University of Pennsylvania
The Wharton School
MGMT 970
Applied Research Methods for Management
Fall 2016
Fridays 9am-12pm SH-DH 2034

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 weekly empirical assignments (67%) which will be graded on a check/check plus basis.

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.

Weekly Assignments
Instructions for each assignment are in the syllabus. Assignments should be submitted by email to the professor teaching the class by 8am on the morning of each 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; (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.

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.

WEEK 1: Friday 2nd September
Course Intro and a Refresher on OLS Regressions (Iwan Barankay)

Please complete this exercise and submit the answers via email to Barankay@wharton.upenn.edu
Deadline is Friday, September 2 at 8am.

Review your material on OLS Regressions from Statistics

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

If you are new to STATA, review one or more online tutorials available at http://www.stata.com/links/resources-for-learning-stata/

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

*ASSIGNMENT (**note: this assignment is due on Friday 9th September**)*
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.

WEEK 2: Friday 9th September (Iwan Barankay barankay@wharton.upenn.edu)
Panel data (in linear, discrete choice or count, fixed vs. random effects, clustering, autocorrelation including spatial, pcse, GMM, …)

Readings

There are slides on the faculty-hub site to refresh your knowledge on Regression analysis. You need to review this before class.

We will cover chapter 5 from Angrist and Pischke "Mostly Harmless Econometrics" You can access a *.pdf of the book at
http://www.development.wne.uw.edu.pl/uploads/Main/recrut_econometrics.pdf

Assignment

WEEK 3: Friday 16th September (Matthew Bidwell mbidwell@wharton.upenn.edu**)**

Choice Models

Readings

- Pampel (2000): Logistic Regression, a Primer. This is a Sage book on quantitative methods. You can either buy it from Amazon, read it on reserve at the Annenberg library, or borrow my copy (I'll lend it on a first come, first serve basis).
- Hoetker, Glenn. 2007. "The Use of Logit and Probit Models in Strategic Management Research: Critical Issues." *Strategic Management Journal* 28:331–43.
- Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)." *Statistical Science* 199–231.

Assignment
I've uploaded a dataset to the facultycafe page as well, and a word document that briefly summarizes the data. There is also an associated paper ("Shifts and Ladders") on the website where we describe the data in more detail. I would like you to do the following, preferably in STATA or R:

1. Create a dummy variable that takes the value 1 if the job was left for personal reasons or to seek a new job elsewhere (jobexitpersonal and jobexitnewjob), is missing if the job was left involuntarily or to start a new company (jobexitletgo, jobexittempover, jobexitnewventure), and is 0 otherwise (these should all be internal moves).
2. Use three different statistical techniques to estimate the determinants of whether a job is left by internal mobility or voluntary external mobility (a binary choice)
3. Estimate how the **log of the number of subordinates** affects the **probability** that the job will be left by internal mobility rather than voluntary external mobility.
4. Provide a means of assessing which of the three different statistical techniques provides the best fit for the data.

WEEK 4: Friday 23rd September (Minyuan Zhao myzhao@wharton.upenn.edu)

Count Models

Readings

Hausman, J. A., B. H. Hall, Z. Griliches. 1984. Econometric models for count data with an application to the patents–R & D relationship. Econometrica 52: 909–938.

Cameron, A. C., P. Trivedi. 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. Journal of Applied Econometrics, 1: 29-53.

Greene, W. 2013 Cross Sections, Panel Data, and Microeconometrics. Pages 810-829.

Optional (Applications)

Agrawal A., C, Rosell, and T, Simcoe. 2014. How do tax credits affect R&D expenditures by small firms? Evidence from Canada

Murray, F. 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 Organizatio*n, 63: 648-687.

Zhao, M. Conducting R&D in countries with weak intellectual property rights protection. 2006. *Management Science*, 56 (7): 1185–99.

Assignments

I have uploaded a small dataset on webcafe, adapted from McCullagh, P. and Nelder, J. (1989) Generalized Linear Models, 2nd Edition. Chapman and Hall, London.

1) Create dummy variables for the ship type, construction period, and operation period.
2) Calculate the descriptive statistics for the data and decide which model is appropriate for the estimation.
3) Use Poisson, Negative Binomial and their variations (zero-inflated, truncated, hurdle model, etc.) to fit the data.
4) Assess the quality of fit. Do the results in 3) support your decision in 2)?

WEEK 5: Friday 30th September (Witold Henisz henisz@wharton.upenn.edu)

Survival/Failure/Event History

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.

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.

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 assuming a Weibull and exponential distribution.
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?

