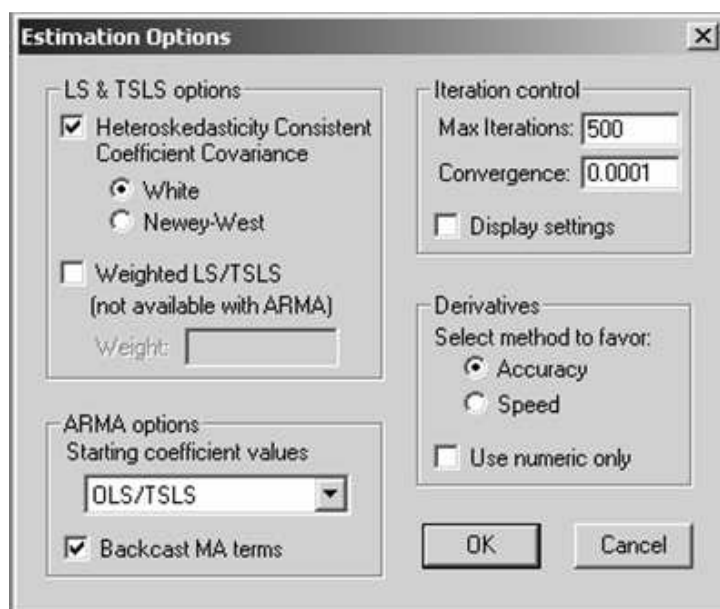


Final Exam (second):INTRODUCTORY ECONOMETRICS (WS 2005/06)

TASK 1

Consider the following EViews screenshot:



- (i) What does the option *heteroskedasticity consistent coefficient covariance* mean? Explain. Which assumptions are required (and which are not required) to compute standard errors of your parameters in this fashion ? (10)
- (ii) To estimate a wage equation

$$\ln(wage) = b_0 + b_1 school + b_2 age + \varepsilon_i$$

where *school* and *age* are the years of schooling and the age in years of individual *i*. A researcher uses EViews. One of the screens (see page 2) looks as follows:

Explain which estimation strategy the researcher is proceeding and what the reasons for choosing this strategy could be. (10)

The estimation produces the following results:

$$\ln(wage) = \underset{(3.51)}{7.24} + \underset{(3.99)}{0.06} school + \underset{(2.97)}{0.03} age$$

$$n = 435 \quad \bar{R}^2 = 0.54$$

(t-statistics in parentheses)

Interpret the estimation results and the significance of the parameter estimates! (10)

- (iv) In a cross sectional linear regression model $y_i = x_i' \beta + \varepsilon_i$ the error variance is not spherical $E(\varepsilon_i^2 | \mathbf{X}) \neq \sigma^2$, but the error variance is a function of x_i ,

$$E(\varepsilon_i^2 | \mathbf{X}) = V(x_i) = \sigma_i^2.$$

The cross correlations of the errors terms are zero. Suppose, σ_i^2 is known. Explain an estimation strategy to estimate the parameter vector β for such a model. (10)

TASK 2

In a paper titled "Blue Laws" Michael Burda and Philippe Weil (2004) investigate the effect of blue laws on unemployment. Blue laws are restrictions on shop-opening hours, most commonly imposed on Sunday trading. They set up a model in which the dependent variable is the of natural log of the rate of unemployment in state i . The explanatory variables also contain a dummy variable *blue* that equals one if the state had a restriction on sunday shop-opening hours and is zero else. The other explanatory variables and including a constant are collected in a vector x_i . Running an OLS regression yields the following results:

Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C (constant)	-	0.0002	1.0749	0.2825
blue (blue law dummy)	0.012	0.002	-	-
other variables (output omitted)				
R-squared	0.56	Mean dependent var		0.06

- (i) Parts of the estimation output got lost while transferring the data. Fill out in the results table the missing data and interpret the economic meaning of the parameter estimates and their statistical significance. Offer a policy implication of the estimation result. **(15)**
- (ii) The parameter vector x_i in Burda and Weil's regression contains three explanatory variables, so together with the blue law dummy and the constant we have a total of $K = 5$ regressors. Compute the bounds of the 0.95 confidence interval for the blue law dummy b **(5)** and provide an interpretation of the result (what does the 0.95 confidence interval tell us?). **(5)**
- (iii) When presenting the results, a discussand criticised that some of the explanatory variables used by Burda and Weil in the vector of explanatory variables x_i are nearly multicollinear. The discussand suggested removing some of the variables from the vector of explanatory variables. What does the discussand mean by the term "near multicollinearity" and what are the problems associated with it? Put yourself in the situation of having to respond to that critique. What would your response be towards the suggestion put forth by the referee? What would a more appropriate remedy against the problem raised by the discussand be? **(10)**
- (iv) As a matter of fact, Burda and Weil did not estimate the parameters via OLS because they argued that an important assumption of the classic linear regression model was violated. Which assumption are Burda and Weil referring to? **(5)** Discuss the consequences of the violation of this assumption also with respect to the robustness of the policy implications of the OLS regression results **(5)**. What would an appropriate alternative estimation strategy be? Describe this estimation strategy and how you would implement it in the concrete example. **(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)** Show that strict exogeneity $E(\varepsilon_i | X) = 0$ implies that $E(x_i \varepsilon_i) = 0$. **(5)**
- (ii) 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 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)**