University of Pennsylvania
The Wharton School
MGMT 970
Applied Research Methods for Management
Fall 2011
Wednesday 1:30-4:30pm, Bowman Room

Course Description:
Students taking the course will be introduced to the seminal readings on a given method, have a hands-on discussion regarding their application often using a paper and dataset of the faculty member leading the discussion. The goal of the course is to make participants more informed users and reviewers of a wide variety of methodological approaches to Management research.

Course Requirements
The course grade will be based on class participation (33%) and the submission of an empirical paper or draft that uses a method covered in class for ―review‖ by the faculty member teaching that method (67%). The paper need not be written for this class and should ideally be a paper that you are working on for your second year paper requirement, dissertation proposal or another external research project.

Participation
Regular attendance and participation are critical to your successful completion of this course. You should complete the assigned readings and assignments prior to each class. You are encouraged to prepare for class with your colleagues; however, each member of the class should be fully conversant in the material—expect to participate in every class.

Policy on Auditors
Advanced students who do not wish to enrol as full participants are welcome to audit the course, under the following conditions: (1) you commit to attend at least 75% of the class sessions, and inform me in advance which sessions you will attend; (2) you complete the assigned readings and assignments for the classes that you attend; and (3) you participate fully in the sessions that you attend, including doing a ―fair share‖ of class discussion. Auditing students are excused from submitting a class paper and should not expect a ―review‖ from the instructors of the class.

Reading Materials:
There is no bulkpack for this class. PDF versions of the readings will be put on webCafé or distributed in hard copy in the week prior to class. Data for assignments will also be distributed via WebCafe.

Prerequisites:

MGMT 953, Research Methods or a similar course covering the Philosophy and Design of Social Science Research.

Oxley, J., J. Rivkin, M. Ryall and the Strategy Research Initiative, (2010) ―Recognizing and Encouraging High Quality Research in Strategy‖ pp. 10-14

Students unfamiliar with STATA should review the materials in
https://faculty-cafe.wharton.upenn.edu/eRoom/depts/DoctoralMaterials/0_980da and, if needed,
http://wps.aw.com/aw_stock_ie_2/50/13016/3332253.cw/index.html

9/7: Introduction and a Refresher on OLS Regressions (Iwan Barankay)

Reading

Review your material from your Statistics class on OLS regression

Assignment

There is a STATA data file, Teachingratings.dta and a pdf file, Teachingbeauty.pdf, describing the contents of the data.

Otherwise you can always type "help 'keyword'" to get help on a keyword, e.g. "help summarize" or "help browse"

1) What is the mean and standard deviation of each variable in the data
2) Create a new variable where all the observations have value one.
3) Run a regression of *female* on the variable you created under 2). Interpret the regression result.
4) Run a regression of *course_eval* on *beauty*. Test for the significance of *beauty*. Interpret the magnitude of the coefficient.
5) Run a regression of *cours_eval* on *beauty* and *female*. Did the coefficient on *beauty* change? Explain clearly why.
6) Using the regression output in 5) interpret in two different ways the magnitude of the estimated coefficient.
7) How is the variable *onecredit* coded. Run a regression of *course_eval* on *beauty, female and onecredit.* Interpret the estimated coefficient on *onecredit.*
8) Which of the other regressors do you think should also be included in the model. Do they affect the estimated coefficient on *beauty?*
9) Are there any variables you wish you could have in the model that are not in the data set? Explain clearly why.

9/14: Discrete Choice (Matthew Bidwell)

Readings

1) Pampel, Fred C. 2000.―Logistic Regression – A Primer‖. Sage: Quantitative Applications in the Social Sciences 07-132. Pages 1-54 [TO BE DISTRIBUTED IN HARD COPY IN CLASS on 9/8]

2) Hoetker, G. 2007. The use of logit and probit models in strategic management research: critical issues. *Strategic Management Journal* **28** 331-343.

3) Bidwell, M. 2010. Why Has Job Mobility Increased? Unions, Organizational Size and the Growth of External Hiring

Assignment

1) Replicate Tables 2 and 3 (don't worry if you can't get them exactly)

2) Provide three different ways to evaluate the magnitude of important effects from Table 2. What are the strengths and weaknesses of the different approaches?

3) We will discuss in detail some of the decisions that I made in setting up the analyses. What do you think about the specific approach I have taken to addressing the issues I am interested in? What other approaches to analyzing the data might I have taken, given the constraints of the data available to me?

9/21 Count (David Hsu)

Reading (Focal)

Furman, JL & S. Stern. 2011. "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," American Economics Review, 101(1933-1963).

