MSc Economics
Econometrics, BST169

ANSWERS

Autumn Semester 2008
Three Hours
Answer FOUR questions

1.

(a) Explain briefly the meaning of a 'consistent estimator' and indicate why consistency is a desirable property of an estimator.

(b) Under what circumstances is the OLS estimator of β in the model

$$y = X\beta + u$$

consistent?

(c) Show that the OLS estimator

$$s^2 = \frac{e'e}{T-K}$$

of σ^2 in the above model is consistent

(d) Discuss ONE example of a situation where the OLS estimator is inconsistent. What estimator would you use in this case?

2. What is meant by an Instrumental Variables (IV) estimator? Examine the circumstances in which an Instrumental Variables estimator is useful.

3. Suppose that β in the linear model

$$y = X\beta + u \quad u \sim N\left(0, \sigma^2 I\right)$$

is partitioned into β_1 and β_2 and that X is conformably partitioned into X_1 and X_2 so the model can be written

$$y = X_1\beta_1 + X_2\beta_2 + u$$

Let b_1^U and b_2^U be the unrestricted OLS estimators of β_1 and β_2 and let b_1^R be the OLS estimator of β_1 under the null hypothesis, $H_0 : \beta_2 = 0$.

(a) Use the normal equations to show that

$$b_1^R - b_1^U = (X_1'X_1)^{-1} X_1'X_2 b_2^U$$

Interpret this result.

Answer:

$$(X_1'X_1)^{-1} X_1'y = b_1^U + (X_1'X_1)^{-1} X_1'X_2 b_2^U \Longrightarrow b_1^R = b_1^U + (X_1'X_1)^{-1} X_1'X_2 b_2^U$$

(b) The Least Squares estimator of β_2 is

$$b_2^U = [X_2'M_1X_2]^{-1} X_2'M_1 y$$

where

$$M_1 = I - X_1 (X_1'X_1)^{-1} X_1'$$

Interpret this result.

Answer: Frisch-Waugh Theorem

(c) Show that, under the null hypothesis, both b_1^R and b_1^U are unbiased. Does this result mean that one should include as many regressors as possible?

Answer:

$$E\left(b_1^R\right) = E\left[(X_1'X_1)^{-1} X_1' (X_1\beta_1 + u)\right] = \beta_1$$

By the Frisch-Waugh Theorem

$$\begin{aligned} b_1^U &= (X_1'M_2X_1)^{-1}X_1'M_2 y \\ &= (X_1'M_2X_1)^{-1}X_1'M_2 (X_1\beta_1 + u) \\ &= \beta_1 + (X_1'M_2X_1)^{-1}X_1'M_2 u \end{aligned}$$

so

$$E\left(b_1^U\right) = \beta_1$$

Restricted estimator is efficient.

(d) Construct and justify a Wald test for the null hypothesis that $\beta_2 = 0$.

Answer: The general form of the Wald test with q restrictions is

$$\left(\hat{\theta} - \theta_0\right)' I\left(\hat{\theta}\right)\left(\hat{\theta} - \theta_0\right) \sim \chi_q^2$$

We can write

$$\hat{\theta} = R\hat{\beta} = [0, I] \begin{bmatrix} b_1^U \\ b_2^U \end{bmatrix} = 0$$

where $R\hat{\beta}$ has the covariance matrix, $RV\left(b^U\right)R' = V\left(b_2^U\right)$. Hence,

$$W = (b_2^u)' \left[\sigma^2 (X_2'M_1X_2)^{-1}\right]^{-1} (b_2^u)$$

Replacing σ^2 with its consistent estimator, s^2

$$W = (b_2^u)' \left[s^2 (X_2'M_1X_2)^{-1}\right]^{-1} (b_2^u) = qF \xrightarrow{a} \chi_q^2$$

4. Consider the model

$$\begin{aligned} y &= X\beta + u \\ u &\sim N\left(0, \sigma^2\Omega\right) \end{aligned}$$

where Ω is a known symmetric positive definite matrix.

(a) Write down the log likelihood for a random sample of size T and solve for the Maximum Likelihood estimators of β and σ^2 given Ω.
Answer:

$$\begin{aligned} L &= -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\sigma^2\Omega\right| - \frac{1}{2}(y - X\beta)'\left(\sigma^2\Omega\right)^{-1}(y - X\beta) \\ &= -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2}\ln|\Omega| - \frac{1}{2}\frac{(y - X\beta)'\Omega^{-1}(y - X\beta)}{\sigma^2} \end{aligned}$$

