

RESEARCH METHODS IN ORGANIZATIONAL SCIENCE II

SECTION 600

FALL, 2000

Class: Monday, 8:30-11:00 a.m. Room 427 Wehner

Instructor: Javier Gimeno

Office: 433S Wehner (845-3881)

Office Hours: By appointment

Course Objectives

1. To provide students with an overview of research methodology applicable to studying non-experimental phenomena within and between organizations. Thus, 686 is mainly a "macro" research seminar.
2. To introduce skills and concepts necessary to conduct independent research.
3. To develop ability to critically read and evaluate the research of others and to understand the peer review process.
4. To provide exposure to archival and field study approaches used by macro researchers.
5. To demonstrate research methodology skills and knowledge acquisition by:
 - a) Building and maintaining a dataset during the semester
 - b) Using the dataset to analyze several problems involving conversion of hypotheses into statistical tests, running these tests, and writing the results.
 - c) Gain first-hand practice in using some important statistical methodologies.
 - d) Completing an individual research project proposed in Research Methods in Organization Science I.

Course Overview

Each class meeting will have an assigned set of readings on the topics listed below. Students will be responsible for participating in all class discussions. In addition to readings, the assignment will suggest some questions to guide your thinking and preparation. Those questions will also be used for class discussion. Although you do not need to turn in your answer for those questions, please be ready to discuss them.

In contrast with other seminars, we will not have students summarize the papers covered. However, each day I will assign a student to prepare a summary of the class discussion, highlighting the main points in the readings and the discussion in a way that will be helpful as a reminder/study aid to the rest of the class. These summary write-ups will be handed out to the instructor and the rest of the class one week after the meeting date.

Each student will be expected to complete 1 problem assignment each week. These will be assigned the week prior to the due date, and will typically involve data manipulation and/or analysis.

Management 686

The major paper for the course will present an independent research project carried out over the past 3 semesters (Spring, Summer, Fall). This should be the outcome of research proposed in Research Methods in Organization Science I. If for some reason you need to deviate from your proposal in MGMT 687, please consult with me. Preferably, the paper should be written and formatted in the style of the Academy of Management Journal.

The final exam will cover the material in the course. According to the new prelim rules, you will not be tested on research methods on the prelim. It is my responsibility to test your individual knowledge of the subject matter with a comprehensive prelim-style exam.

Books Used

Several books will be used throughout the semester. The following is a list of books that will be required, and/or are recommended.

Required books:

- Davis, James A. 1985. *The Logic of Causal Order*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-055. Newbury Park, CA: Sage Publications.
- Jaccard, James, Robert Turrisi and Choi K. Wan. 1990. *Interaction Effects in Multiple Regression*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-072. Newbury Park, CA: Sage Publications.
- Sayrs, Lois W. 1989. *Pooled Time Series Analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-070. Newbury Park, CA: Sage Publications.
- McDowall, David, Richard McCleary, Errol E. Meidinger, and Richard A. Hay, Jr. 1980. *Interrupted Time Series Analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-021. Newbury Park, CA: Sage Publications.
- Finkel, Steven E. 1995. *Causal Analysis with Panel Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-105. Newbury Park, CA: Sage Publications.
- Bryk, Anthony, and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Advanced Quantitative Techniques in the Social Sciences series, volume 1. Newbury Park, CA: Sage Publications.
- Long, J. Scott. 1983. *Confirmatory Factor Analysis: A Preface to LISREL*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-033. Newbury Park, CA: Sage Publications.
- Long, J. Scott. 1983. *Covariance Structure Models: An Introduction to LISREL*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-034. Newbury Park, CA: Sage Publications.
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences series, volume 7.

Newbury Park, CA: Sage Publications.

Breen, Richard. 1996. *Regression Models: Censored, Sample Selected, or Truncated Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-111. Newbury Park, CA: Sage Publications.

Yamaguchi, Kazuo. 1991. *Event History Analysis*. Applied Social Research Methods series, volume 28. Newbury Park, CA: Sage Publications.

Scott, John. 1991. *Social Network Analysis: A Handbook*. Newbury Park, CA: Sage Publications.

Highly Recommended General Texts:

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation*. Boston: Houghton Mifflin Co.

Greene, William H. *Econometric Analysis*. Prentice Hall. (latest edition).

Highly Recommended Advanced Texts in Specific Methodologies:

Baltagi, Badi H. 1995. *Econometric Analysis of Panel Data*. Chichester, UK: Wiley.

Hsio, Cheng. 1986. *Analysis of Panel Data*. Econometric Society Monographs, No. 11. Cambridge, UK: Cambridge University Press.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs, No. 3. Cambridge, UK: Cambridge University Press.

Blossfeld, Hans-Peter, and Götz Rohwer. 1995. *Techniques of Event History Modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis*. Structural Analysis in the Social Sciences series, volume 8. Cambridge, UK: Cambridge University Press.

Software Used

As part of this course, you will need to complete some assignments regarding data manipulation and statistical analysis. I do not have any preferences about the software that you use for the analysis.

Most quantitative research involves two steps: data manipulation and statistical analysis. For data manipulation, you need to be familiar enough with one piece of software that will allow you flexible manipulation of data. My suggestion for this are SAS and Stata. I believe that SAS is the best software for data manipulation, as well as for connecting to external sources (Compustat, CRSP, etc). I will provide SAS assistance, and the problem assignments will often talk about SAS code. Stata is also very effective data manipulation

Management 686

software (in some areas, better than SAS). However, it is perhaps not as general and all-inclusive as SAS.

For the statistical analysis, you are free to use whatever software you want, and in fact I expect that you will become proficient in moving your database around to alternative software packages to benefit from the software specialization in particular methodologies. For instance, Stata is particularly good statistical software for several of the methodologies that we will use, and I will encourage you to transfer your data to Stata for analysis. However, I will not expect you to change to a new package as long as you can do the assignment in your preferred package.

