Quantitative Research Methods Comprehensive Exam

READING LIST

UNIVERSITY OF TORONTO,

DEPARTMENT OF SOCIOLOGY

Last update: February, 2019

Area's Standing Committee, 2018-2019: Blair Wheaton (chair), Andrew Miles, Ethan Fosse.

Other faculty in the area (alphabetical order): Monica Alexander, Irene Boeckmann, Ethan Fosse, Andrew Miles, David Pettinicchio, Markus Schafer, Scott Schieman, **Blair Wheaton, Geoffrey Wodtke**

REQUIRED READINGS

Total number of reading units (excluding supplemental): 173 units

The reading list includes required readings and supplementary readings. Supplementary readings include references to works that will facilitate further study of a particular topic. Supplementary sections also include related work by faculty in the Department of Sociology at the University of Toronto. Students must read all 173 required reading units plus 12 supplementary reading units selected from any two topics, for a total of 185 reading units. Although there are some exceptions, units typically determined as follows:

- 1 large book > 100 pages = 5 reading units
- Small books and parts of larger books = 2-3 reading units, depending on amount of pages assigned
- 1 journal article or book chapter = 1 reading unit

MODULES

1. Background Readings (12 units)
2. Generalized Linear Models and Related Methods (20 units)
3. Causal Models, Graphical Chains, and Structural Equations (44 units)
4. Multilevel Models and Contextual Effects (30 units)
5. Methods for Longitudinal Data (32 units)
6. Sampling and Inference Issues (20 units)
7. Big Data, Online Data and Sampling (15 units))

1. Background Readings (12 units)

Kieran Healy. 2019. Data Visualization. Princeton University Press. (5)

Moore, Will H. and David A. Siegel. 2013. A Mathematics Course for Political and Social Research. Princeton, NJ: Princeton University Press. (5)

Nettler, Gwynne. 2003. Boundaries of Competence: How Social Studies Makes Feeble Science. New Brunswick, N.J.: Transaction (chapters 1-2) (1)

Raftery, A. E. (2001). Statistics in Sociology, 1950-2000: A Selective Review. Sociological Methodology, 31, 1-45. (1)

***Supplemental Readings* :**

Aneshensel, Carol S. 2002 Theory Based Data Analysis for the Social Sciences. CA: Pine Forge.

Blalock, H.M. 1968. "The measurement problem: a gap between the languages of theory and research." Pp. 5-27 in Blalock, Hubert M. and Ann Blalock (eds.) Methodology in Social Research. New York: McGraw-Hill.

Cortina, Jose M. 1993. "What is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology*, 78: 98-104.

Gill, Jeff. 2006. Essential Mathematics for Political and Social Research. Cambridge, UK: Cambridge University Press.

Nettler, Gwynn. 1970. *Explanations*. New York: McGraw-Hill.

Suen, H.K. 1990. "Basic Concepts of Item Response Theory." Pp. 83-98 in H.K. Suen (ed.) Principles of Test Theories. Lawrence Erlbaum Associates.

Wheaton, B. 2003. "When methods make a difference". *Current Sociology.* 51(5) 543-72.

Xie, Yu. 2007. "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." Research in Social Stratification and Mobility 25(2):141–56.

2. Generalized Linear Models and Related Methods (20 units)

2(a). Linear and Generalized Linear Models (**12 units**)

Abbott, Andrew. 1988. "Transcending General Linear Reality." Sociological Theory 6(2):169. (1)

Frank, Kenneth A. 2000. "Impact of a Confounding Variable on a Regression Coefficient," Sociological Methods and Research, 29(2): 147-194. (1)

Long, J. S. 1997. Regression models for categorical and limited dependent variables. Thousand Oaks, CA: Sage.. (5)

McCullagh, P. and J.A. Nelder. 1989. Generalized Linear Models (2nd Edition). New York: Chapman & Hall. (5)

2 (b). Interpretation and Presentation of Statistical Models **(3 units)**

Firth, David. 2003. "Overcoming the reference category problem in the presentation of statistical models," Sociological Methodology, 33:1-18. (1)

Fox, J. 1987. "Effect Displays for Generalized Linear Models," Pp. 347-61 in Sociological Methodology, vol. 17, edited by C. C. Clogg. Washington, DC: American Sociological Association. -1997. (1)