WEEK 6: 14th October (Exequiel Hernandez exequiel@wharton.upenn.edu)

Network Methods

Readings

Zaheer, Akbar, Remzi Gozübüyük, and Hana Milanov. 2010. "It's the Connections: The Network Perspective in Interorganizational Research." *Academy of Management Perspectives* 24(1):62–77.

Vasudeva, Gurneeta, Akbar Zaheer, and Exequiel Hernandez. "The embeddedness of networks: institutions, structural holes, and innovativeness in the fuel cell industry." *Organization Science* 24.3 (2013): 645-663.

Assignment

You will work with the "Fuel Cell 1992.xls" and "Fuel Cell 1992 for UCINET.xls" files. [NOTE: If you happen to have your own data and would like to use it, be my guest!]. These files represent a single year of the dataset used for the paper by Vasudeva, Zaheer, and Hernandez.

These files contain data on fuel cell R&D alliances between organizations for the year 1992. Each row represents a unique alliance dyad (e.g. row 3 is an alliance between Alcatel and General Electric). Columns D and E simply list the numerical ID for each firm in the dyad (e.g. Alcatel's ID is 2, GE's is 12). In network language, each row represents a tie or an "edge" between firms. Each firm is a "node" in the network. If you want to read a good, short tutorial on network nomenclature you can read "Fundamentals of Social Network Analysis" by David Knoke. This is not required since many of the basics are fairly intuitive.

You will work with these data in UCINET, which is the most commonly used software for network analysis by management scholars. You can a free trial copy from https://sites.google.com/site/ucinetsoftware/downloads. It only works on Windows.

Please complete the following tasks:

1. Input the data into UCINET (Data > Import Excel > DL-type formats). Use the "Fuel Cell 1992 for UCINET.xls" for this (the other file is just for you to know the names of the firms involved).

From the options on the right menu, select "Edgelist1" from the drop down menu as the data format. Leave the other options as default, except that you will save your file once by clicking on "force symmetry" and once by leaving "force symmetry" unchecked. I'll call the former the symmetric matrix and the latter the asymmetric matrix.

Now display both network matrices (Data > Display). What's the difference between the two? Given what you know about alliances, which of the two makes the most sense to use for analysis?

(Optional: It really helps to visualize network data. I'd encourage you to play around with one of the files in NetDraw by going to Visualize > NetDraw)

2. Calculate the following network variables:

- Degree centrality (Network > Centrality and Power > Degree). Do it for both the symmetric and asymmetric files. What does the output mean? What kinds of research questions could a researcher answer by using degree centrality as a variable? Can you calculate degree centrality—unstandardized and standardized—manually for node #8 (Deutsche Aerospace)?

[To help with this, read through the "Centrality and Prestige" file created by David Knoke. It's very short.]

- Constraint (Network > Ego Networks > Structural Holes). Do this only for the symmetric matrix. What does the output mean? What kinds of research questions could a researcher answer by using constraint as a variable? Can you calculate constraint manually for node #8 (Deutsche Aerospace)?

[To help with this, read through the "Calculating Access to SH – Burt (2008)." This is the appendix to one of Ron Burt's books on structural holes. You only need to read the first 3.5 pages. Read the rest if you're interested in interpreting all the output you get from UCINET, but it's not required.]

- BONUS: Calculate Bonacich's Power Centrality measure using the symmetric matrix (Network > Centrality and Power > Bonacich Power), which is typically used to capture status (e.g. Podolny, 2001 in AJS). What does the output mean? This measure is complex so don't try and calculate it manually.

If you're finding this so fascinating that you want more, play around with other network measures. UCINET's help menu provides a description of each variable along with a reference to the original journal article from which it was taken. If you want even more:

David Knoke's website for "Social Network Analysis Theories & Methods" (http://www.soc.umn.edu/~knoke/pages/SOC8412.htm). David has made his materials available for free.

Hahneman & Riddle's "Introduction to Social Network Methods", a free online book (http://faculty.ucr.edu/~hanneman/nettext/) with UCINET instructions.

Borgatti, Everrett, & Johnson's "Analyzing Social Networks" book (not free), which also has UCINET instructions.

WEEK 7: 21st October (Anoop Menon armenon@wharton.upenn.edu)

Computational Text Methods

Readings:

Read: Grimmer, J. and Stewart, B. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Analysis, pp 1-31. http://stanford.edu/~jgrimmer/tad2.pdf

Assignment:

1. Think of a research question that you are interested in which requires, or can benefit significantly from, large-scale textual analysis. Keep this question in mind as you approach the following:

a. Which method(s) appears most relevant for your question?

b. What concrete next steps would you need to take on the text processing front?