Evaluation

Class participation & summaries	25%
Research problems	25%
Final Exam	25%
Research paper	25%

Tentative Schedule

(please note that the following dates may change to accommodate travel schedule. I have made an effort to already re-schedule those days when I won't be able to be here.)

M., Aug. 28	Introduction; Design and Specification Issues (I): causal order, control variables, critical tests.
M., Sept. 4	Design and Specification Issues (II): mediators, moderators, interaction effects, dummy variable models.
<i>F., Sept. 15</i>	Longitudinal Analysis of Continuous Dependent Variables: Panel Data Analysis Techniques.
M., Sept. 18	Longitudinal Analysis of Continuous Dependent Variables: Introduction to Dynamic Modeling.
M., Sept. 25	Issues of Multiple Levels of Analysis: Introduction to HLM.
M., Oct. 2	Measurement Issues. Exploratory and Confirmatory Factor Analysis.
M., Oct. 9	Introduction to Structural Equation Modeling.
M., Oct. 16	Models for Binary, Count and Limited Dependent Variables: Logit, Probit, Multinomial Logit, Tobit, Poisson, Negative Binomial.
M., Oct. 23	Models for dealing with Selection Bias: Survival Bias; Sample Selection Bias; Self-Selection Models.
<i>F., Oct. 27</i>	Modeling Discrete Change over Time: Introduction to Event History Modeling (I)
M., Oct. 30	Event History Modeling (II)
M., Nov. 6	Event History Modeling (III); Financial Event Studies
M., Nov. 13	Network Analysis (I)
M., Nov. 20	Network Analysis (II); contagion/diffusion models.
M., Dec. 4	Papers due (5 p.m.)
F., Dec. 8	Final Exam (10-noon)

INTRODUCTION

CLASS 1

PREPARATION QUESTIONS:

- ❖ Please prepare a 5-10 minutes presentation of the research project that you begun in MGMT 687, with special attention to your hypotheses, research design and specification of your model. Please also bring a copy of your paper so that I can be familiar with your project.
- ❖ Prepare the following questions for class discussion:
 - The regression model includes an error term. What is that error term supposed to represent? Why do we assume that the error term is distributed as a normal distribution?
 - What are “control variables”? Why are control variables important? What’s wrong with having too many control variables? What’s wrong with having too few control variables?
 - Every time you submit a paper, there will always be a reviewer that will make a fuss about a variable that is missing in your specification even though it is, according to the reviewer, well known to affect your dependent variable. The reviewer wants you to add the variable, but, honestly, it is a pain to collect new variables after the fact. Under what conditions could you make a methodologically credible argument that the lack of the control variable does not pose a threat to the statistical conclusion validity of your analysis?
 - Stinchcombe describes the concept of the “crucial experiment”. What is it? Could you provide an example of a “crucial experiment” that you have seen in the literature?
 - (Please think about both sides of this questions.) Suppose that your theory predicts a positive (alternatively, negative) relationship between X and Y. Yet, you find a negative (alt., positive) correlation. You suspect that the difference between theory and correlation is that you are not controlling for other factors that may influence the DV. You have a list of potential control variables, with information about their correlation with X and Y. How would you identify what missing control variables are likely to explain the unexpected results?
 - How do issues of specification relate to the threats to validity? (You may need to review the papers and notes from MGMT 687 about threats to validity).
- ❖ Application to your research project:
 - Think about your research project, and about the main relationship that you expect to find. Come up with at least 2 different alternative explanations for that relationship that do not involve your suggested mechanism. How could you control for those alternative theories? How could you build a critical experiment? You may want to mention this during your presentation.
 - Davis depicts theoretical models graphically in systems of boxes and arrows. In your presentation, use a path diagram to represent your model.

READINGS:

Davis, Murray S. 1971. That's Interesting! *Philosophy of Social Science*, 1: 309-344.

Locke, Karen, & Karen Golden-Biddle. 1997. Constructing opportunities for contribution: Structuring intertextual coherence and 'problematizing' in organizational studies. *Academy of Management Journal*, 40(5): 1023-1062.

Davis, James A. 1985. *The Logic of Causal Order*. Newbury Park, CA: Sage Publications.

Greene, William H. 1997. "Specification Analysis." Pp. 399-404 in *Econometric Analysis*. Upper Sadle River, NJ: Prentice Hall.

Hitt, Michael A., Javier Gimeno, and Robert E. Hoskisson. 1998. "Current and Future Research Methods in Strategic Management." *Organizational Research Methods* 1:6-44.

McWilliams, Abigail, and Donald Siegel. 2000. "Corporate Responsibility and Financial Performance: Correlation or Misspecification?" *Strategic Management Journal*. 21(5): 603-609.

Stinchcombe, Arthur L. 1968. "The Logic of Scientific Inference." Pp. 15-56 in *Constructing Social Theories*. Chicago, IL: University of Chicago Press.

MEDIATORS, MODERATORS, AND INTERACTION EFFECTS

CLASS 2

PREPARATION QUESTIONS:

