

**Special Topic: Data Analysis in OB Research
Spring 2020, B Term**

Instructor: Le (Betty) Zhou, Ph.D.

Office: CSOM 3-283

Email: zhoul@umn.edu

Class meeting time & location: 1-4pm Tuesday, CSOM 3-166

Office hours: By appointment

Course Overview & Expectations

The goals of this course are (1) to provide an overview of data analyses tools commonly used in OB research, (2) to build the foundation for future self-directed learning of related and more advanced methods, and (3) to help students make progress in developing their own program of research. This course is designed for PhD students interested in empirical research using quantitative methods in Organizational Behavior (OB) field.

Class activities include pre-class readings (required and optional), class lecture, discussion, and homework. Students should complete pre-class preparation work (see details below) prior to the first class meeting. Each week, students should complete required readings before class. Class meetings will include lectures and discussions. Slides or handouts of the lectures will be shared with students before class. All class material, including slides and homework assignments, are intellectual properties of the instructor and should not be shared online or with others without the instructor's permission. Students are expected to participate in class discussion and proactively ask questions if they have any. Students should individually complete each week's homework after class, which is designed to help students practice the methods covered in the course. Ultimately, the students should be the driver of their learning process throughout this course.

Demonstrations in class will use SPSS, *Mplus*, and R. For homework assignments, students can use any software packages which they are familiar with and have the capacity for the analyses needed. Regardless of the statistical packages used, students should submit their homework in a WORD document that includes a manuscript-like report of their analyses and results, using text, figures, and tables. Students can follow the style of reporting in the "Analytic Strategy" and "Results" sections in a management/OB journal they are familiar with (e.g., *Journal of Applied Psychology* or *Academy of Management Journal*).

Grade & Assignments

Pre-class test <i>*Note: Students who perform poorly in this test will be recommended to withdraw from this course.</i>	10 points
Class participation: Complete required readings before class + Participate in class discussion	20 points
Homework: Each week's homework assignment (10 points each) should be completed and emailed to me before 5pm CT Sunday. Late submission will not receive any feedback from me or any points.	70 points

Total: 100 points

Schedule

Pre-class Preparation

Learning goals: Understand the conceptual meaning of the following concepts and be able to apply them to solve problems in social science research:

- Mean, median, mode
- Variance, standard deviation
- Probability distribution
- z score
- Population, sample, sampling distribution, estimate, standard error, confidence interval
- Null hypothesis significance test, p value
- Type I & Type II errors, effect size, statistical power
- Independent samples t -test, paired t -test
- ANOVA, planned comparison/constrast, post hoc tests
- Correlation, covariance
- Ordinary least squares (OLS) regression

Assignment:

- Khan Academy self-directed learning material shared by the instructor
- Complete pre-class test before 5pm Sunday 3/15. Test will be shared on Monday 3/9. This test is open notes.

Week 1 (3/17): SEM/general latent variable modeling framework

Learning goals: Understand foundational structural equation modeling (SEM) concepts; develop the foundation for applying SEM framework to more advanced analysis (e.g., multilevel SEM).

- Latent variable, observed variable, measurement error
- Model fit indices, model comparison
- Multiple-group analysis
- General latent variable modeling framework
- Maximum likelihood estimation
- Missing data
- *Bayes' theorem, Bayesian approach¹

Required readings:

- Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). Structural equation modeling in management research: A guide for improved analysis. *Academy of Management Annals*, 3, 543-604.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychology Bulletin*, 103, 411-423.

¹ We will cover topics noted by an asterisk if we have time.

- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1, 104-121.
- Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, 17(4), 372-411.
- Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, 41, 390-420.

Optional readings:

- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879-903.
- Spector, P. E., Rosen, C. C., Richardson, H. A., Williams, L. J., & Johnson, R. E. (2019). A new perspective on method variance: A measure-centric approach. *Journal of Management*, 45(3), 855-880.
- Cortina, J. M., & Landis, R. S. (2011). The earth is *not* round ($p = .00$). *Organizational Research Methods*, 14, 332-349.
- Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, 15, 722-752.

Assignment: Homework 1

Week 2 (3/24): Moderation & congruence analyses

Learning goals:

- Interaction, moderation
- Simple slope, region of significance
- Graphic presentation of moderation analysis results
- Polynomial regression
- Graphic presentation of polynomial regression results (response surface)

Required readings:

- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed., Chapter 7, pp. 255-300). Mahwah, NJ: Lawrence Erlbaum.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed., Chapter 9, pp. 354-388). Mahwah, NJ: Lawrence Erlbaum.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437-448.
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal*, 36, 1577-1613.

