UNIVERSITÄT TÜBINGEN
Wirtschaftswissenschaftliche Fakultät
Abteilung Statistik und Ökonometrie
Professor Dr. Joachim Grammig

24. February 2006

Final Exam (first):INTRODUCTORY ECONOMETRICS (WS 2005/06)

TASK 1

A researcher estimates using EVIEWS the following model proposed by Glosten/Harris:

$$\Delta P_t = \mu + c\Delta Q_t + z_0 Q_t + z_1 Q_t V_t + \varepsilon_t$$

where P_t is the transaction price, $\Delta P_t = P_t - P_{t-1}$, Q_t is the trade indicator, $\Delta Q_t = Q_t - Q_{t-1}$ and V_t is the traded volume. The estimation results produced by EVIEWS have been corrupted in that some information got lost:

Included observations: 2777

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-	0.0002	1.0749	0.2825
ΔQ	0.0048	0.0002	-	-
Q	0.0029	-	9.1181	0.0000
QV	0.0008	0.0001	-	-
R-squared	0.3849	Mean dependent var		0.0001
Adjusted R-squared	0.3842	S.D. dependent var		0.0149
S.E. of regression	0.0117	Akaike info criterion		-6.0569
Sum squared resid	0.3796	Schwarz criterion		-6.0484
Log likelihood	8414.021	F-statistic		578.3119
Durbin-Watson stat	2.4004	Prob(F-statistic)		0.0000

(i) Fill out in the results table the missing data and interpret the economic meaning of the parameters and their statistical significance. **(10)**

(ii) The researcher also estimated the following alternative specification:

Included observations: 2777

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0002	0.0002	0.9113	0.3622
DQ	0.0048	0.0002	20.7421	0.0000
Q	0.0041	0.0003	14.4375	0.0000
R-squared	0.3707	Mean dependent var		0.0001
Adjusted R-squared	0.3702	S.D. dependent var		0.0149
S.E. of regression	0.0118	Akaike info criterion		-6.0348
Sum squared resid	0.3884	Schwarz criterion		-6.0284
Log likelihood	8382.35	F-statistic		816.9383
Durbin-Watson stat	2.3957	Prob(F-statistic)		0.0000

Using various model selection criteria: Which of the two specifications would you prefer? Discuss! **(10)**

(iii) Compute an F-statistic for the null hypothesis that $z_1 = 0$ in the first regression (use the information of both results tables). Describe how you would conduct such an F-test using EViews? **(10)**

TASK 2

(i) The OLS estimation of a linear regression model yields:

$$y_i = \underset{(1.51)}{121.17} + \underset{(0.99)}{0.15}\, x_{1i} - \underset{(-0.97)}{23.91}\, x_{2i} + \underset{(4.35)}{3.98}\, x_{3i}$$
$$n = 236 \quad \bar{R}^2 = 0.8046 \quad F(2, 11) = 27.767 \quad \hat{\sigma} = 39.118$$
(t-statistics in parentheses)

The estimates are quite imprecise (argue why!). Give a suggestion how you would you proceed to deal with the problem of the imprecise estimates? **(10)**

(ii) A researcher suggests to estimate the following model that excludes the second regressor from the equation. This yields:

$$y_i = \underset{(1.84)}{115.13} - \underset{(2.65)}{0.13}\, x_{1i} + \underset{(6.87)}{3.34}\, x_{3i}$$
$$n = 236 \quad \bar{R}^2 = 0.7835 \quad F(2, 12) = 23.467 \quad \hat{\sigma} = 40.187$$
(t-statistics in parentheses)

The results seem improved (why?). Discuss the problems related with dismissing the seemingly insignificant variable x_2 from the model. **(10)**

TASK 3

(i) Show that in a linear regression model the ε_i are serially uncorrelated provided that $g_i = x_i\varepsilon_i$ is a martingale difference sequence and there is a constant in the model. **(5)**

(ii) Show, that strict exogeneity $E(\varepsilon_i|X) = 0$ implies that $E(x_i\varepsilon_i) = 0$. **(5)**