Supplementary Readings

Cameron, A. Colin & Pravin Trivedi. 1986. Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests. Journal of Applied Econometrics, 1: 29-53.

Cameron, A. Colin & Pravin Trivedi. 1998. Regression Analysis of Count Data. New York, NY: Cambridge University Press.

Hausman, Jerry, Bronwyn Hall, & Zvi Griliches. 1984. Econometric Models for Count Data with an Application to Patents-R&D Relationship. Econometrica, 52(909-938).

F. Murray and S. Stern. 2007. "Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis," Journal of Economic Behavior and Organization, 63: 648-687.

Simcoe, Tim Stata module to Estimate Fixed-Effects Poisson (Quasi-ML) Regression with Robust Standard Errors; download at: http://en.scientificcommons.org/20874411. Or, in Stata, type "ssc install xtpqml"

Woolridge, Jeffrey M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, pp. 645-683.

Assignment
To prepare for our session, I would like you to read the Furman and Stern (AER, 2011) paper, which I have uploaded to the class webcafe. It is oriented toward science policy, but gives a nice illustration of applied count data analysis. You can find the dataset and Stata code used to generate their results at: http://www.aeaweb.org/issue.php?doi=10.1257/aer.101.5
by scrolling down to their article and clicking on "download data set". Also, if you do not have prior exposure to econometric count data models, please read the Woolridge (2002) chapter and/or the Cameron and Trivedi (1986) article (both are available in electronic form in the class webcafe). For further background, I have also uploaded the Hausman, Hall & Griliches (1984) paper, which derives panel data estimators for count data models. The Cameron and Trivedi (1998) book is a standard reference in this domain (consult the library or buy the book if you are interested). Finally, the reference to Stata code implements robust standard errors for fixed effects Poisson models, a feature which does not come standard in Stata (Stata has standard commands for fixed and random effects Negative Binomial

panel models). The Murray and Stern (JEBO, 2007) paper is optional, and is another application of count data models.

My plan for the session is to mainly cover the application of count data models within the context of research study design. To that end, we will discuss and debate the Furman and Stern paper from the standpoint of choices they made in their study design. In the second part of class, I will motivate a separate research question that is centered more on business policy. We will then form several teams and work in parallel in brainstorming research study designs. We will then reconvene and discuss and debate the merits of each of the study designs. Please come prepared to actively participate.

9/28 Panel data (in linear, discrete choice or count, fixed vs. random effects, clustering, autocorrelation including spatial, pcse, GMM, …) (Iwan Barankay)

Readings
Chapter 5 from Angrist and Pischke ―Mostly Harmless Econometrics‖

Assignment
?

10/5 Dealing w/ endogeneity: Selection, instruments, propensity score matching (Evan Rawley)

Readings

1. Campa, Jose Manuel and Simi Kedia (2002) 'Explaining the Diversification Discount' *Journal of Finance* 57(4):1731-62.
2. David, Guy, Dan Polsky and Evan Rawley (2010) 'Integration and Task Allocation: Evidence from Patient Care' *Working Paper*

Assignment

This class will focus on two methods commonly used to "deal with endogeneity": matching on observables and instrumental variables.

The papers are attached in webCafé. Campa and Kedia is an excellent example of the power of instrumental variables (IV)—it attacks a classic result in the diversification literature that did not "deal with endogeneity" very well. We will use this paper to discuss the strengths and weaknesses of IV.

My paper with Guy David and Dan Polsky uses a new matching technique, called Coarsened Exact Matching (CEM), that is superior to propensity score matching in some ways and is easy to apply. Besides discussing CEM we will also use the paper to discuss propensity score matching and discuss statistical techniques for evaluating whether the results of matching are "good."

I‘ve listed some optional readings below for students who are interested in more background material on dealing with endogeneity.

Matching

Levine, David I., and Michael W. Toffel. "Quality Management and Job Quality: How the ISO 9001 Standard for Quality Management Systems Affects Employees and Employers." *Management Science* 56, no. 6 (June 2010): 978-996.

Stefano M. Iacus, Gary King, Giuseppe Porro. Causal Inference Without Balance Checking: Coarsened Exact Matching. October 12, 2009. Mimeo.

Heckman correction

J. Myles Shaver. 1998. Accounting for Endogeneity When Assessing Strategy Performance: Does Entry Mode Choice Affect FDI Survival? *Management Science*, Vol. 44, No. 4., pp. 571-585.