where

$$\begin{aligned} (y - X\beta)'\Omega^{-1}(y - X\beta) &= (y - X\beta)'\left(\Omega^{-1}y - \Omega^{-1}X\beta\right) \\ &= y'\Omega^{-1}y - y'\Omega^{-1}X\beta - \beta'X'\Omega^{-1}y + \beta'X'\Omega^{-1}X\beta \\ & \quad y'\Omega^{-1}y - 2y'\Omega^{-1}X\beta + \beta'X'\Omega^{-1}X\beta \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= -2X'\Omega^{-1}y + 2X'\Omega^{-1}X\beta = 0 \\ \frac{\partial L}{\partial \sigma^2} &= -\frac{T}{2\sigma^2} + \frac{1}{2}\frac{(y - X\beta)'\Omega^{-1}(y - X\beta)}{\sigma^4} \end{aligned}$$

$$\begin{aligned} \hat{\beta} &= \left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}y \\ \hat{\sigma}^2 &= \frac{\left(y - X\hat{\beta}\right)'\Omega^{-1}\left(y - X\hat{\beta}\right)}{T} \end{aligned}$$

(b) Show that the ML estimators of β and σ^2 can be found from an OLS regression in transformed variables, y^* and X^*.
Answer: Since Ω is p.d. we have $\Omega^{-1} = P'P$, hence

$$\begin{aligned} \left(X'P'PX\right)^{-1}X'P'Py &= \left((PX)'(PX)\right)^{-1}(PX)'Py \\ &= \left(X^{*\prime}X^*\right)^{-1}X^{*\prime}y^* \end{aligned}$$

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\left(y - X\hat{\beta}\right)'P'P\left(y - X\hat{\beta}\right)}{T} \\ &= \frac{\left[P\left(y - X\hat{\beta}\right)\right]'\left[P\left(y - X\hat{\beta}\right)\right]}{T} \\ &= \frac{\left(y^* - X^*\hat{\beta}\right)'\left(y^* - X^*\hat{\beta}\right)}{T} \end{aligned}$$

(c) Suppose that

$$y_t = \beta x_t + u_t$$

while Ω is a diagonal matrix with elements

$$\omega_t = \exp(\gamma z_t)$$

FInd the Maximum likelihood estimators of β, γ and σ^2 and suggest a test for $\gamma = 0$.

Answer: We have

$$L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{\gamma}{2}\sum_{t=1}^{T} z_t - \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - \beta x_t)^2}{\sigma^2 \exp(\gamma z_t)}$$

so

$$\begin{aligned}
\frac{\partial L}{\partial \beta} &= \sum_{t=1}^{T}\frac{(y_t - \beta x_t)\, x_t}{\sigma^2 \exp(\gamma z_t)} = 0 \Longrightarrow \hat{\beta} = \sum_{t=1}^{T}\frac{y_t x_t}{x_t^2} \\
\frac{\partial L}{\partial \gamma} &= -\frac{1}{2}\sum_{t=1}^{T} z_t + \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - \beta x_t)^2 z_t}{\sigma^2 \exp(\gamma z_t)} = 0 \Longrightarrow \sum_{t=1}^{T} z_t \left[1 - \sum_{t=1}^{T}\frac{\left(y_t - \hat{\beta} x_t\right)^2}{\hat{\sigma}^2 \exp(\hat{\gamma} z_t)}\right] = 0 \\
\frac{\partial L}{\partial \sigma^2} &= -\frac{T}{2\sigma^2} + \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - \beta x_t)^2}{\sigma^4 \exp(\gamma z_t)} = 0 \Longrightarrow \hat{\sigma}^2 = \sum_{t=1}^{T}\frac{\left(y_t - \hat{\beta} x_t\right)^2}{T \exp(\hat{\gamma} z_t)}
\end{aligned}$$

The LR test is

$$\begin{aligned}
\hat{L} &= -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\hat{\sigma}^2 - \frac{\hat{\gamma}}{2}\sum_{t=1}^{T} z_t - \frac{T}{2} \\
\tilde{L} &= -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\tilde{\sigma}^2 - \frac{T}{2} \\
-2\ln\lambda &= T\ln\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} - \hat{\gamma}\sum_{t=1}^{T} z_t \sim \chi_1^2
\end{aligned}$$

One might also use the Breusch-Pagan Test or a Wald test.

5. Examine the advantages and disadvantages of using OLS to estimate a regression equation when the error term is heteroscdastic.
Answer: OLS is unbiased and consistent but inefficient. However, it is reasonably simple to use and the covaraince matrix can be adjusted. When the form of the heteroscedasticity is unknown, and a two-step estimator is not necesarily better in small samples.

6. Comment on **TWO** of the following:

(a) Testing for Structural Change

(b) Choosing between Non-Nested Models

(c) Lagrange Multiplier tests

(d) Testing for Autocorrelated Errors

(e) Measures of Predictive Accuracy

7.

(a) Explain what is meant by 'identification' in the context of a simultaneous equations model and give the conditions under which the parameters of an equation are identified.