- ❖ Using the ideas discussed by Murray Davis (1971) and Locke & Golden-Biddle (1997) [which you read last week], write an outline of the introduction of your paper. How are you making your paper interesting? How are you "problematizing" your contribution?
- ❖ Mediating relationships are often difficult to test, since they involve multiple dependent variables. Can you clearly state the steps that one should go through to test whether the relationship between X and Y is fully mediated by Z? How about partially mediated?
- ❖ Be ready to discuss the similarities and differences of these concepts:
 - Contingency theory
 - Fit
 - Moderator relationship
 - Interaction effect
- ❖ Very often, we assume that a contingency theory should imply a moderator relationship. Is that assumption granted? Are there other ways in which a contingency theory can be tested?
- ❖ Very often, we assume that a moderator relationship should be tested with an interaction effect. Is that assumption granted? Are there other ways in which a moderator relationship can be tested?
- ❖ Should you include the main effects when you test for interaction effects? Why? Why not? What are the pros and cons of including the main effects?
- ❖ In a model with interaction effects, how should we interpret the main effects? (tricky question, watch out!).
- ❖ I recently run a regression model with an interaction effect. If I left the variables in their original state, only the interaction effect was significant. On the other hand, if I centered the variables before taking the interaction, the main effects were different, and all three effects (two mains and one interaction) were significant. What results should I report? How should I interpret the results that I get? How should I interpret my weird change of significance?
- ❖ Suppose that you have a moderator relationship between a binary variable (0,1) and a continuous variable. How would you go about specifying the relationship in your model? Would you center? Both variables? Only one? None?
- ❖ Another interesting thing in moderator relationships relates to graphing the results. This helps in the interpretation of the paper (see previous point). There are several methods to do this. Cannella and Rowe use a very popular method. Do you know of other ways to display moderator relationships? Please search a couple of journals in your area over the last year for alternative ways to present moderator relationships and interaction effects. If you decide to graph your interactions effects, does it matter at what points you graph the relationship. (hint: range).

READINGS:

Jaccard, James, Robert Turrisi and Choi K. Wan. 1990. *Interaction Effects in Multiple Regression*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-072. Newbury Park, CA: Sage Publications.

Aiken, Leona S., & Stephen G. West. 1991. 1991. Testing and Probing Three-Way Interactions. In *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications. [excellent book!]

Drazin, Robert, and Andrew H. Van de Ven. 1985. Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30: 514-539.

Baron, R., and Kenny, D. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

James, Larry, and Jeanne Brett. 1984. Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, 69, 307-321.

Applications readings (read the theory quickly, focus on the method applications):

Cannella, Albert A. Jr. and W. Glenn Rowe. 1995. Leader capabilities, succession, and competitive context: A study of professional baseball teams. *Leadership Quarterly*, 6: 69-88.

Rowe, W. Glenn, Albert A. Cannella Jr., and Ira Harris. 1998. In search of meaning: Is the Fortune reputation survey a source of new information to investors? Working paper, Texas A&M University.

PANEL DATA ANALYSIS

CLASS 3

PREPARATION QUESTIONS:

- ❖ What is panel data? Find an example in your literature of someone using panel data in their research. Bring a copy of the paper to class.
- ❖ Why would anybody want to collect panel data as opposed to cross-sectional or time-series data? What are the advantages of panel data? What are the problems?
- ❖ How important are panel data techniques for experimental researchers? And for non-experimental researchers?
- ❖ What are the problems of analyzing panel data using the traditional OLS approach – i.e., pooling all observations and running an OLS regression?
- ❖ There are mainly two popular approaches for analyzing panel data (a) the Kmenta approach – also known as the Parks approach, and technically, the "cross-sectionally heteroskedastic and timewise autoregressive model" [what a mouthful] –, and (b) the panel data models – fixed effects or random effects. Can you state the assumptions in each model in mathematical form? What do those assumptions mean in real, practical terms? In your opinion, under what conditions would one be better than the other?
- ❖ What are the differences between the following panel estimators: (a) the standard pooled OLS estimator, (b) the within estimator, and (c) the between estimator? Why are they called that way? How would you calculate these estimators in your package of choice? (don't actually do it, but determine what steps you would need to undertake).
- ❖ What are the differences between fixed effects and random effect estimators? What are the pros and cons of each? How would you know if you need to use one or the other? While knowing the formulas is a plus, try to understand the insight behind the formulas.
- ❖ Traditional panel data models only tend to assume individual differences in the intercepts (individual-specific effects). Why not differences in the slopes of the independent variables? What would be the limitations for developing models where individuals – or time periods – have different slopes? How could those models be estimated? [we will look at this in the the multi-level modeling class].
- ❖ In groups of two students, please explore how you could estimate the following models in the following statistical packages. You will then report your findings in class. Models: fixed effects (LSDV), random effects, Kmenta model, Hausman test. Packages: SPSS, SAS, Stata, LIMDEP.

READINGS:

- Sayrs, Louis W. 1989. *Pooled Time Series Analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-070. Newbury Park, CA: Sage Publications.
- Greene, William H. 1997. *Econometric Analysis*. Third edition. Chapter 14. Upper Saddle River: Prentice Hall.
- Kmenta, Jan. 1986. *Elements of Econometrics*. Second edition. Sections 12-1 and 12-2. New York: MacMillan.
- Baltagi, Badi H. 1995. *Econometric Analysis of Panel Data*. Chapter 1. Chichester, UK: Wiley.
- Beck, N. and Katz, J.N. 1995. "What to do (and not to do) with time-series cross-section data". *American Political Science Review*, 89: 634-647.
- Bowen, Harry P. and Margarethe F. Wiersema. 1999. "Matching method to paradigm in strategy research: Limitations of Cross-sectional analysis and some methodological alternatives", *Strategic Management Journal*, 20: 625-636.

Catch up readings on heteroscedasticity and autocorrelation:

I would like everybody that is not well versed on the (1) diagnosis and (2) correction for heteroscedasticity and autocorrelation to read on that topic. If your knowledge of statistical notation and math is good enough, go directly to Greene. If not, read Gujarati, and then skim through Greene looking for diagnosis tools that were not covered in Gujarati. Particularly, look at the Breusch-Pagan test and the test for groupwise heteroscedasticity in Greene. These are important tests, and are often used in the context of panel data analysis.

Chapters 11 and 12 of Gujarati, Damodar. *Essentials of Econometrics*.

or

Chapters 12 and 13 of Greene, William H. *Econometric Analysis*.