2. For your question, think of some sample text documents.

a. Acquire 3 such documents.

b. Paste them (separately) into Stanford Core NLP, at http://nlp.stanford.edu:8080/corenlp/process

c. Does the output seem sensible? What seems to work, and what does not?

d. Explore.

3. Imagine reviewing a paper in your domain of interest that uses computational text processing techniques.

a. What would be some of the context specific concerns you might have regarding the application of computational text processing techniques to that domain?

b. What robustness checks or validations would you ask for?

WEEK 8: 28th October (Ann Harrison annh@wharton.upenn.edu)
Dealing w/ endogeneity: Selection, instruments, propensity score matching (Ann Harrison)

Readings

You need to read the following papers to get a sense of the varied approaches to dealing with potential endogeneity bias:

William Easterly and Ross Levine, "Tropics, Germs and Crops: How Endowments Influence Economic Development", Journal of Monetary Economics, 2003 50(1), 3-40.
Ann Harrison and Jason Scorse, "The Nike Effect: Anti-Sweatshop Activists and Labor Market Outcomes in Indonesia", American Economic Review (on my website).

Esther Duflo and Rohini Pande, "Dams", The Quarterly Journal of Economics (2007), 122(2): 601-646.

Corinne A. Moss-Racusina,b, John F. Dovidiob, Victoria L. Brescollc, Mark J. Grahama,d, and Jo Handelsmana, "Science faculty's subtle gender biases favor male students", Proceedings of the National Academy of Sciences, September 2012.

Assignment

1. You have been given the dataset used by Easterly and Levine (EL). Reproduce the Tables used by Easterly and Levine in their article. In class, I will be using STATA to show my work.
2. What happens to the coefficient of interest when EL switch from OLS to IV. Why?
3. How good of an instrument do EL have? How do you know?
4. Think of your favorite research question or theory. In the specification that you have in mind to test that question (or theory) what are the likely sources of endogeneity bias? Think of an approach (instrument, randomized trials, other approaches) that could address this problem.

WEEK 9: 4th November (Emilie Feldman feldmane@wharton.upenn.edu)
Propensity Score Matching and Coarsened Exact Matching

Readings

Belen Villalonga (2004), Does Diversification Cause the Diversification Discount"? *Financial Management*. 33(2) 5-27

Matthew Blackwell, Stefano Iacus, Gary King and Giuseppe Porro (2009) cem: Coarsened Exact Matching in Stata. *The Stata Journal* 9(4): 524-526

Assignment

Please be sure to install the propensity score matching and coarsened exact matching files in STATA using the following commands:

- ssc install psmatch2, replace
- ssc install cem, replace

Also, please download the Stata file, NSW.dta, from the course webpage.

Background: The National Supported Work (NSW) Experiment was a randomized experiment carried out in the mid-1970's to examine whether a job training program improved the earnings of low-skilled workers. Since it was a real experiment, there was a control group and a treatment group whose treatments were randomly assigned, which means that the experiment could be used to consistently identify the causal effect of receiving job training on future earnings. Thus, the treatment reflected whether or not workers received on-the-job training, and the outcome was the 1978 real income earned by the workers. The results of the NSW Experiment showed that the mean difference between the treatment and control groups (unadjusted for any other variables) was $1,794 in 1978 real income (the standard error was $633). Using a regression-adjusted model, the mean difference between the treatment and control groups was $1,672 in 1978 real income (the standard error was $638). Thus, on-the-job training caused the workers who received it to enjoy higher future earnings.

An economist named Robert Lalonde took the data on the treated cases from the experiment (the workers who participated in the job training program), threw out the real experimental controls (the workers who did not receive on-the-job training), and constructed a new set of controls from observational data (from the Current Population Survey and the Panel Study on Income Dynamics). He then ran a variety of regression models to estimate the effect of job training on future earnings, and he compared his results to those of the randomized experiment. What he found was troubling: the regression results scarcely resembled the true experimental findings—they often had the wrong sign and were two to three times the size of the real effect.

The dataset you have, NSW.dta, contains the original treated cases from the NSW Experiment (the workers who participated in the job training program) and the control cases that Lalonde incorporated (a large number of people from the CPS and PSID, a nationally-representative survey of income and employment). Please answer the following questions, and email me your responses to them by 8am on Friday, November 4:

Ordinary Least Squares

1. What is the estimated treatment effect of the NSW program on 1978 real income using a basic OLS regression with no control variables? With control variables? Why are these so different than the "true values" of $1,794 ($1,672) that came out of the NSW Experiment?
2. Construct summary statistics for the treated and control group. How do these help explain the results you found in question 1?