(b) Consider the following model

$$
\begin{array}{rcll}
C_t & = & \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \alpha_3 \left(W_t^P + W_t^G \right) + u_t^C & \text{Consumption Equation} \\
I_t & = & \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \beta_3 K_{t-1} + u_t^I & \text{Investment Equation} \\
W_t^P & = & \gamma_0 + \gamma_1 X_t + \gamma_2 X_{t-1} + \gamma_3 A_t + u_t^W & \text{Private Wage Bill} \\
X_t & = & C_t + I_t + G_t & \text{Equilibrium Condition} \\
P_t & = & X_t - T_t - W_t^P & \text{Private Profits Identity} \\
K_t & = & I_t + K_{t-1} & \text{Evolution of the Capital Stock}
\end{array}
$$

where C_t, is private sector consumption, I_t is private sector investment spending, W_t^P is the private sector wage bill, P_t s private sector profits, X_t is GNP, G_t, is government non-wage spending, W_t^G s the government's wage bill, T_t s indirect business taxes plus net exports and A_t.is a time trend, The endogenous variables appear on the left hand side of the above equations. The exogenous variables are G_t, W_t^G,T_t, A_t and the constant term while K_{t-1}, $P_{t-1}$and X_{t-1} are predetermined.

i. Use the Order Condition to assess the identifiability of the first three equations of the model.

ii. Explain why there is no identification issue with the last three equations.

iii. Use the Order Condition to assess the identifiability of Consumption Equations and, on the basis of your results, suggest an appropriate estimation technique.

Answer: The coefficient array is,

	C_t	I_t	W_t^P	X_t	P_t	K_t	c	G_t	W_t^G	T_t	A_t	K_{t-1}	P_{t-1}	X_{t-1}
C_t	1	0	$-\alpha_3$	0	$-\alpha_1$	0	$-\alpha_0$	0	$-\alpha_3$	0	0	0	$-\alpha_2$	0
I_t	0	1	0	0	$-\beta_1$	0	$-\beta_0$	0	0	0	0	$-\beta_3$	$-\beta_2$	0
W_t^P	0	0	1	$-\gamma_1$	0	0	$-\gamma_0$	0	0	0	$-\gamma_3$	0	0	$-\gamma_2$
X_t	-1	-1	0	1	0	0	0	-1	0	0	0	0	0	0
P_t	0	0	1	-1	1	0	0	0	0	1	0	0	0	0
K_t	0	-1	0	0	0	1	0	0	0	0	0	-1	0	0

The parameters of a structural equation are identified under the Order Condition if the total number of variables, endogenous or predetermined, excluded from the equation is at least as large as the number of endogenous variables in the system less one.

$$(K - K_j) + (G - G_j) = G - 1$$

Here $K = 8, G = 6$

	K_j	G_j	$K - K_j$	$G - G_j$	$(K - K_j) + (G - G_j)$	$G - 1$
Consumption Equation	3	3	5	3	8	5
Investment Equation	3	2	5	4	9	5
Private Wage Bll	3	2	5	4	9	5

Rank Condition for Cons Eq. The following matrix has to have rank 5.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -\beta_3 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 & -\gamma_3 & 0 & -\gamma_2 \\ -1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

For instance,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 & 0 \end{bmatrix} = - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{bmatrix} = - \begin{bmatrix} 1 & 0 \\ 0 & -\gamma_1 \end{bmatrix} = \gamma_1$$

1. The regressions in the following table are based on a sample of 474 employees of a US bank. The variables are defined as follows: *logsal* - the (natural) logarithm of an individual's salary, l*ogsalbegin* is the (natural) logarithm of an individual's salary when first employed, *educ* is completed years of education, *gender* is a dummy variable which is 1 for males and 0 for females and *minority* is a dummy variable which is 1 for a 'minority' employee and 0 for non-minorities.

   (a) Using a 5% level of significance, carry out a test of the null hypothesis that the coefficients on logsalbegin, educ, gender and minority in Regression A are all zero.

   (b) Calculate the Akaike Information Criterion for Regression A.

   (c) Using a 1% level of significance, carry out a Wald test and an LM test of the hypothesis that the coefficients on gender and minority in Regression A are both zero.

   (d) Use the results in Regressions C, D and E to test the hypothesis that the coefficients in Regression C are the same for minority employees and non-minority employees.

(e) How would your interpretation of the coefficients on gender and minority change if Regression A did not include a constant?
Answer:

Carry out a test of the null hypothesis
that the coefficients on logsalbegin, educ, gender and minority in Regression A are all zero.

ANS	TSS=SSR/(1-Rsq)=	74.67468
	F(4,469)=	481.3216

Calculate the Akaike Information Criterion for Regression A

ANS	AIC=	-293.562	or	-0.61933

Carry out a Wald test and an LM test of the hypothesis
that the coefficients on gender and minority in Regression A are both zero

ANS	F(2,469)=	4.234867

Use the results in Regressions C, D and E to test the hypothesis that the coefficients in Regression C are the same for minority employees and non-minority employees

ANS	F(4,466)=	3.413198	prob val	0.033761