Optional readings:

- Edwards, J. R. (2009). Seven deadly myths of testing moderation in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 143-164). New York: Routledge.
- Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes*, 64, 307-324.
- Zyphur, M. J., Zammuto, R. F., & Zhang, Z. (2016). Multilevel latent polynomial regression for modeling (in) congruence across organizational groups: The case of organizational culture research. *Organizational Research Methods*, 19(1), 53-79.

Assignment: Homework 2

Week 3 (3/31): Mediation analysis

Learning goals:

- Mediation
- Indirect effects
- Integrating mediation and moderation in single-level models
- *Categorical outcome variable, count outcome variable

Required readings:

- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58, 593-614.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422-445.
- Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, 12, 1-22.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, 12, 23-44.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed., Chapter 13, pp. 479-535). Mahwah, NJ: Lawrence Erlbaum.

Optional readings:

- MacKinnon, D.P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93-115.

- Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, 18(3), 233-239.
- Taylor, A. B., MacKinnon, D. P., & Tein, J. Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, 11(2), 241-269.

Assignment: Homework 3

Week 4 (4/7): Multilevel analysis I

Learning goals: This week we will discuss multilevel concepts and basic idea of multilevel modeling.

- Levels-of-analysis issues in conceptualization and operationalization of constructs
- Aggregation tests, ICC(1), ICC(2), r_{wg}
- Basic multilevel model

Required readings:

- Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and empirical processes. In K. J. Klein and S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (chapter 1, pp. 3-90), San Francisco: Jossey-Bass.
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83, 234-246.
- LeBreton, J. M. & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11, 815-852.
- Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management*, 23, 723-744.

Optional readings:

- Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, 7, 400-417.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (pp. 3-67). Sage.

Assignment: Homework 4

Week 5 (4/14): Multilevel analysis II

Learning goals: Continuing from last week, we will discuss

- Centering
- Integrating moderation and mediation in multilevel models

- *Social relations model

Required readings:

- Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24(5), 623-641.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, 12(2), 121-138.
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15, 209-233.
- Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, 12(2), 347-367.
- Knight, A. P., & Humphrey, S. E. (2019). Dyadic data analysis. In S. E. Humphrey and J. M. LeBreton (Eds.), *The Handbook for Multilevel Theory, Measurement, and Analysis* (pp. 423-447). Washington, DC: American Psychological Association.

Optional readings:

- Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, 21(2), 189-205.
- Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, 97(5), 951-966.

Assignment: Homework 5

Week 6 (4/21): Longitudinal data analysis I

Learning goals: This week we start from some longitudinal data analyses tools that are already popular among OB researchers:

- Cross-lagged model
- Latent growth model
- Latent change score model

Required readings:

- Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505-528.
- Liu, Y., Mo, S., Song, Y., & Wang, M. (2016). Longitudinal analysis in occupational health psychology: A review and tutorial of three longitudinal modeling techniques. *Applied Psychology: An International Review*, 65(2), 379-411.
- Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS)

and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, 1, 421-483.

- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577-605.

Optional readings:

- Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, 5, 362-387.
- Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, 38, 529-569.

Assignment: Homework 6

Week 7 (4/28): Longitudinal data analysis II

Learning goals: Continuing from last week, we will discuss

- Centering in longitudinal analyses
- Time series analysis
- Intensive longitudinal data analysis
- Survival analysis
- *Latent categorical variable analysis and application in longitudinal analyses

Required readings:

- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 583-619.
- Jebb, A. T., & Tay, L. (2017). Introduction to time series analysis for organizational research: Methods for longitudinal analyses. *Organizational Research Methods*, 20, 61-94.
- Zhou, L., Wang, M., & Zhang, Z. (in press). Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods*.
- Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 18(2), 155-195.
- Wang, M., & Hanges, P. (2011). Latent class procedures: Applications to organizational research. *Organizational Research Methods*, 14, 24-31.

Optional readings:

- Hamaker, E. L., & Muthén, B. (2019). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods*.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis* (chapters 9-15). Oxford University Press.

- Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, 10, 635-656.
- Wang, M., & Chan, D. (2011). Mixture latent Markov modeling: Identifying and predicting unobserved heterogeneity in longitudinal qualitative status change. *Organizational Research Methods*, 14, 411-431.

Assignment: Homework 7

Week 8 (5/5): CARMA binge watch (in CSOM 4-300V)

12:00pm-12:10pm: set up, lunch

12:10pm-1:20pm: Big data concepts – Sang Eun Woo

1:25pm-2:35pm: Big data analytics – Fred Oswald

2:40pm-3:50pm: Causal identification through a cumulative body of research in the study of strategy and organizations – Myles Shaver