Instructor: Prof. Anna Alberini  
Lectures: T – Th 2:00 – 3:15  
Office hours: Tuesday 9:00 – 10:30 and 3:30 – 5:00.  


Requirements: One midterm, final, plus a 10-page paper reporting on an empirical analysis of your choice (each worth one-third of the final grade).  

Homework: Regular homework sets will be assigned (see below). They will not be graded, but you are strongly encouraged to do the homework to prepare for the midterm, final and prelim exams. Solutions to the homework sets will be provided.  

Suggested Topics:  

1. A review of basic distribution theory. (Univariate distributions and moments; bivariate distributions and moments; conditional and marginal distributions; the bivariate normal distribution; truncated distributions; our first derivation of the linear regression model) [Greene, §3.1, 3.2, 3.4.1, 3.5, 3.6 (skip 3.6.1)]  

Homework: Chapter 3, problems 4, 7, 9, 14, 17, 18, 22.  
Reading: Goodstein article  

2. Statistical Inference. (Sampling frame: random, cluster, stratified and endogenous sampling; basic properties of estimators: unbiasedness, consistency, convergence in distribution, efficiency; interval and point estimation) [Greene, §4.1, 4.2, 4.3, 4.4.1, 4.4.2, 4.4.3]  

Homework: Chapter 4, problems 1, 2, 3  

3. Maximum likelihood estimation. (Derivation of the likelihood function and of the ML estimates; properties of the ML estimates; Wald, score and likelihood ratio tests) [Greene, §4.5]  

Homework: Chapter 4, problems 8, 9, 10, 12, 15, 17
4. The classical linear regression model. (The model, assumptions, OLS, derivation of the OLS estimates, distribution of the OLS estimates, coefficient of determination, prediction) [Greene, §5.1, 5.2, 5.3, 5.4, 5.5]

Homework: Chapter 5, problems 1, 3, 6, 9, 15. Additional problem: Derive the ML estimates of the coefficients of the model \( y_i = \beta_0 + \beta_1 x_i + \epsilon_i \), assuming that the errors are normal with mean 0 and variance \( \sigma^2 \).

5. The classical (multiple) linear regression model using matrix notation. (Notation, derivation of the OLS estimates; partitioned regression; F test of linear restrictions on the parameters; general and specific applications: testing all slope coefficients of a regression model, testing structural change, testing equality of variances; comparison between the F test and the Wald and score tests; testing non-linear restrictions on the parameters; choosing the functional form for the independent variables: the J test; using dummy variables in the right-hand side of the model; [Greene, §6.1, 6.2, 6.3.1, 6.3.3, 6.3.4, 6.5.2, 7.2.2, 7.3, 7.4, 7.6, 7.7, 8.2 up to 8.2.4 included]

Homework: Chapter 6, problems 10, 11, 12
Reading: Grossman and Krueger paper; another paper TBA

6. Large sample properties of the OLS estimates. (Conditions for consistency; asymptotic normality) [Greene, §10.1, 10.2, 10.3]

7. Violation of the assumptions of the classical model I:

A. Omission of variables (Effect of excluding variables on the estimates and on the R square; testing specification with the F test and the Hausman test; model selection criteria)
B. Collinearity
C. Stochastic regressors and their effect on the estimated coefficients. Errors in variables; use of instruments and Hausman test.
D. Non-normal errors; unspecified error distribution. The Gauss Markov theorem.

8. Violation of the assumptions of the classical model II:

A. Heteroskedasticity (effects of ignoring it; correct calculation of the covariance matrix of the estimates; GLS estimation; testing for heteroskedasticity using the White test, the Goldfeld-Quandt test and the Breusch-Pagan test) [Greene, §14.1, 14.2, 14.3]

Homework: Chapter 14, problems 1, 2, 3. Additional problem: Derive the ML estimates of the parameters of the model \( y_i = \beta_0 + \beta_1 x_i + \epsilon_i \), assuming that the errors are normal with mean 0 and heteroskedastic variance \( \sigma_i^2 = \alpha_0 + \alpha_1 z_i \).

Reading: TBA
B. **Serial Correlation.** (some examples of serially correlated error terms: autoregressive errors of the first order, moving average errors of the first order; effects of ignoring serially correlation; correcting for serial correlation; methods for estimating the coefficient of serial correlation: Cochrane-Orcutt, Hildreth-Lu, ML; autoregressive conditional heteroskedasticity). [Greene, §15.1, 15.2, 15.4, 15.5, 15.9]

Homework: Chapter 15, problems 1, 2, 3, 7.
Reading: TBA

9. Systems of equations:

A. **The seemingly unrelated regression (SUR) model.** (basic notation; structure of the error covariance matrix; GLS estimation; conditions under which GLS is the same as OLS equation by equation) [Greene, §17.1, 17.2.12, 17.2.2]

Homework: Chapter 17, problems 1, 2, 3, 4.
Reading: TBA

B. **Systems of simultaneous equations.** (introduction to simultaneous equations; a special case: triangular recursive systems; general notation; obtaining the reduced form; identification conditions; estimation techniques: instrumental variables, two and three-stages least squares)

Readings: