This course provides an introduction to advanced statistical analysis techniques within the regression paradigm—such as limited dependent variable models, survival analysis, sample selection equations and panel data models. There are three distinguishing features of this course:

1. Rather than treat each of these topics as isolated tools in a toolbox, we will approach them as extensions to the traditional linear regression model that respond to its statistical and theoretical shortcomings.

2. This will be a hands-on course—each technique will be taught through application to a dataset.

3. This will not be a “proofs” course—while I do expect some familiarity with linear algebra, probability theory and multivariate calculus, the focus will be on understanding the theory behind different estimation techniques, rather than mathematical manipulation of underlying models.

Student weekly assignments (done in groups of 2/3) will cover 50% of the grade; 30% will come from an individual project, and the remaining 20% will be assigned based on class participation. The individual project does not require any data collection or analysis—rather, individual students will select a problem in management research of interest to themselves, and review relevant empirical papers. Their contribution will be to

- evaluate the empirical results in the area
- identify the strengths and weaknesses of the empirical techniques employed
- suggest enhancements in research design or analysis techniques
- If data is available, I strongly recommend that students use this opportunity to analyse it using the different techniques covered in this course

Students will present their work in class in week 5 and 10. There will be 10 sessions in all. The pace will be quite rapid, with weekly assignments and preparation of readings requiring between 5-10 hours per week.

Required Texts:

1. T.A and Campbell, D.T. Quasi-experimentation: design and analysis issues for field settings

2. Wooldridge, J. Introductory econometrics: A modern approach

3. Survival Analysis using the SAS system, P.D. Allison, SAS Institute

Reference:

1. Econometric Analysis, W. Greene
2. Basic Econometrics, D. Gujarati

---

1 Sufficient for you to be able to get some intuition about the underlying mathematical formulation.
Session I: So what’s wrong with cross sectional OLS? (May 1)

- Unobserved heterogeneity & distributional assumptions
- Overview of non-normality, heteroscedasticity, correlated errors, collinearity problems.

**Readings:**

Chapter 2 Cook, T.A and Campbell, D.T. Quasi-experimentation: design and analysis issues for field settings

Skim Chapters 2-5 in Wooldridge to catch up on OLS


Session II: Experimental Design: The Gold Standard (May 9)

- The power of randomised assignment and standardized treatments
- ANOVA/ANCOVA

Guest Faculty: Madan Pillutla

**Readings:**

Chapter 3, 4 & 8, Cook, T.A and Campbell, D.T. Quasi-experimentation: design and analysis issues for field settings


TBA

Session III: Interactions effects in OLS (May 16th)

- Moderating & mediating effects; Estimation and interpretation

**Readings:**

Chapters 1-3 from Aiken, L. S. and West, S.G. Multiple Regression: Testing and interpreting interactions


Session IV: Limited dependent variable models (May 23rd)

- Introduction to ML estimation
- Logit, probit, tobit, poisson and negative binomial regressions

Readings:

Chapter 17 from Wooldridge


Session V: Student Presentations (May 30th)

Mid-project presentations

Session VI: Censored observations (June 6th)

- Censoring
- Parametric and non-parametric models for censored data
- Cox regressions

Readings:

Chapter 2, 5 and 9 from Survival Analysis using the SAS system, P.D. Allison , SAS Institute


Session VII: Adjusting for selection biases (June 13th)

- Instrumental variables and 2-stage least squares.
- Heckman’s correction

Readings:
Chapters 15 and 17 from Wooldridge


**Session VIII Panel data with continuous dependent variables (June 20th)**

- Fixed effects and Random effects
- Specification tests

**Readings:**

Chapters 13 and 14 from Wooldridge


**Session IX: Panel data with Limited dependent variables (June 27th)**

- Conditional fixed effects
- Population averaged models

**Readings:**

Chapter 8 of Logistic Regression using the SAS system, by P.D. Allison, SAS Institute


TBA

**Session X: Student Presentations (July 4th)**

Final project presentations
Note on software packages:

Feel free to work with whichever package you are most comfortable with- the data will be provided in excel format. You might also want to cultivate the skill of porting your data to different packages depending on the specialized strengths of the package. If you are starting out to learn a package- I find Stata to be a complete environment for my work. Also, I can probably provide some hand on guidance with Stata that I cannot for other packages.