Applied Econometrics Using Stata, SS 2015

February 2, 2016

Jun. Prof. Dr. Lars Hornuf and Cornelia Strüwing

The exam consists of two parts. You have 45 minutes to complete the first part of the exam. This part comprises a pencil-and-paper test and you may not use anything besides a pencil. We will notify you when the first part of the exam ends and collect your answer sheets. Thereafter, we will move to the computer-based exam, for which you will have another 75 minutes. During this part of the exam, you may use your textbooks, course materials and personal notes. Please remember that cheating at the exam is an academic offence and will be punished to the fullest extent. We both wish you good luck!

Part I

1. OLS Estimation (14 Points)

(a) Assume we want to explain the hourly wage of bank managers. The explanatory variables are university education (measured in years) and gender (a dummy variable equal to 1 for males and 0 otherwise). Explain what econometricians understand under the following assumptions: linearity in parameters, random sampling, no perfect collinearity, zero conditional mean, and heteroskedasticity (10 Points).

(b) Does it make sense to include university education (measured in month) or innate ability of the managers as additional explanatory variables in the model? Explain your response. (4 Points).

2. Panel Data and Duration Models (15 Points)

(a) As a policy evaluation tool we often use the difference-in-difference estimator. Explain the idea behind this estimator and write down a simple regression equation to illustrate the estimator. Explain the variables you must include in such a regression in an abstract manner. (7 Points)

(b) What is the difference between a balanced and unbalanced panel? Name two reasons why a panel dataset might be unbalanced. What consequences could unbalancedness have for your analysis? (6 Points)

(c) What do we understand under a proportional hazard, which is assumed under the Cox proportional hazard model? Why might there be no proportionality? (2 Points)

3. Limited Dependent Variables (10 Points)

(a) Explain advantages and disadvantages of the linear probability model. (3 Points)

(b) Logit and Probit are models for binary response variables. How do these models differ from the linear probability model? Explain the difference between the link function in Logit and Probit. (3 Points)

(c) How can we interpret partial effects for discrete and continuous explanatory variables in Logit and Probit models? How does the interpretation differ from simple OLS regressions? (3 Points)