Propensity Score Matching

3. What is the estimated treatment effect of the NSW program on 1978 real income using a propensity score matching model with nearest-neighbor matching and a logit regression to compute the propensity scores? How does it compare to the OLS results you found above, and to the "true values" of the experimental results?
4. What is the estimated treatment effect of the NSW program on 1978 real income using a propensity score matching model with caliper matching (use a caliper of 0.01) and a logit regression to compute the propensity scores? How do these results compare to the OLS results you found above, to the "true values" of the experimental results? Why do the nearest neighbor matching results differ from the caliper matching results?
5. For both the nearest-neighbor and caliper matching results, how does the balance look overall? Are there any variables in particular that look unbalanced even after matching? What, if anything, should you do with the t-test column? Why? What should you do when balance is not achieved?
6. For both the nearest-neighbor and caliper matching results, what is the estimated treatment effect of the NSW program on 1978 real income using a basic OLS regression on the matched sample you identified in your propensity score matching model? How do these results compare to the OLS results you found above, and to the true values of the NSW Experiment? Why do the nearest neighbor matching results differ from the caliper matching results?

Coarsened Exact Matching

7. What is the estimated treatment effect of the NSW program on 1978 real income using a coarsened exact matching model with automated coarsening? How does it compare to the OLS results you found above, and to the "true values" of the experimental results?
8. Re-run your coarsened exact matching model using manually-coarsened values of the educ variable into the following buckets: Grade School = 0-6, Middle School = 7-8, High School = 9-12, College = 13-16, Graduate School = 16+. What is the estimated treatment effect of the NSW program on 1978 real income? How does it compare to the OLS results you found above, and to the "true values" of the experimental results?
9. Compare the CEM models you ran in response to Questions 7 and 8 in terms of: (a) the overall imbalance of the models; (b) the number of matched observations; and (c) the estimated treatment effects. What explains these differences?

WEEK 10: 11th November (Luis Rios luisrios@wharton.upenn.edu)

More Endogeneity: Difference in Differences and Natural Experiments

Readings

Chapter 6 from Angrist and Pischke "Mostly Harmless Econometrics" You can access a *.pdf of the book at http://www.development.wne.uw.edu.pl/uploads/Main/recrut_econometrics.pdf

Cook, Thomas D. ""Waiting for life to arrive": a history of the regression-discontinuity design in psychology, statistics and economics." *Journal of Econometrics* 142.2 (2008): 636-654.

Angrist, Joshua and Victor Lavy (1999) "Using Maimonides' Rule To Estimate The Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* v.114 pp.533-575. (SKIM this to get a sense of how the empirics are set –up)

Marx, Matt, Deborah Strumsky, and Lee Fleming. "Mobility, skills, and the Michigan non-compete experiment." *Management Science* 55.6 (2009): 875-889. (SKIM this to get a sense of how the empirics are set –up)

Imbens, Guido W. "An economist's perspective on Shadish (2010) and West and Thoemmes (2010)." (2010): 47. (SKIM this to develop intuition on the larger conversation about these methods).

Assignments

D-in-D:

Examine the impact of Hurricane Katrina on labor market outcomes of evacuees using Current Population Survey (CPS). The March 2005 CPS serves as a "pretest" detailing labor market outcomes for the 2004 calendar year, and the March 2006 CPS is the "posttest." The CPS is designed to interview people 8 times over a 16 month period. You will receive data in advance.

1. Describe the following outcomes for evacuees in the March 2006 sample: weeks worked last year, earnings, unemployment compensation, and self-reported health status. Report means for continuous variables and the distribution across categories for ordinal ones. This description is the equivalent of a quasi-experiment with no control group and no pretest. Can you conclude anything meaningful from this basic analysis?

2. Perform a simple before/after comparison, testing the hypothesis of equality for the same four variables, comparing the "pre-evacuation" outcomes of individuals destined to become evacuees (in the March 2005 sample) with the post-evacuation

outcomes of evacuees (in the March 2006 sample). Estimate OLS or ordered-Probit regressions (your choice) that control for age, race, sex, and educational attainment.
Do controls make a big difference for the estimates?

3. Now perform a simple treatment/control comparison, using non-evacuees as the control group, using only the March 2006 sample. In other words, perform a quasi-experiment with a control group but no pretest, for the same four variables.

