}{T-k_1} \left(E\left[\varepsilon' M_{X_1} \varepsilon\right] + (X_2\beta_2)' M_{X_1} (X_2\beta_2)\right) \\
&= \frac{1}{T-k_1} \left(\sigma_\varepsilon^2 tr(M_{X_1}) + (X_2\beta_2)' M_{X_1} (X_2\beta_2)\right) \\
&= \frac{1}{T-k_1} \left(\sigma_\varepsilon^2 (T-k_1) + (X_2\beta_2)' M_{X_1} (X_2\beta_2)\right) \\
&= \sigma_\varepsilon^2 + \frac{1}{T-k_1} \left((X_2\beta_2)' M_{X_1} (X_2\beta_2)\right) \geq \sigma_\varepsilon^2
\end{aligned}
$$

Question 3

Question

(d) The question here is a bit misleading: "uncorrelated" is not a strong enough assumption to conclude in this context. Assume instead that $X_1'X_2 = 0$ (which would be the sample counterpart of x_{1t} and x_{2t} uncorrelated if we knew that either of them has mean 0, e.g. if we consider deviations from the mean.).

Answer

- $E(\hat{\gamma}) = \beta_1 + \underbrace{(X_1'X_1)^{-1}X_1'X_2\beta_2}_{bias} = \beta_1$ i.e. the bias disappears.
-
$$\begin{aligned} E(s_1^2) &= \sigma_\varepsilon^2 + \frac{1}{T-k_1}\left((X_2\beta_2)'M_{X_1}(X_2\beta_2)\right) \\ &= \sigma_\varepsilon^2 + \frac{1}{T-k_1}\left((X_2\beta_2)'I_n(X_2\beta_2)\right) \end{aligned}$$

  i.e. the bias is at its worst.

Question 3

- What I suspect Vassilis meant:
- The coefficients of a regression of each variable in X_2 on the variables in X_1 are all zero (except maybe the constant).
- Then the bias in $\hat{\gamma}$ is $(X_1'X_1)^{-1}X_1'X_2\beta_2$. Recall that $(X_1'X_1)^{-1}X_1'X_2$ is a $k_1 \times k_2$ matrix whose kth column is the OLS estimator of the regression of the kth column of X_2 on the variables in X_1. If there is a constant in X_1, then the bias in the $k_1 - 1$ last elements of $\hat{\gamma}$ will be zero iff the variables in X_1 have no explanatory power over the variables in X_2.. OR if β_2 is actually zero, e.g. if the omitted variables do not explain the variations in y once you control for X_1.

Question 3

- Textbook example is much simpler to think about, with X_1 containing only one variable beside the constant, and X_2 only one variable. E.g. in a wage equation, think of X_1 as education and of X_2 as intrinsic ability.
- $w_i = \alpha + \beta \ educ_i + \gamma \ ability_i + \varepsilon_i$.
- The problem is that we have at the same time:
  - X_1 and X_2 are (likely) positively correlated.
  - X_2 is (likely) also a determinant of wages BEYOND its effect on education.
- So that if we omit ability in our regression, the coefficient on education will be upwards biased, as it will capture some of the effect of ability on wages.
- In this simple case, expression for omitted variable bias simplifies to: $\frac{cov(ability, education)}{var(education)} * \gamma$.

Question 4

Question

- $y = X\beta + \varepsilon$ with $X : T \times k$ fixed in repeated samples. $A1$, $A2$, $A3f$ hold, and $E(\varepsilon\varepsilon') = \Sigma$ positive definite and symmetric. (i.e. may be heteroskedasticity and autocorrelation in the error term.)
- We consider the usual OLS estimator for β, and estimate its covariance matrix as $s^2(X'X)^{-1}$.

Question 4

Question

(a) $\hat{\beta}$ unbiased.

Answer

$$
\begin{aligned}
E(\hat{\beta}) &= E\left((X'X)^{-1}X'y\right) = E\left((X'X)^{-1}X'(X\beta + \varepsilon\right) \\
&= \beta + (X'X)^{-1}X'E(\varepsilon) = \beta
\end{aligned}
$$

Question 4

Question

(b) $V(\hat{\beta}) = (X'X)^{-1}X'\Sigma X(X'X)^{-1}$.

Answer

$$
\begin{aligned}
V(\hat{\beta}) &= V\left((X'X)^{-1}X'(X\beta + \varepsilon)\right) = V\left((X'X)^{-1}X'\varepsilon\right) \\
&= E\left((X'X)^{-1}X'\varepsilon)(X'X)^{-1}X'\varepsilon)'\right) \\
&= \left((X'X)^{-1}X'\right)E(\varepsilon\varepsilon')\left(X(X'X)^{-1}\right) \\
&= (X'X)^{-1}X'\Sigma X(X'X)^{-1}
\end{aligned}
$$

Question 4

Question

(c) Does the Gauss-Markov theorem allow us to compare the matrix found in (b) and $\sigma_\varepsilon^2(X'X)^{-1}$?

Answer

Gauss-Markov does not tell us anything about how $(X'X)^{-1}X'\Sigma X(X'X)^{-1}$ compares with $\sigma_\varepsilon^2(X'X)^{-1}$. In this context (i.e. with this version of $A4$), the latter matrix does not mean anything interesting, and is actually not even well defined as we haven't defined σ_ε. More interestingly, it can be shown that the GLS estimator has covariance matrix equal to $(X'\Sigma^{-1}X)^{-1}$, which can be shown to be smaller than $V(\hat{\beta}_{OLS})$ (see Vassilis' solution for a proof).