Stat 520 Applied Econometrics I

Instructor: Linda Zhao Izhao@wharton.upenn.edu, 470 Huntsman Hall (JMHH), 215-898-8228 Lectures: 12:00-1:30, Tuesday and Thursday, F94 JMHH Office Hours: 3:00-5:00 Monday or by appointment TA: Kai Zhang zhangk@wharton.upenn.edu, 433 JMHH, Office Hours: 1:30-3:30 Friday Webcafe: https://webcafe.wharton.upenn.edu/stat/

Text book: 1) Linear models with R, Julian Faraway, Chapman & Hall/CRC, 2005

2) Extending the linear model with R, Faraway, Chapman & Hall/CRC, 2006.

References:

- 1) Applied regression analysis and other multivariable methods, Kleinbaum, Kupper, Nizam and Muller, fourth edition, Duxbury, 2008
- 2) *Statistical models theory and practice*, Freedman, Cambridge University Press, 2005
- 3) *Mathematical statistics with applications*, Wackerly, Mendenhall and Scheaffer, sixth edition, Duxbury 2002
- <u>Objects of the course</u>: This is primarily a methodology course. Real data sets will be used to demonstrate the methods and clarify the use of the techniques. At the end of the semester students are expected to be able to analyze a data set (quantitative or qualitative) **wisely** and **conclude correctly**. Regression based techniques will dominate the course, including simple linear model, multivariate regression, discrete choice data. We will however start the course with a quick introduction/review of basic statistical techniques needed such as the central limit theorem, confidence intervals and hypotheses tests.
- <u>Software</u>: R and JMP will be used throughout the lectures. You are encouraged to use R but it is not mandatory for you to use either one of them. You may stay with any software that you already have known.

Logistics: Regular homework assignments and projects, one midterm and a final exam.

Topics:

Preparation for the course

- 1) Probability and distributions, especially Normal, t, F and Chi-squared distributions. The central limit theorem
- 2) Confidence intervals and hypotheses tests

Simple and multiple regressions

- 1) Estimation, Inferences
- 2) Model diagnostics
- 3) Model selection
- 4) One way ANOVA, two way ANOVA and ANCOVA

Discrete choice models

- 1) Binomial data; logit and probit models
- 2) Contingency tables
- 3) Multinomial data

Panel Data and Time Series (time permitted)