

PSC 603: Applied Data Analysis II

Spring 2020

Lecture: TH 2-3:30pm

Course website: <https://sites.google.com/a/stonybrook.edu/mperess/teaching/lm2020>

Prof. Michael Peress

michael.peress@stonybrook.edu

SBS N743

TA: Alecia Nepaul

Overview: In this course, we will examine the linear regression model and its variants.

The course has two goals: (1) to provide students with the statistical theory of the linear model, and (2) to provide students with skills for analyzing data. The linear model is a natural starting point for understanding regression models in general, inferences based on them, and problems with our inferences due to data issues or to model misspecification.

The model's relative tractability has made it an attractive tool for political scientists, resulting in volumes of research using the methods studied here. Familiarity with the linear model is now essentially required if one wants to be a consumer or producer of modern political science research.

Recommended Textbooks:

William Greene—Econometric Analysis, Any Edition

Kennedy, Peter—A Guide to Econometrics

Angrist, Joshua D., and Jorn-Steffen Pischke—Mostly Harmless Econometrics

J. Johnston and J. DiNardo—Econometric Methods

Jeffrey M. Wooldridge—Introductory Econometrics: A Modern Approach

Fumio Hayashi—Econometrics

Jeffrey M. Wooldridge–Econometric Analysis of Cross Section and Panel Data

The Matrix Cookbook–

(http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf)

Grading: Grading (assuming good attendance) will be based on a midterm (33%), a final (33%), and six (plus or minus one) homework assignments (33%).

Software: In class, we will make use of stata and r for data analysis

Course Outline: I plan on covering the following topics throughout the semester. The list may be contracted or expanded based on how long things are taking.

A. Bivariate Regression

B. Linear Algebra

C. Ordinary Least Squares (OLS)

1. Finite sample properties of OLS
2. Large sample properties of OLS
3. Justifications for the OLS estimator
4. Hypothesis testing with OLS

D. Specification for OLS and Problems with OLS

1. Model fit
2. Omitted and irrelevant variables
3. Prediction
4. Heteroskedasticity and non-normality
5. Outliers and multicollinearity
6. Functional form

E. Endogeneity and Instrumental Variable Estimation

1. Endogeneity
2. Two-stage least squares
3. Measurement error
4. Simultaneous equations
5. LIML and other estimators

6. Weak instruments

F. Time Series Models

1. Lagged dependent variables
2. ARMA errors
3. Newey-West covariance matrix
4. Unit roots
5. Structural change

G. Panel Data Models

1. Fixed effects
2. Random effects and the Hausman test
3. Clustered standard errors
4. Hierarchical models
5. Dynamic panel data models
6. IV estimators for panel data models

H. Nonparametric Techniques

1. Kernel regression
2. Regression discontinuity
3. Nearest neighbor matching