Introductory Econometrics WS 2006/07
Questions for review

First lecture

1. Ragnar Frisch's statement in the first Econometrica issue basically says that Econometrics is not (only) applied statistics and not (only) mathematics applied to economics. What is it, then?

2. How does the fundamental asset value evolve over time in the Glosten/Harris(1988) model. Which parts of the equation are associated with public and which parts with private information that influences the fundamental asset value.

3. When do we call an asset price "efficient"?

4. Explain the components of the equations that give the buy (bid) and the (ask) price in the Glosten/Harris model.

5. Explain how in the Glosten/Harris the market maker anticipates the impact that a trade event exerts on the fundamental asset price when setting buy and sell prices.

6. Why should it be interesting for a) an investor and b) for a stock exchange (like the New York Stock Exchange) to attach real numbers to the parameters of the Glosten/Harris (1988) model (z_0, z_1, c μ).

7. What does the spread mean? Why is the spread associated with (implicit) transaction costs.

8. Which objects (variables and parameters) in the Glosten Harris model are a) observable to the market maker, but not to the econometrician b) observable to the econometrician.

9. The final equation of the Glosten/Harris model contains observable objects, unknown parameters and an unobservable component. What are these? What is the meaning of the unobservable variable in the final equation?

10. Why did we transform the equations for the market maker's buy and sell price that the market maker posts at point in time t?

11. In the "Mincer equation" from human capital theory (discussed in the lecture) what are the observable variables and what is the meaning of the unobservable component?

12. Why should the government be interested in estimating the parameters of the Mincer equation from human capital theory? Why do we need statistical hypothesis testing in that context?

13. Explain why we can conceive $ln(WAGE_i)$, S_i $TENURE_i$ and $EXPR_i$ in the Mincer equation as random variables.

Second lecture

1. Discuss the key difference of experimental data and the economic data that is typically available.

2. Why is that difference important in the context of ceteris paribus/causal statement like "if x increases, ceteris paribus, then y increases".

3. Take the Mincer equation as an example. Why are the years of schooling "endogenous", i.e. the result of an economic decision process that is not made explicit in the equation. What is the role of the disturbance term ϵ in this context?

4. Provide another economic example (like the crime example or the liberalization example, but one that is not given in the lecture) where endogeneity is present. Cast your example in a CLRM. What is the economic interpretation of ϵ in your model?

5. There is a special name for model parameters which have an economic meaning. What do we call them?

6. Prove your algebraic knowledge and write the CLRM in matrix and vector notation. Indicate the dimensions of scalars, vectors and matrices explicitly!

7. Which objects in the CLRM are observable and which are not?

8. Why do we want that in the scatterplot (= points mapped in a $(x_i'b,y_i)$ diagram) the points cluster closely around the 45 degree line?

9. We are in a CLRM setting, i.e. $y_i = x_i'\beta + \epsilon$. Consider choosing $b = \beta$ in a diagram that takes y_i on the ordinate and $x_i'b$ on the abscissae. Sketch how the diagram would look like when a) $\epsilon \sim N(0,1)$ b) $\epsilon_i \sim N(0,20)$ and c) $\epsilon_i = 0 \quad \forall \quad i$.

10. Imagine you choose as an estimate for β a b that yields a scatterplot where all points lie below the 45 degree line in the $(x_i'b,y_i)$ scatterplot. Why is that troubling in the first place? How would you have to alter your choice of b to change that plot?

11. In the lecture we had a possible choice of $b = (X'X)^{-1}X'y$. Write out that expression in detail using the definition of the matrix X and the vector y. Can you write the right hand side expression using $x_i = x_{i1}, x_{i2}, \ldots, x_{iK}$ and y_i instead of the matrix X and the vector y? You have to use summations $\sum_{i=1}^{n}$ then!