Silber, J.H., P.R. Rosenbaum and R.N. Ross. 1995. "Comparing the Contributions of Groups of Predictors: Which Outcomes Vary With Hospital Rather than Patient Characteristics," Journal of the American Statistical Association, 90 (429): 718. (1)

2(c). Nonparametric Regression (**5 units**)

Gareth, James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. An Introduction to Statistical Learning. New York: Springer. (5)

Supplementary Readings

Agresti, Alan. 2010. Analysis of Ordinal Categorical Data, Second Edition. Hoboken, New Jersey: John Wiley & Sons.

Aldrich, J.H. and F.D. Nelson. 1984. Linear Probability, Logit, and Probit Models. Beverly Hills, CA: Sage Publications.

Andersen, Robert. 2008. Modern Methods for Robust Regression. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-152. Thousand Oaks, CA: Sage.

Barron, David N. 1992. "The Analysis of Count Data: Overdispersion and Autocorrelation." Sociological Methodology, Vol. 22. pp. 179-220.

Becker, M. 1994. "Analysis of Cross-Classification of Counts Using Models for Marginal Distributions: An Application to Trends in Attitudes on Legalized Abortion." Sociological Methodology, P. Marsden (ed.). Oxford: Blackwell.

Bishop, Y., S. Feinberg, and P. Holland. 1975. Discrete Multivariate Analysis. Cambridge: MIT Press.

DeMaris, Alfred. 1992. Logit Modeling. Newbury Park, CA: Sage.

Fox, John, and Robert Andersen. 2006. "Effect Displays for Multinomial and Proportional-odds Logit Models," Sociological Methodology 36(1) 225-255

Fox, John. 2000. Multiple and Generalized Nonparametric Regression. (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-130). Thousand Oaks, CA: Sage.

Harrell, Frank E, Jr. 2001. Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer.

Long, J. S. & Freese, J. 2005. Regression models for categorical dependent variables using Stata. 2nd Edition. College Station TX: Stata Press.

Powers, Daniel A. and Yu Xie. 2000. Statistical Methods for Categorical Data Analysis. Academic Press.

3. Causal Models, Graphical Chains, and Structural Equations (42 units)

3(a). Causal Models (*13 units*)

Alwin, D. F. and R. M. Hauser. 1975. "The Decomposition of Effects in Path Analysis." *American Sociological Review*, 40:37-47. (1)

Cook, Thomas D., and Donald T. Campbell. 1979. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally. (3)

Cox, David R., and Nanny Wermuth. 1993. "Linear Dependencies Represented by Chain Graphs." Statistical Science 8: 204–18. (1)

Duncan, O. D. 1975. *Introduction to Structural Equation Models*. New York: Academic. Chapters 1-3. (2).

Elwert, Felix and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology* 40(1):31–53. (1)

Elwert, Felix. 2014. "Graphical Causal Models." Pp. 245-274 in Handbook of Causal Analysis for Social Research, edited by Stephen L. Morgan. New York: Springer. (1)

Stolzenberg, R. M. 1980. "The Measurement and Decomposition of Causal Effects in Nonlinear and Nonadditive Models." *Sociological Methodology*, 459-479. (1)

Vanderweele, Tyler. 2015. Explanation in Causal Inference. Oxford: Oxford University Press. Chapters 1-2. (2)

Winship, C. and R. D. Mare. 1983. "Structural Equations and Path Analysis for Discrete Data." *American Journal of Sociology*, 89:54-110. (1)

3(b) Causal Inference (12 units)

Gangl, Markus. 2014. "Partial Identification and Sensitivity Analysis." Pp. 377-402 in Handbook of Causal Analysis for Social Research, edited by Stephen L. Morgan. New York: Springer. (1)

Manski, Charles F. 1993. "Identification Problems in the Social Sciences," Sociological Methodology, Vol. 23. pp. 1-56. (1)

Morgan, Stephen L. and Christopher Winship. 2015. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge: Cambridge University Press. Second Edition. (5)

Sobel, Michael E. 1996. "An Introduction to Causal Inference," Sociological Methods and Research, 24 (3): 353-379. (1)