4. Use the data to perform a difference-in-difference style estimate, using both pre- and post-evacuation outcomes and observations both on Katrina evacuees and others. Estimate OLS regression models (ordered Probit is OK for health status) with and without controls.

RD:

For this assignment, you'll use regression discontinuity techniques to analyze whether the failure to receive a bonus motivated NC teachers to improve their performance in the subsequent school year. You will receive dataset in advance.

1. Use the Stata *rd* command to find Wald estimates of the local average treatment effect of receiving a bonus (or alternatively, being denied a bonus) on school performance in the subsequent year (avgrow06). ***To install this command, type "ssc install rd" into the command line in your Stata window.*** Report bootstrapped standard errors for your Wald estimates. Do you find any evidence of a statistically significant effect? Compare the standard errors obtained using different bandwidths in the underlying local linear regression. Which bandwidth choice—default (denoted lwald), twice the default (lwald200), or half the default (lwald50)—yields the most precise estimates of the Wald statistic?
2. Use the Stata rd command once again to produce the following graphs:
a. A plot of the discontinuity in treatment status when avgrow05 switches from negative to positive. Verify that there is a significant discontinuity present in this graph.
b. A plot of the outcome variable against avgrow05, showing the discontinuity (if any) at the bonus threshold. BONUS: Replicate this plot using the three default bandwidth choices. Verify that the local linear regression is smoother when a wider bandwidth is employed.
c. Plots of the following four covariates: enroll, pblack, phisp, and pfrl. An assumption underlying regression discontinuity analysis is that relevant pretreatment covariates do not vary discontinuously at the treatment threshold. Are your graphs consistent with this assumption? Use an informal "eyeball" test.
4. Perform an RD analysis parametrically. Regress avgrow06 against a cubic in avgrow05, the pretreatment characteristics pblack, phisp, pfrl, enroll, elemschl, midschl, highschl, and the bonus indicator. Can you reject the hypothesis that schools receiving the bonus perform equally as well as schools that fail to receive the bonus in the subsequent year? If so, which schools are doing better in the following year -- those that receive the bonus or those that do not?

WEEK 11: 18th November (Nancy Rothbard nrothbard@wharton.upenn.edu)
Factor Analysis & Structural Equation Modeling

Readings

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.

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.

WEEK 12: 2nd December (Drew Carton carton@wharton.upenn.edu**)**

Multilevel Modeling

Readings:

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. (2000). The application of hierarchical linear modeling to organizational research. (Note: Read 489 – 497 first, then read from the beginning of the chapter to 489.)

Carton, A. M., & Rosette, A. S. (2011). Explaining bias against black leaders: Integrating theory on information processing and goal-based stereotyping. Academy of Management Journal, 54(6), 1141-1158.

Assignment:

I will send you data and instructions for SPSS in which you should conduct all tests suggested by Hofmann et al. (2000) – from those needed to establish the need for HLM to a test of HLM itself. Namely, conduct the following tests:

1. In this dataset: Which are the Level 1 and which are the Level 2 variables?
2. Justifying HLM
    a. One way ANOVA. Describe the amount of variation between versus within groups. Is HLM justified?
    b. Beyond the one way ANOVA, are there other reasons, based on the hypotheses and the measures, that we may need to conduct a mixed model?
3. Random-coefficient regression
4. Intercepts-as-Outcomes Model
5. Slopes-as-Outcomes Model

WEEK 13: 9th December (Samir Nurmohamed nurmo@wharton.upenn.edu)
Experimental Methods

Readings

Colquitt, Jason A. "From the editors publishing laboratory research in AMJ: A question of when, not if." Academy of Management Journal 51.4 (2008): 616-620.

Hulbert, Lorne. "Experimental research designs." Breakwell, Glynis M. (Ed), (2004). Doing social psychology research. , (pp. 10-53). Leicester, England: British Psychological Society; Malden: Blackwell Publishing, viii, 379 pp.

Nurmohamed, S. Article on ethical leadership and reviewer comments (to be distributed).

Assignment

The purpose of this session is to introduce students to designing and conducting laboratory experiments to answer research questions related to topics in the organizational sciences. To prepare for the session, students will read articles on how to design effective laboratory experiments (Hulbert), and the specific challenges, as well as opportunities, that lab experiments present in organizational research (Colquitt). Using these readings, students will be asked to develop a lab study to address reviewer comments to an article on ethical leadership (to be distributed) and present their designs to the class. We will end the session by discussing some of the questions related to the research ethics of lab experiments.