DYNAMIC MODELING

CLASS 4

PREPARATION QUESTIONS:

- ❖ What are the advantages of explicitly dynamic modeling (i.e., when lagged variables are used) relative to simply longitudinal research?
- ❖ Dynamic analysis is good for testing direction of causality. Suppose that you have X and Y that are correlated, but you do not know which causes which. Apart of course from theory, would dynamic modeling help you figure out the order of causality? What would you do?
- ❖ What are the pros and cons of the different possible approaches to measuring and testing change (first differences, residual gains, lagged DVs).
- ❖ What is “regression to the mean”? Why is it a problem? What are the implications of regression to the mean processes for dynamic modeling?
- ❖ Make sure that you understand the process of identification of an ARIMA(p,d, q) process.
- ❖ Suppose that you want to test the effect of a one-time change of a qualitative change (shock) on a continuous DV over time. How would you approach the modeling of the change effect? (Hint: look at section 3 of McDowall et al).
- ❖ Suppose that you want to test the effect of a change in a quantitative (continuous) variable on a continuous DV. How would you approach it? (Hint: think about. We’ll work through an application in class).

READINGS:

McDowall, David, Richard McCleary, Errol E. Meidinger, and Richard A. Hay, Jr. 1980. *Interrupted Time Series Analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-021. Newbury Park, CA: Sage Publications.

(pages 1-44) Finkjel, Steven E. 1995. *Causal Analysis with Panel Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-105. Newbury Park, CA: Sage Publications.

(pages 44-69) Menard, Scott. 1991. *Longitudinal Research*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-076. Newbury Park, CA: Sage Publications.

For those of you that want to do their assignments in SAS, it will be helpful to read:

“Chapter 2: Working with Time Series Data”, from *SAS/ETS User’s Guide*.

MULTIPLE LEVELS OF ANALYSIS

CLASS 5

PREPARATION QUESTIONS:

- ❖ Either (a) find a paper in your area that deals with a multiple levels of analysis issues, or (b) think about a research question in your area that deals with multilevel issues. What are the important levels that you tend to find in your literature? What is the nature of cross-level influences?
- ❖ What are the similarities between multilevel analysis and panel data analysis? What are the differences? Could you write a fixed-effects panel model as a multi-level model? And a random-effects one?
- ❖ What are the differences between the following concepts? What role do they play in hierarchical linear modeling?
 - Fixed effects ANOVA vs. random effects ANOVA.
 - Varying coefficients vs. random coefficients.
- ❖ Write down a two-level hierarchical structure. Then, write that model in the form of a single equation (by substituting the coefficients by their equations). What is the structure of the multilevel model? Why couldn't you estimate it with OLS? What is the nature of the error term?
- ❖ Why is centering such an important topic in multilevel research? What are the consequences of grand-mean centering vs. group-mean centering?
- ❖ How can you evaluate if your multilevel model is correct? How would you evaluate if there is significant parameter variance across groups?
- ❖ Determine what software you would like to use for your forthcoming multilevel assignment (SAS MIXED or HLM). You may browse at the sources discussed below to gauge the capabilities of the programs. Please send me an e-mail at least 24 hours before the class with your preference.

READINGS (I SUGGEST THAT YOU READ THEM IN THIS ORDER)

- Klein, Katherine J., Fred Dansereau, and Rosalie J. Hall. 1994. Levels issues in theory development, data collection, and analysis. *Academy of Management Review*, 19: 195-229.
- Hoffmann, David A. 1997. An overview of the logic and rationale of hierarchical linear models. *Journal of Management*, 23: 783-804.
- (Chapters 1, 2, 4, 5, 8 and 9 are required) Bryk, Anthony S., and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models*. Advanced Quantitative Techniques in the Social Sciences, 1. Newbury Park, CA: Sage Publications.
- Kreft, Ita, and Jan de Leeuw. 1998. 'Frequently Asked Questions', Chapter 5 in *Introducing Multilevel Modeling*, pp. 105-146. Thousand Oaks, CA: Sage Publications.
- Hoffman, David A., Mark A. Griffin, and Mark B. Gavin. (in press). "The application of Hierarchical Linear Modeling to Organizational Research", in K.J. Klein and W.W.J. Kozlowski (Eds.) *Multilevel theory, research, and methods in organizations*. Society for Industrial and Organizational Psychology Frontiers book series, Jessey-Bass Inc.

The second part of your reading assignment (which I suggest that you delay until you actually run your hierarchical linear model) depends on your choice of software. There are two software solutions that I recommend. Depending on which one you choose, you will need to read on that software so that you are familiar with it. The software solutions are SAS PROC MIXED and HLM 4.04 Student Version (which is available free at <http://www.ssicentral.com/other/hlmstu.htm>). You may use other HLM version or software if you can get a copy.

If you decide to use SAS PROC MIXED, I recommend the following readings (available from me):

- PROC MIXED User's Guide documentation.
- Singer, Judith D. "Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models", *Journal of Educational and Behavioral Statistics*, 1998, 24(4), 323-355. (thanks Matt!).
- Littell, Hamon C., George A. Milliken, Walter W. Stroup, and Russell D. Wolfinger. 1996. "Random Coefficient Models", pp. 253-266. Chapter 7 in *SAS System for Mixed Models*. Cary, NC: SAS Institute

If you decide to use HLM, I recommend the following:

- HLM manual. I don't have a copy, but someone may have it in the department.
- Examples of HLM applications in the HLM website at www.ssicentral.com/hlm/mainhlm.htm.

MEASUREMENT ISSUES & FACTOR ANALYSIS

CLASS 6

PREPARATION QUESTIONS:

- 1) What are the differences between a construct and a measure? What is reliability? How does one know if the measures are reliable?
- 2) From a philosophical point of view, what are problems with factor analysis techniques?
- 3) What are the differences between exploratory and confirmatory factor analysis? What does confirmatory factor analysis add to the equation? Why is that important?
- 4) One of the problems in exploratory factor analysis is that everybody sees what they want to see. How does confirmatory factor analysis (CFA) change this? What is “confirmatory” about CFA?
- 5) Working through either the Bagozzi et al. paper or the Long book, try to figure out how CFA contributes to the estimation of the MTMM model.
- 6) Please spend extra care on reading about identification issues in the Long book. That is probably the most critical issue in CFA and LISREL models, so make sure that you understand where the CFA parameters come from, and why there is only one solution rather than many possible solutions.