(iii) To test the permanent income hypothesis which states that $C_i^* = kY_i^*$ where C_i^* is permanent consumption and Y_i^* is permanent income you usually encounter measurement errors for the two variables. The variables observed are:

$$C_i = C_i^* + c_i \text{ and } Y_i = Y_i^* + y_i$$

The measurement errors are assumed to have zero mean and are uncorrelated with permanent income and consumption and $E(y_i c_i) = 0$. The model in terms of observable variables reads as:

$$C_i = kY_i + u_i$$

Derive the cross moment $E(Y_i u_i)$. What does your result imply for the consistency of the OLS estimator? **(10)**

(iv) Consider a linear model $y_i = x_i'\beta + \varepsilon_i$. Under which assumptions do we have: $E(\mathbf{b}) = \beta$, where $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Sketch the proof. **(10)**

(v) What are the assumptions one needs to show that $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ is consistent and asymptotically normal? Sketch the proof for asymptotic normality of the OLS estimator. **(10)**

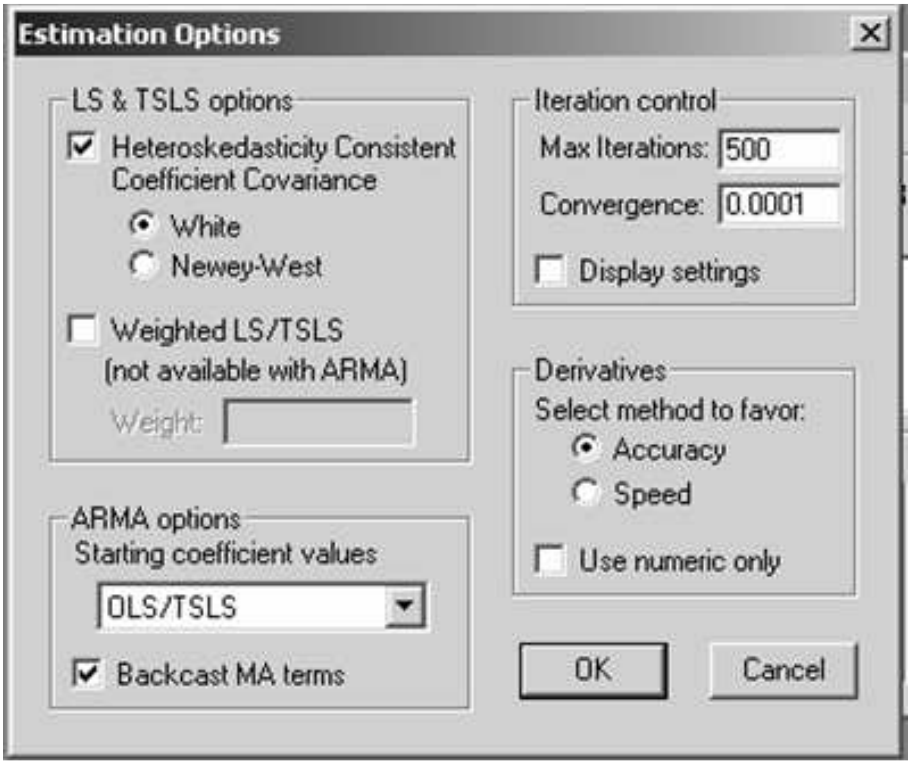
TASK 4

(i) Name three reasons that the assumption $E(x_i\varepsilon_i) = 0$ is violated in a linear regression model and state the consequences for the OLS estimator if this assumption is violated. What do we say about the regressors if $E(x_i\varepsilon_i) = 0$ holds and if $E(x_i\varepsilon_i) = 0$ does not hold? **(10)**

(ii) Which econometric methodology would you propose to estimate the parameters of a linear model $y_i = x_i'\beta + \varepsilon_i$ if the assumption $E(x_i\varepsilon_i) = 0$ is violated? Explain your estimation strategy using a specific example. **(10)**

TASK 5

Consider the following EVIEWS screenshot:



(i) What does the option *heteroskedasticity consistent coefficient covariance* mean? Explain. On which assumptions of linear regression model is the calculation of those standard errors based on? Which assumptions implying the BLUE property of the OLS estimator are NOT present? **(10)**