Special Issue of European Sociological Review on Causality. 2001. Volume 17 (1). [John H. Goldthorpe. "Causation, Statistics, and Sociology," 1-20; Ulrich Pötter and Hans-Peter Blossfeld. "Causal Inference from Series of Events," 21-32; Elja Arjas. "Causal Analysis and Statistics: A Social Sciences Perspective," 59-64; D.R. Cox and Nanny Wermuth. "Some Statistical Aspects of Causality," 65-74] (4)

3(c). Structural Equations with Latent Variables (19 units)

Bielby, W.T. and R. Matsueda. 1991. "Statistical Power in Nonrecursive Models." Sociological Methodology 21:167-197. (1)

Bollen, K. 1989. Structural Equations with Latent Variables. New York: Wiley. (5)

Bollen, K., and J. S. Long. 1993. Testing Structural Equation Models. Newbury Park, CA: Sage. (5)

Kline, Rex B. 2015. Principles and Practice of Structural Equation Modeling, 4th Edition. Guildford Press. (5)

Masyn, Katherine E. "Latent Class Analysis and Finite Mixture Modeling." The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis, edited by Todd D. Little. Oxford. (2)

Muthen, B. O. 1984. "General Structural Equation Models with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators." Psychometrika, 49:115-132. (1)

Supplementary Readings

Alwin, D. F. and D.J. Jackson. 1980. "Measurement Models for Response Errors in Surveys: Issues and Applications." Pp. 68-119 in *Sociological Methodology,* K.F. Schuessler (ed.). San Francisco: Jossey-Bass.

Berry, W. D. 1984. *Nonrecursive Causal Models*. Sage Series on Quantitative Applications in the Social Sciences. Newbury Park, CA: Sage.

Bollen, Kenneth A. 1995. "Structural Equation Models That are Nonlinear in Latent Variables: A Least-Squares Estimator. " *Sociological Methodology,* Vol. 25. pp. 223-251.

Duncan, Otis Dudley. 1966. "Path Analysis: Sociological Examples." American Journal of Sociology 72(1):1–16.

Jae-On, Kim. 1984. "An Approach to Sensitivity Analysis in Sociological Research." *American Sociological Review*, 49:272-282.

Hout, M. , O. D. Duncan, and M. Sobel. 1984. "Association and Heterogeneity." *Sociological Methodology* 14:145-184.

Jasso, Guillermina. 1996. Exploring the Reciprocal Relations between Theoretical and Empirical Work: The Case of the Justice Evaluation Function. *Sociological Methods and Research,* 24 (3): 253-303.

McCutcheon, A. 1987. *Latent Class Analysis*. Newbury Park: Sage.

Muthen, B. O. 1978. "Contributions to Factor Analysis of Dichotomous Variables." *Psychometrika*, 43: 551-560.

Raftery, A. 1986. "Choosing Models for Cross-Classifications." *American Sociological Review* 51:145-146.

Reiser, M. and K. Schuessler. 1991. "A Hierarchy for Some Latent Structure Models." *Sociological Methods and Research* 19: 419-465.

Reilly, Terence; O'Brien, Robert M. 1996. "Identification of Confirmatory Factor Analysis Models of Arbitrary Complexity: The Side-by-Side Rule." *Sociological Methods and Research* 24 (4): 473-491.

Spirtes, Peter; Richardson, Thomas; Meek, Christopher; Scheines, Richard; Glymour, Clark. 1998. Using Path Diagrams as a Structural Equation Modeling Tool, *Sociological Methods and Research*, 27(2): 182-225.

4. Multilevel Models and Contextual Effects (30 units)

Baumert (eds.), Modeling Longitudinal and Multilevel Data. Mahwah, NJ: Lawrence Erlbaum. (5)

Chou, Chih-Ping, Peter M. Bentler, and Mary Ann Pentz. 2000. "A Two-Stage Approach to Multilevel Structural Equation Models: Application to Longitudinal Data." Pp. 33-51 in Todd Little, K.U. Schnabel, and J. (1)

DiPrete, Thomas A. and Jerry D. Forristal. 1994. "Multilevel models: methods and substance." Annual Review of Sociology 20: 331-357. (1)

Gelman, Andrew and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, UK: Cambridge University Press. (5)