Villalonga, Belen. Does Diversification Cause the "Diversification Discount"? *Financial Management*, Vol. 33, No. 2. (Summer, 2004), pp. 5-27.

Instrumental variables

Nevo, Aviv. 2000. ―Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry,‖ *The RAND Journal of Economics*, 31(3), 395-421, 2000. Reprinted in P. Joskow and M. Waterson ed., *Emprical Industrial Organization*, Edward Elgar, 2004.

Novak, Sharon and Scott Stern. 2008. How Does Outsourcing Affect Performance Dynamics? Evidence from the Automobile Industry. *Management Science* Vol. 54, No. 12, December 2008, pp. 1963-1979

10/12 Survival/Failure/Event History & event studies (Witold Henisz)

Readings

1) Kiefer, Nicholas M. (1988) ―Duration Data and Hazard Functions‖ Journal of Economic Literature 26(2): 646-679.

2) Box-Steffensmeier, Janet M. (1997) ―Event History Models in Political Science‖ American Journal of Political Science 41(4): 1414-1461

3) Allison, Paul D. (2010) ―Survival Analysis‖ Pp. 413-425 in The Reviewer‘s Guide to Quantitative Methods in the Social Sciences, edited by Gregory R. Hancock and Ralph O. Mueller. New York: Routledge.

4) Henisz, W. J. & Delios, A. (2001). ―Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996.‖ Administrative Science Quarterly, 46(3): 443-75.
5) Jensen, M. (2006) ―Should We Stay or Should We Go? Accountability, Status Anxiety, and Client Defections.‖ Administrative Science Quarterly, 51(1):97-128.

Data Assignment

Using the dataset examining leadership mortality (of a sample of national political leaders) available in the eRoom answer the following questions. Use STATA help and manuals to help you as needed.

NB: Definition of LOST
0 - still in power in 1987
1 - Constitutional loss of power
2 - Natural Death
3 - Other

1) What is the origin time for each leader (i.e., the time at which a leader begins to be at risk for being deposed)? Explain.

2) How many leaders
   a. survived one year?
   b. are right censored?

3) Construct a life table for the data.

4) Plot the survivor function

5) Plot a hazard function

6) Do your plots lead you to favor a parametric or partial likelihood approach to modeling this data? Why?

7) What do you think the appropriate functional form to measure the likelihood of losing power over time is? Explain.

8) What determines the likelihood of losing power? Does it depend on the type of loss (i.e., natural death, constitutional or non-constitutional transfer)?

Discussion Questions

1) Conceptually discuss the variables that the authors used to stset the data in the Henisz & Delios and Jensen papers. To answer this questions you need to have a clear sense of the data structure and the types of data needed to determine the origin, length and termination of a spell for data with time varying independent variables.

2) Conceptually discuss the variables that you would use to stset the data in a paper of interest to you that is amenable to event history/survival analysis.

3) What functional form and other modeling choices would you make for this dataset? Why?

10/19 Experiments (Lab & Field) (Adam Grant)

Readings

1) Aronson, E., Wilson, T. D., & Brewer, M. (1998). Experimentation in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), The handbook of social psychology, Volume 2 (4th Ed.) (pp. 99-142). New York: McGraw-Hill.

2) Cooper, W. H., & Richardson, A. J. 1986. Unfair comparisons. Journal of Applied Psychology, 71: 179-184.

3) Latham, G. P., Erez, M., & Locke, E. A. 1988. Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez-Latham dispute regarding participation in goal setting. Journal of Applied Psychology, 73: 753-772.

4) Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. Psychological Bulletin, 112, 160-164.

Assignment

1) List at least two causal hypotheses that you are interested in testing your own research.

2) For one hypothesis, write out the design of a laboratory experiment to test it, including the task, how you would manipulate the independent variable, and how you would measure the dependent variables.

3) For the other hypothesis, write out the design of a field experiment or quasi-experiment to test it, including the organizational setting, how you would create or track changes in the independent variables, and how you would measure the dependent variables.

10/26 Factor Analysis & Structural Equation Modeling (Nancy Rothbard)

Readings and Reflections

1) Edwards, J. R. & Bagozzi, R. P. 2000. On the Nature and Direction of Relationships Between Constructs and Measures. Psychological Methods, 5(2): 155-174.

- o The Edwards and Bagozzi (2000) article addresses underlying factor analysis issues from a theoretical perspective.
- o As you read it, think about your own data sets and whether you have reflective or formative indicators of a latent construct and what the theoretical and methodological implications of that are.