REQUIRED READINGS:

Long, J. Scott. 1983. *Confirmatory Factor Analysis: A Preface to LISREL*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-033. Newbury Park, CA: Sage Publications.

Bagozzi, Richard P., Youjae Yi, and Lynn W. Phillips. 1991. Assessing Construct Validity in Organizational Research. *Administrative Science Quarterly*, 36: 421-458.

Fryxell, Gerald E., and Jia Wang. 1994. The Fortune corporate 'reputation' index: Reputation for what? *Journal of Management*, 20: 1-14.

Gould, Stephen J. 1981. *The Mismeasure of Man*. New York: W.W. Norton & Co. pp. 234-320.

CATCH UP READINGS:

I will assume that all of you are familiar with basic principal components and exploratory factor analysis techniques. If this is not true, you should probably read the Kim and Mueller Sage book (I don't have a copy, please find it in the library). Also, I will assume that everybody is familiar with the basic concepts of validity and reliability, including basic notions of classical test theory and basic measures such as Cronbach's alpha. If you are not, you may want to browse through Carmines and Zeller.

Carmines, Edward G. and Richard A. Zeller. 1979. *Reliability and validity assessment*. Newbury Park, CA: Sage.

Kim, J., and C.W. Mueller. 1978. *Introduction to factor analysis: What it is and how to do it*. Newbury Park, CA: Sage.

INTRODUCTION TO SYSTEMS MODELING

CLASS 7

PREPARATION QUESTIONS:

- ❖ The basic structure of an structural equation model is captured in these three equations:
$$\eta = B\eta + \Gamma\xi + \zeta$$
$$x = \Lambda_x\xi + \delta$$
$$y = \Lambda_y\eta + \epsilon$$

What is the meaning of that mumbo-jumbo? What parameters and matrices are being estimated in a structural equation model? What are the meaning of those parameters and matrices?
- ❖ How do we assess fit in an structural equation model? Fit with respect to what? In particular, pay attention to the Anderson and Gerbing method described in Hoskisson et al. and Anderson & Williams. What is the idea behind this 2-stage fit assessment? What are the steps? What models are compared to what models? What is the underlying logic of this approach?
- ❖ One of the most important areas of contention in SEM is about model testing and goodness-of-fit indicators. The chapter by Hu and Bentler provides a great review of some of those indicators and issues of contention. Please pay attention to those indicators, since we will use them in the assignment.
- ❖ Structural equation modeling is often used to estimate model with reciprocal causation. Why not use OLS models? What is different for non-recursive models that make them problematic to estimate with regular regression?
- ❖ SEM methodology is particularly flexible, in that you can write many other methods as SEM models. Let's try if you can write the following models as SEM models:
 - An OLS regression.
 - Two OLS regressions with correlated error terms.
 - Two OLS regressions where the DV of one equation is an IV for another equation, and vice versa (reciprocal causation).
 - A confirmatory factor analysis with two correlated factors.
 - A regression model with latent variables as IVs and DVs.
 - A cross-lagged regression model with latent variables as IVs and DVs.
- ❖ Compare the two SEM papers that you have read Hoskisson et el. And Anderson and Williams in terms of their writing and explanations. How well do they follow the guidelines by Hoyle and Panter?

MOST IMPORTANT REQUIRED READING:

Prior to your assignment, you should thoroughly read the documentation for whatever software you will use. Only by carefully reading the manuals you will be familiar with all the options and possibilities of your preferred software.

REQUIRED READINGS:

Long, J. Scott. 1983. *Covariance Structure Models: An Introduction to LISREL*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-034. Newbury Park, CA: Sage Publications.

(pages 45-92) Finkel, Steven E. 1995. *Causal Analysis with Panel Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-105. Newbury Park, CA: Sage Publications.

Anderson, James C., and David W. Gerbing. 1988. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Journal of Applied Psychology*, 103(3), 411-423.

Hu, Li-Tze, and Peter M. Bentler. 1995. 'Evaluating Model Fit', in Rick H. Hoyle (Ed.) *Structural Equation Modeling: Concepts, Issues, and Applications*. Sage Pubs: Thousand Oaks, CA., pp. 76-99.

Hoyle, Rick H., and Abigail T. Panter. 1995. 'Writing about Structural Equation Models', in Rick H. Hoyle (Ed.) *Structural Equation Modeling: Concepts, Issues, and Applications*. Sage Pubs: Thousand Oaks, CA., pp. 158-176.

Applications:

Hoskisson, R.E., Johnson, R.A., & Moesel, D.D. 1994. Corporate divestiture intensity in restructuring firms: Effects of governance, strategy, and performance. *Academy of Management Journal*, 37: 1207-1251 (you already read this paper for your strategy seminar. Now focus on the methods).

Anderson, Stella E. and Larry J. Williams. 1996. Interpersonal, Job, and Individual Factors Related to Helping Processes at Work. *Journal of Applied Psychology*, 81(3): 282-29.

RECOMMENDED READING:

Williams, Larry J. and Philip M. Podsakoff. 1989. 'Longitudinal Field Methods for Studying Reciprocal Relationships in Organizational Behavior Research: Toward Improved Causal Analysis'. *Research in Organizational Behavior*, vol. 11, 247-292. (Very good introduction to dynamic causal analysis using SEM, but it overlaps a bit with Finkel).

