

Mgmt 469

Empirical Methods in Strategy

Lies, Damned Lies, and Statistics

Most managers lack real world experience working with data, above and beyond the basics of spreadsheet analysis. As a result, managers either perform inadequate evaluations of the data themselves, or turn the data over to statisticians who may lack the relevant institutional knowledge. Ideally, data-driven evaluation should combine both skills (data analysis and business expertise). This course is intended to help you obtain skills with data analysis. Bearing this in mind, this course has several goals:

- To familiarize you with the process of assembling and analyzing a data set
- To expose you to good research methods
- To reinforce statistical techniques learned in your DS class and also introduce new statistical techniques
- To increase your comfort level working with data
- To introduce Stata, a powerful yet user-friendly empirical research software package
- To enable you to assess research and perform convincing research of your own.
- To help you distinguish convincing research from damnable lies.

Pedagogy

This is a *clinical* course. You may be familiar with the educational process in other clinical fields, such as medicine. A famous dictum in medical education is "see one, do one, teach one." We will adhere to the same philosophy. I will give several lectures during the quarter, but you will spend most of your classroom time (and virtually all of your time outside the classroom) performing and discussing empirical research. You will design and implement three research projects that I have chosen to highlight a range of key research issues. You will also be asked to read and discuss several empirical research papers.

The course is divided into four modules. The first module consists of lectures that build rapidly from basic to advanced research methods. The centerpiece of each remaining module is the analysis of a real world data set. Working in small groups, you will answer specific questions by developing and implementing an empirical model. We will discuss your modeling ideas during in-class "lab sessions." You should also feel free to visit me during my office hours:

Office Hours (Leverone 608) -- TW: 12:15-2pm

I am also available most other times during the week, especially in the morning. Just come up and if my door is open, feel free to poke your head in.

Each project module will end with several group presentations, during which we will discuss methods for presenting empirical results.

Statistical software

Your statistical analyses must be performed in **Stata**. Stata Version 10 has been installed on selected computers in the Kellogg computer lab and at McManus. I will have much more to say about Stata (including providing a Stata primer) in a later class. While Stata is very easy to use when compared with comparably powerful software, you will find yourself spending a disproportionate amount of time during the first two or three weeks mastering the basics of Stata. I hope that the upfront fixed time investments pay off by dramatically reducing the marginal cost of doing good empirical work down the road.

Why learn Stata when you know Excel? Excel is not designed for sophisticated analyses you will want to perform in practice. Moreover, you cannot readily program in Excel; this makes it difficult for you (or others) to reproduce your work. Reproducibility is a hallmark of good research! Most firms that do a lot of empirical research use powerful statistical packages like Stata, and Stata is the package of choice among Kellogg faculty.

Stata offers a graphical user interface (GUI). The GUI is pretty good and we will use it from time to time. But most students end up typing commands the old-fashioned way. I will rely mainly on typed commands in class.

Stata's graphics are very good, but not in the same league as Excel. Feel free to use Excel for graphics. I will show you how to quickly move data back and forth between Stata and Excel.

If you have never used a statistical software package before, you will find it a bit difficult to adjust to Stata at first. You will need to learn a new "language" and "rules of grammar". Rest assured that by mid-quarter, you will be performing sophisticated analyses at lightning speed. In fact, most people who use Stata swear by it. If you decide you would like to purchase it for your own computer, let me know. It is not cheap (\$150-\$200), but it is a worthwhile, long-term investment.

Course Schedule (Note: There is no class on 3/27; there is a makeup class on 4/4)

Topics in *italics* have corresponding readings posted to the course page.

Session 1 (3/30): Introduction to Course; *Ten Steps towards Convincing Research*

Session 2 (4/2): *Using Stata; Regression Basics*

Students should attempt the CAPM tutorial either individually or with a partner. All necessary files are posted to the course page.

Session 3 (4/6): *Regression Basics; Getting Your Data Ready*

The yogurt project will be introduced during this class session. Students should commence work on the “yogurtsmall” data set.

Session 4 (4/9): *Choosing the Right Variables; Fixed Effects Models*

Technical Note: *Performing a Chow (partial-F) test*

Session 5 (4/13): *Maximum Likelihood* and “Lab session”

We will hold an in-class “lab” session discussing the results of your preliminary work on the yogurtsmall data set. Students may commence work on the “yogurtall” data set.

Technical Note: *Influential Observations*

Session 6 (4/16): *Nonlinear Relationships; Programming in Stata*

Session 7 (4/20) Read and discuss the following research papers. A link is provided on the course page:

- Henderson, R. and I. Cockburn (1994) “Measuring Competence? Exploring Firm Effects in Pharmaceutical Research” *Strategic Management Journal* 15: 63-84. Feel free to skip the following parts of the paper: the section entitled “Specification of the Econometric Model”; the discussion of Table 4 on page 77; the Appendix.

Technical Note: *Using ANOVA to Complement Regression*

Session 8 (4/23): Lab session for yogurt project

Technical Note: *Dummy Variables and Interactions*

Session 9 (4/27): Student presentations: Yogurt Project

After the presentations, I will introduce the Hospital Services project. Students may commence work on this project at this time.

Session 10 (4/30): *Regression Noise; Discrete Variables*

Session 11 (5/4): *Discrete Variable (continued)* Accessing U.S. census data.

Session 12 (5/7) Lab session for Hospital Services project

Technical Note: Hypothesis Testing in GLM Models

Session 13 (5/11): Read and discuss two research papers in case packet:

- Kuemmerle, W. (1998) “Optimal Scale for Research and Development in Foreign Environments” *Research Policy* 27: 111-26.
- Gans, J., Hsu, D. and S. Stern, 2007, “The Impact of Uncertain Intellectual Property Rights on the Market for Ideas: Evidence from Patent Grant Delays” *NBER Working Paper* 13234

Session 14 (5/14): Student Presentations: Hospital Services Project

After the presentations, I will introduce the hospital performance and airline no-show projects. (Students may select either project.)

Session 15 (5/18): *Causality*

In addition to this lecture, I will discuss the hospital industry and data in further detail.

Session 16 (5/20): *Benchmarking*

Session 17 (5/21) Lab session for hospital and airline no-show projects

Session 18 (5/28) Final Exam

Session 19 (6/1) Presentation of Final Projects I

Session 20 (6/4) Presentation of Final Projects II

Grading

Your course grade will be based on the projects, class participation, and final exam, broken down as follows:

Yogurt	15%
Hospital Services	15%
Final Project	15%
Final Exam	35% (Note: The final exam is held in class during <i>week 9</i>).
Class Participation	20%

I will assign you to groups for the first project. In general, each group will consist of students of like ability. Otherwise, the most advanced students tend to dominate a group, denying their group mates the maximum opportunity for hands-on learning. However, I will avoid creating groups consisting entirely of “poets”. You may form your own groups for the second and third projects. I strongly encourage you to reform your groups – you will learn more by working with different classmates.

Class participation is vital to the learning experience. Class participation promotes the collaborative nature of empirical research. Be sure to listen to your classmates as they discuss their work. If you can help others become stronger researchers through your questions and comments, then you will receive a strong participation grade. If you say nothing, then you are not helping others and you will receive a low participation score. Also be sure to come to each “Research Paper” class session having more than a passing knowledge of the assigned papers. I will cold call!