2) Chapter 8 ―Hypothesis Testing‖ in Kline, R. (2010). Principles and Practice of Structural Equation Modeling, Third Edition.

- o Read this chapter to become familiar with the various model statistics used to interpret structural equation models. We will go over these in the session and talk about questions you might have. And what caveats there might be to the advice he gives regarding model evaluation.

3) Rothbard, N. (2001). Enriching or Depleting? The dynamics of engaging in work and family roles. Administrative Science Quarterly, 46: 655-684.

o This is one of my empirical papers that uses structural equation modeling. Take a look at the front end to see what the hypotheses are, but you only need to carefully read the methods and results sections.
o Look at the model statistics that are reported.
o Look at Table 2 which gives you the information you would need to reconstruct the confirmatory factor model.
o The methods and results section talk about using instrumental variables to identify a non-recursive (i.e. reciprocal) model. Bring questions you have about model identification.

Data Assignment

As you prepare for the session, if you have a data set that you are working on, please bring some data for us to work with (have it in electronic form so we can cut and paste it into some Lisrel syntax during the session. Please limit the number of constructs to 5 and bring a correlation matrix at the item level and means and standard deviations of each item.

- For example, if you have a 3 item scale of job satisfaction, a 3 item scale of organizational commitment, a 4 item scale of intrinsic motivation, and a single item measure of performance, you would need a correlation matrix of the 11 items that represent these 4 constructs.
- If you have a 3 item job satisfaction scale a 3 item organizational commitment scale, a single item performance evaluation measure, a single item indicator of gender and a single item indicator of age, then your correlation matrix would have 9 items that would represent 5 constructs.

Think about a hypothesis you have about how these constructs will relate to one another. For example, Intrinsic motivation will lead to greater (a) job satisfaction, (b) organizational commitment and (c) performance. Or Intrinsic motivation will lead to greater job satisfaction, which will in turn lead to greater organizational commitment.

11/2 Hierarchical Linear Modeling (Jennifer Mueller)

Readings

To prepare for this class you only need to read the following two articles posted in Web Café:

a. Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. Journal of Management. Special Issue: Focus on hierarchical linear modeling, 23(6), 723-744.

b. Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics, 23(4), 323-355.

The other readings are NOT required but I'll use them to supplement our class discussion so I've made them available for you.

Assignment

The goal of this session is to introduce you to multi-level modeling (MLM) for purposes of publishing research in the social sciences. Because we will be covering a broader range of topics, we will not be able to delve too deeply into any one single topic. However, you should complete the assignment below so that you can spend more time focusing on the application(s) of MLM which you may find most valuable. Please note, I'll demonstrate MLM using SAS PROC Mixed for this course and not HLM (HLM is a

statistical software that performs multi-level modeling) – so no need to buy HLM (although you can download a student version of HLM from the website for free). You will not need to use or bring your laptops to class. If you would like to bring specific examples to class to build and test real time, please email me (jennm@wharton.upenn.edu) and I'll try to incorporate this into the class (assuming we are covering material relevant to your question).

Come to class prepared with your own data (or research idea) which you think might require MLM. Write out at least one hypothesis you wish to test. To test your hypothesis, identify which of the six MLM applications (from the list above) you would employ and identify why. If you are not sure which application would be best to use (or if your question requires yet another application of MLM not listed above), give your best guess about the application this question would require.

- What is between-level variance?
- What is within-level variance?
- Which of the applications of MLM above requires that you first meet the criterion of showing sufficient between-level variance in your outcome?
- What statistic would you use to show that you have sufficient between-level variance in your outcome?

Class content:
1. Controlling for group level variance (or covariance) when analyzing level 1 data
2. Intercepts as outcomes
3. Slopes as outcomes
4. Multi-level mediation
5. Multi-level moderated mediation
6. Panel data – longitudinal data analysis

11/9 Networks (Lori Rosenkopf, Katherine Klein and Mathis Schulte)

There are three components to your preparation for our session.