MODELS FOR CATEGORICAL VARIABLES

CLASS 8

PREPARATION QUESTIONS:

- ❖ Suppose that, in a piece of research, your DV is a categorical binary outcome (yes-no; stay-go; etc).
 - What are the possible methodologies that you can use to study such DV? List both the ones discussed in the readings and others that you know about. List both univariate and multivariate methods.
 - What's wrong with estimating an OLS models with a dummy dependent variable?
- ❖ What are the differences between the following models? What are their assumptions?
 - Linear probability model.
 - Logit model.
 - Probit model.
 - Ordered logit model.
 - Multinomial logit model.
- ❖ Suppose now that your DV is a count variable.
 - Give me an example in your research in which you are likely to encounter count dependent variables.
 - What are the possible methodologies that you can use to study a count DV? List both the ones discussed in the readings and others that you know about. List both univariate and multivariate methods.
 - What's wrong with estimating an OLS models with a count dependent variable?
- ❖ What are the differences between the following models?
 - Poisson model.
 - Negative Binomial model.
- ❖ The estimation of the models covered today is always done using maximum likelihood estimation. Compare ML (e.g., logistic regression) and OLS estimation in these dimensions:
 - Underlying distributional assumptions.
 - Measures of goodness-of-fit.
 - Test of hypotheses (what statistics, what tests?).
- ❖ Suppose that you have a 5-item Likert-scale question as a dependent variable. Evaluate the appropriateness of using the following methods to analyze that dependent variable. What would be best? What would be wrong? What would be more or less conservative?
 - OLS
 - Multinomial logit
 - Ordered logit
 - Poisson or negative binomial.

READINGS: THEORY.

(Chapters 1-6 and 8) Long, J. Scott. 1997. *Regression models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences Series, 7. Thousand Oaks, CA: Sage.

READINGS: APPLICATIONS. (SKIM THEORY-FOCUS ON METHODS SECTIONS).

Folta, T.B. Governance and Uncertainty: The Tradeoff between Administrative Control and Commitment. *Strategic Management Journal*, 19(11): 1007-1028.

Henderson, R. and Ian Cockburn. 1994. Measuring competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal*, 15 (Winter Special Issue): 63-84.

**CENSORED AND TRUNCATED MODELS.
SAMPLE-SELECTION AND SELF-SELECTION MODELS.**

CLASS 9

PREPARATION QUESTIONS:

- ❖ What is the difference between censored and truncated samples?
- ❖ What is the Tobit model? How does it differ from a probit model? How does it differ from a regression model? Can you think of an application of the tobit model in your line of research?
- ❖ Why are we concerned about sample selection bias? How does sample selection bias influence the internal and external validity of our analysis and results?
- ❖ There is an old maxim in statistics that one “should not sample on the dependent variable”. Essentially, this maxim dissuades from going out and picking a group of observations because they exhibit very high or low values of the variable that you are interested in. For instance, if you are interested in firms performance, you should not go out and collect a group of best-performing firms and then draw implications from that sample. What is the statistical logit of that maxim? What threats to validity are involved in the practice of sampling on the DV?
- ❖ Suppose that you did a survey of XXX, and you obtained a 50% response rate. A reviewer is making a big fuss that your sample may not be representative of the population? What can you do, either in terms of logical argumentation or statistical argumentation, to defend yourself from that claim? What tests could you carry out to show that the sample is not biased?
- ❖ Several papers in this section state that strategy is essentially self-selected (i.e., firms are likely to take whatever strategy is best for them). What are the implications of that view for the analysis of the strategy-performance relationship? (For micro students: human behavior may be goal oriented; what are the implications about the assessment of the behavior-goal relationship?).
- ❖ What is the infamous “inverse of Mill’s ratio”? How would you go about calculating it?
- ❖ Think about designing a study in your area of research where you could apply the theory and methods of self-selection.

READINGS: THEORY.

(Chapter 7) Long, J. Scott. 1997. *Regression models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences Series, 7. Thousand Oaks, CA: Sage.

Breen, Richard. 1996. *Regression Models: Censored, Sample-Selected, or Truncated Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-111. Thousand Oaks, CA: Sage.

Berk, R.A. 1983. An introduction to sample selection bias in sociological data. American Sociological Review, 48: 386-398.

Masten, S.E. 1993. Transaction costs, mistakes, and performance: Assessing the importance of governance. Managerial and Decision Economics, 14: 119-129.

Shaver, J. M. 1998. Accounting for endogeneity when assessing strategy performance: does entry mode choice affect FDI survival? Management Science, 44(4): 571-585.

READINGS: APPLICATIONS. (SKIM THEORY-FOCUS ON METHODS SECTIONS).

Gimeno, J., Folta, T.B., Cooper, A.C., & Woo, C.Y. 1997. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. Administrative Science Quarterly, 42: 750-783.

Poppo, L. and T. Zenger. 1998. Testing alternative theories of the firm: Transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services. Strategic Management Journal, 19(9): 853-877.

INTRODUCTION TO EVENT HISTORY ANALYSIS

CLASS 10

PREPARATION QUESTIONS:

- ❖ What is event history analysis? What is the difference relative to other longitudinal methods?
- ❖ What is the hazard rate? Is it a probability? How does it relate to the survivor function? Suppose that you had a continuous-time model of job quitting? What would be the relationship between the hazard rate (hazard function) and the probability that a person will quit during a given year?
- ❖ What is censoring? What is left-censoring? What is right censoring? Which one(s) should you care about?
- ❖ What are the differences between continuous-time and discrete-time methods?
- ❖ What are the differences between (1) nonparametric, (2) semiparametric, and (3) parametric event history models? Could you classify the models that you have read about in these three categories? What are the assumptions underlying each? What are their weaknesses/strengths?
- ❖ What is the “proportional hazards” assumption? Under what conditions would this assumption be violated?
- ❖ What are the life-table and Kaplan-Meyer methods of event history? What do they estimate?
- ❖ Morita et al. distinguish between “ANOVA-analog” and “regression-analog” survival models. How are those models different?