1. Three readings in the assigned readings folder:
   - Kilduff and Brass (current review of both micro and macro network research)
   - Snijders et al. (conceptual overview of SIENA and a simple micro-level example of its use)
   - Gulati et al. (macro-level application examining network structure)

   There are also several other articles for your discretionary use available in the reference readings folder:
   - SIENA manual
   - UCINET manual
   - Schulte, Cohen, Klein (companion piece to data analysis below)
   - Rosenkopf and Padula (companion piece to data analysis below)
   - Van de Bunt and Groenewegen (macro-level SIENA application)

2. Choose an article using social network analysis that is published in a top-tier management or discipline (e.g., sociology, psychology, economics) journal. This could be an article you've read before, or a new article you've just discovered. To insure that everyone takes a different article, please post your choice as a comment to these directions below (in webCafé) and place a copy of the article in the reference folder. By Tuesday night, 11/8, at 9 pm, please write and email us a very brief description (in four sentences or less) of your article's main point or research question, making clear why social networks matter given the author(s)' research focus. In short, why did the author(s) take a social network approach versus some other approach? Please also note in bullet point format:

- How the network boundary was determined
- What types of ties constitute the network
- The unit of analysis and the network measures at this level
- Whether the network variables are predictors or outcomes
- The type of regression/analysis and the software employed

3. Data analysis: This folder contains some network data (data1109.zip) and directions on how to analyze it (UCINET Tutorial.docx).

11/16 Simulations (Nicolaj Siggelkow)

Readings

Nicolaj Siggelkow and Jan W. Rivkin. 2005. "Speed and Search: Designing Organizations for Turbulence and Complexity." Organization Science, 16, pp. 101-122.

Nicolaj Siggelkow and Jan W. Rivkin. 2009. "Hiding the Evidence of Valid Theories: How Coupled Search Processes Obscure Performance Differences among Organizations." Administrative Science Quarterly 54, pp. 602-634.

Dirk Martignoni and Nicolaj Siggelkow. 2010. "When it Pays to be Neurotic or to Have Blind Spots: The Value of Understanding External and Internal Contingencies."

Corp-Dev-Spec : These are the specs that Vikas, Harbir and I sent to a programmer. The "Daedalus" program refers to the program that this programmer wrote for Jan and me.

(optional): If you are interested what became of the specs, here's the paper that resulted: Vikas Aggarwal, Nicolaj Siggelkow, and Harbir Singh. forthcoming. "Corporate Development Choices and Interdependence: Strategic Tradeoffs and Performance Implications." Strategic Management Journal.

Discussion questions

1. What roles can simulation models play? To which purposes are simulation models being used in the first three papers?

2. When are simulation models convincing, when are they not?

3. Dirk and I just received an R&R for our paper from SMJ. What do you think did the reviewers complain about? What would you have complained about?

4. Pick a question that you are interested in. Start sketching out a simulation model. Be as concrete as possible. Ideally, you could hand over your outline to a programmer. What tradeoffs are you facing? What aspect of this exercise did you find most difficult?

11/30 Content Coding (Sigal Barsade)

Readings

Content Coding
Please read the **theory and methods sections only** :

- Barsade, Sigal G. (2002). "The Ripple Effect: Emotional Contagion and its Influence on Group Behavior." Administrative Science Quarterly, 47, 644-675.

Please read the **methods section only** :

- Amabile, Teresa M., Barsade, Sigal G., Mueller, Jennifer S.& Staw, Barry M. (2005). "Affect and Creativity at Work." Administrative Science Quarterly, 50, 367- 403.

Inter-rater Agreement

- Inter-rater agreement - 2 page document by Kristin Smith-Crowe
- Smith-Crowe, K., Burke, M. J., Kouchaki, M., & Signal, S. (2010). Assessing interrater agreement given theoretical and methodological problems in applied psychology and management. *Working Paper*.
- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69,* 85-98.
- Burke, M. J., & Dunlap, W. P. (2002). Estimating interrater agreement with the average deviation (AD) index: A user's guide. *Organizational Research Methods, 5,* 159-172. (**scan only**)

Supplemental

- Meyers, R.A. & Seibold, D. R. Forthcoming. Coding Group Interaction. In *Methods for Studying Small Groups: Interdisciplinary Perspectives* (Eds: Andrea B. Hollingshead & M. Scott Poole).

12/7 Comparative Methods (Mauro Guillen)

Readings

Mark Blaug, ―Kuhn versus Lakatos, or paradigms versus research programmes in the history of economics.‖ History of Political Economy 7(4) (Winter 1975):399-433.

Colin McGinn, ―Looking for the Black Swan.‖ The New York Review of Books 49(18) (21 November 2002). [Review of books on Karl Popper]

Anuja Gupta and Mauro F. Guillén, ―Developing, Testing, and Validating Management Theory with Comparative Case Studies.‖ Working Paper (2010).

John Gerring, Case Study Research (New York: Cambridge University Press, 2007), pp. 65-150.
Ragin, Charles C. 2000. Fuzzy-Set Social Science. Chicago: University of Chicago Press, excerpts.

Assignment

TBD