READINGS: THEORY.

- (Chapters 1-4) Yamaguchi, Kazuo. 1991. Event History Analysis. Applied Social Research Methods series, volume 28. Newbury Park, CA: Sage Pubs.
- Morita, June G, Thomas W. Lee, and Richard T. Mowday. 1989. Introducing survival analysis to organizational researchers: A selected application to turnover research. Journal of Applied Psychology, 74(2): 280-292.
- Morita, June G, Thomas W. Lee, and Richard T. Mowday. 1993. The regression-analog to survival analysis: A selected application to turnover research. Academy of Management Journal, 36(6), 1430-1464.
- Box-Steffensmeier, Janet M., and Bardford S. Jones. 1997. Time is of the Essence: Event History Models in Political Science. American Journal of Political Science, 41(4): 1414-1461.
- Blossfeld, Hans-Peter, and Götz Rohwer. 1995. 'Introduction'. Chapter 1 in Techniques of Event history Modeling. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 1-32.

EVENT HISTORY ANALYSIS: CONTINUATION

CLASS 11

PREPARATION QUESTIONS:

- ❖ What are the different ways to handle time-varying variables?
- ❖ What is episode-splitting? Why can one break an event history into multiple episodes? What are the advantages of this property?
- ❖ What are competing risk models? How can these models be estimated in (a) continuous time models, and (b) discrete time models?
- ❖ Some times, we try to model events that are repeated over time. For instance, job quitting can occur multiple times for an individual. How can a researcher handle repeated events?
- ❖ Suppose that you want to test the following hypotheses in an event history model. How would you construct your dataset to do it?
 - The rate of a quitting a job increases during the May-September period.
 - The rate of quitting a job decreases with the age of individual.
 - The rate of quitting a job decreases with the time since the last promotion.
 - The rate of announcing a merger increases with the number of announced mergers in the industry.
 - The rate of announcing a merger increases with the number of announced mergers in the last three years.
- ❖ What are the advantages/disadvantages of discrete time modeling? If you had a model in which you observe the job quitting behavior of individuals to the closest month, would you model this as discrete or continuous?
- ❖ What is time (or duration) dependence? Why is it important? What are the ways in which we can model time dependence in an event history model?
- ❖ What are the different assumptions of the Cox model relative to (a) parametric models and (b) non-parametric models? The Cox model is one of the most popular models in statistics. What are the appealing properties of this model?

READINGS: THEORY.

- (Chapters 5-7) Yamaguchi, Kazuo. 1991. Event History Analysis. Applied Social Research Methods series, volume 28. Newbury Park, CA: Sage Pubs.
- Petersen, Trond. 1991. The statistical analysis of event histories. Sociological Methods and Research, 19: 270-323.
- Allison, Paul D. 1982. Discrete-time methods for the analysis of event histories. In S. Leinhardt (Ed.) Sociological Methodology 1982, 61-98. San Francisco: Jossey-Bass.
- Hachen, David S. 1988. The Competing Risks Model. Sociological Methods and Research, 17(1): 21-54.

READINGS: APPLICATIONS.

(EACH STUDENT WILL BRIEFLY (5 MINUTES) PRESENT THE HIGHLIGHTS OF ONE OF THESE PAPERS, WITH A SPECIAL EMPHASIS ON THE METHODS SECTION. FOR THE OTHER PAPERS, PLEASE SKIM THE THEORY AND READ THE METHODS AND RESULTS)

Amburgey, Terry L., Dawn Kelly, and William P. Barnett. 1993. Resetting the clock: The dynamics of organizational change and failure. Administrative Science Quarterly, 38(1): 51-73.

Barnett, William P., and Anne S. Miner. 1992. Standing on the Shoulders of Others: Career Interdependence in Job Mobility. Administrative Science Quarterly, 37(2): 262-281.

Pennings, Johannes M., Kyungmook Lee, and Arjen van Witteloostuijn. 1998. Human capital, social capital, and firm dissolution. Academy of Management Journal, 41(4): 425-440.

Penning, Johannes M., Harry Barkema, and Sytse Douma. 1994. Organizational learning and diversification. Academy of Management Journal, 37(3): 608-640.

Henderson, Andrew D. 1999. Firm strategy and age dependence: A contingent view of the liabilities of newness, adolescence, and obsolescence. Administrative Science Quarterly, 44(2): 281-314.

Ocasio, William. 1999. Institutional action and corporate governance: The reliance on rules of CEO succession. Administrative Science Quarterly, 44: 384-416.

**EVENT HISTORY MODELING (III)
FINANCIAL EVENT STUDIES**

CLASS 12

PREPARATION QUESTIONS:

- ❖ What are the similarities and differences between event history studies and financial event studies?
- ❖ Why are financial event studies used in management research? What are their advantages and disadvantages relative to other methods for studying the financial performance effects of organizational strategies?
- ❖ What are the main assumptions underlying financial event studies? How could a researchers check that those assumptions hold?
- ❖ Suppose that you wanted to run a financial event study, and you did not have software (like Eventus) that can do it for you. What are the main steps involved in doing a financial event study?

READINGS: FINANCIAL EVENT STUDIES.

Barney, J. B. 1997. Gaining and Sustaining Competitive Advantage. Boston, MA: Addison-Wesley (chapter 2 entitled “What is performance?”, pp. 30-64).

McWilliams, Abigail and Donald Siegel. 1997. Event studies in Management Research: Theoretical and Empirical Issues. Academy of Management Journal, 40(3), 626-657.

Brown Stephen J., & Jerold B. Warner. 1985. Using daily stock returns: The case of event studies. Journal of Financial Economics, 14: 3-31.

Anderson, Ronald C. & D. Scott Lee. 1997. Ownership studies: The data source does matter. Journal of Financial and Quantitative Analysis, 32(3): 311-329 + supplement.

Anand, B.N., & T. Khanna. 2000. "Do firms learn to create value? The case of alliances. Strategic Management Journal, 21: 295-315.

Eventus manual (available in electronic form).

INTRODUCTION TO NETWORK ANALYSIS

CLASS 13

PREPARATION QUESTIONS:

- ❖ Is there a “network theory”? Or is there just a “network analysis” methodology? To what extent does network analysis contribute a new theoretical perspective to management research?
- ❖ Network analysis requires a different way to collect and organize your data. How can network data be collected? What are the main concerns associated with collecting network data?
- ❖ What are the boundaries of a network? How can those boundaries be determined theoretically and empirically?
- ❖ There are three main levels of analysis that can be used at the network level: the individual, the dyad, and the network or group/subgroup level. For a given relation, what measures can be used for describing that relation at each different level? What are the important descriptive measures at the ego level? What are the measures at the dyadic level? What are the network level or group level measures?
- ❖ A major goal of network analysis is to divide large groups into smaller subgroups. What approaches are there to do that. What is the difference between groups based on cohesion (i.e., cliques) and groups based on structural equivalence?
- ❖ Develop one testable hypothesis in your area of interest that could be tested with a network analysis methodology.

READINGS: NETWORK ANALYSIS.

Scott, John. 1991. Social Network Analysis: A Handbook. Newbury Park, CA: Sage Pubs.

Salancik, Gerald R. 1995. WANTED: A good network theory of organization. Administrative Science Quarterly, 40: 345-349.

Marsden, Peter V. 1990. Network data and measurement. Annual Review of Sociology, 16: 435-463.

Galaskiewicz, Joseph, and Stanley Wasserman. Social Network Analysis: Concepts, Methodology, and Directions for the 1990s. Sociological Methods and Research, 22(1): 3-22.

Krackhardt, David, and Daniel J. Brass. 1994. Intraorganizational networks. In Wasserman and Galaskiewicz (Eds.) Advances in Social Network Analysis, 207-229.

Mizruchi, Mark S, and Joseph Galaskiewicz. 1994. Networks of interorganizational relations. In Wasserman and Galaskiewicz (Eds.) Advances in Social Network Analysis, 230-253.

NETWORK ANALYSIS: CONTINUATION

CLASS 14

PREPARATION QUESTIONS:

- ❖ What is structural equivalence? How does it differ from cohesion? Find an example in which you could clearly distinguish between structural equivalence and cohesion.
- ❖ What are structural holes? Why the advantages or benefits associated with being in a position rich in structural holes? How are structural holes measured in practice?
- ❖ What is social contagion/diffusion? How is social contagion/diffusion influenced by cohesion and structural equivalence? What is the theoretical logic for both effects?
- ❖ What statistical models can be used to test hypotheses relating to social contagion and diffusion?
- ❖ Suppose that you want to study a particular dyadic relationship using network data. What are the threats to validity associated with using OLS methods? In particular, what is network autoregression? What are the solutions? What is the Quadratic Assignment Procedure (QAP)? Why is it done?
- ❖ Suppose that you want to know whether a particular relation (x) is being reciprocated. Could you simply look at the correlation between x_{ij} and x_{ji} ? Are there any reasons why this correlation maybe invalid for statistical inference? How could you build an appropriate empirical model to test reciprocity of relationship?
- ❖ Develop one hypothesis about social diffusion, and think about how this hypothesis would be tested. In particular, think about how the model should be specified in order to test the hypothesis.

READINGS: THEORY.

Marsden, Peter V., and Noah E. Friedkin. 1994. Networks Studies of Social Influence. In Wasserman and Galaskiewicz (Eds.) Advances in Social Network Analysis, 3-25.

Lincoln, James R. 1984. Analyzing Relations in Dyads. Sociological Methods and Research, 13(1), 45-76.

Mizruchi, Mark S. 1993. Cohesion, equivalence, and similarity of behavior: a theoretical and empirical assesment. Social Networks, 15: 275-307.

Burt, Ronald S. 1997. The contingent value of social capital. Administrative Science Quarterly, 42: 339-365.

Strang, David, and Nancy B. Tuma. 1993. Spatial and temporal heterogeneity in diffusion. American Journal of Sociology, 99: 614-639.

READINGS: APPLICATIONS OF NETWORK ANALYSIS.

(EACH STUDENT WILL READ AND BRIEFLY (5 MINUTES) PRESENT ONE OF THESE PAPERS, OR ANOTHER NETWORK ANALYTIC PAPER OF THEIR CHOICE, WITH A SPECIAL EMPHASIS ON THE METHODS SECTION. – FOR THE OTHER PAPERS, PLEASE SKIM THE THEORY AND READ THE METHODS AND RESULTS).

Ibarra, Herminia, and Steven B. Andrews. 1993. Power, social influence, and sense making: effects of network centrality and proximity on employee perceptions. Administrative Science Quarterly, 38(2): 277-303.

Krackardt, David, and Lyman W. Porter. 1985. When friends leave: A structural analysis of the relationship between turnover and stayers' attitudes. Administrative Science Quarterly, 30: 242-261.

Greve, Henrich R. 1995. 1995. Jumping ship: The diffusion of strategy abandonment. Administrative Science Quarterly, 40: 444-473.

Gulati, Ranjay. 1995. Social Structure and Alliance Formation Patterns: A Longitudinal Analysis. Administrative Science Quarterly, 40: 619-652.

Lincoln, James R., Michael L. Gerlach, and Peggy Takahashi. 1992. Keiretsu Networks in the Japanese Economy: A Dyad Analysis of Intercorporate Ties. American Sociological Review, 57: 561-585.

Madhavan, Ravindranath, Balaji R. Koka, and John E. Prescott. 1998. Netoworks in transition: How industry events (re)shape interfirm relationships. Strategic Management Journal, 19: 439-459.