Hox, Joop. 2000. "Multilevel analyses of grouped and longitudinal data." Pp. 1532 in Todd Little, K.U. Schnabel, and J. Baumert (eds.), Modeling Longitudinal and Multilevel Data. Mahwah, NJ: Lawrence Erlbaum. (1)

McElreath, Richard. 2018. Statistical Rethinking: A Bayesian Course with Examples in R and Stata. Boca Raton, FL: CRC Press. (5)

Raudenbush, Stephen and Anthony Bryk, 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Second Edition. Thousand Oaks, CA: Sage. (5)

Raudenbush, Stephen W; Sampson, Robert. 1999. "Assessing Direct and Indirect Effects in Multilevel Designs with Latent Variables," Sociological Methods and Research, 128 (2): 123-153. (1)

Robinson, W. S. 1950. "Ecological correlations and the behavior of individuals." American Sociological Review 15:351-7. (1)

Snijders, Tom A. B. and Roel J. Bosker. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2nd ed. Thousand Oaks, C.A.: Sage Publications. (5)

Supplemental Readings

Blalock, H.M. 1984. "Contextual-effects models: theoretical and methodological issues." In R.H. Turner and J.F. Short (eds.), *Annual Review of Sociology* 10: 353-372.

Hannan, Michael T. 1991. *Aggregation and Disaggregation in the Social Sciences*. (Revised Edition) Lexington: D.C. Heath

Iversen, Gudmund R. 1991. *Contextual Analysis*. Newbury Park, CA: Sage.

Mason, William M., G.M. Wong, and B. Entwistle. 1983. "Contextual analysis through the multilevel linear model." Pp. 72-103 in In S. Leinhardt (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Raudenbush, S., C. Johnson, and R. Sampson. 2003. "A Multivariate, Multilevel Rasch Model with Applicaton to Self-Reported Criminal Behavior." *Sociological Methodology*, ed. Ross Stolzenberg.

Raudenbush, Stephen W., and Robert J. Sampson. 1999. "Ecometrics: Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology,* 29: 1-41.

Rovine, Michael J., and Peter C.M. Molenaar. 2001. "A Structural Equations Modeling Approach to the General Linear Mixed Model." Pp. 65-96 in Linda M. Collins and Aline G. Sayer (eds.), *New Methods for the Analysis of Change.* Washington: American Psychological Association.

5. Methods for Longitudinal Data (32 units)

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Los Angeles, C.A.: Sage. (3)

Allison, Paul. D. 2014.. Event History Analysis for Longitudinal Event Data. Newbury Park, CA: Sage. (5)

Barber, Jennifer S., Murphy, Susan A., Axinn, William G., and Jerry Maples. 2000. "Discrete-Time Multilevel Hazard Analysis," Sociological Methodology, Vol. 30. pp. 201-235. (1)

Bollen, Kenneth A. and Jennie E. Brand. 2010. "A General Panel Model with Random and Fixed Effects: A Structural Equations Approach." *Social Forces* 89(1):1–34. (1)

Bollen, Kenneth A. and Patrick J. Curran. 2005. Latent Curve Models: A Structural Equation Perspective. Wiley-Interscience. (5)

Brockwell, Peter J.and Richard A. Davis. 2002. Introduction to Time Series and Forecasting (2nd ed.). N.Y. Springer. (5)

Hosmer, David W., Stanley Lemeshow, and Susanne May. 2008. Applied Survival Analysis: Regression Modeling of Time-to-Event Data, 2nd Edition. Wiley. (5)

Muthen, Bengt. 1997. "Latent Variable Modeling of Longitudinal and Multilevel Data." Sociological Methodology, Vol. 27. pp. 453-480. (1)

Singer, Judith, and John B. Willett. 2003. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press. (5)

Teachman, Jay, and Mark D. Hayward. 1993. "Interpreting Hazard Rate Models." *Sociological Methods and Research*, 21:340-372. (1)

Supplemental Readings

Allgulander, C. and L.D. Fisher. 1986. "Survival Analysis (or time or event analysis), and Cox Regression Models: Methods for the Longitudinal Psychiatric Research." *Acta Psychiatrica Scandinavica*, 74:529-535.

Allison, Paul. 1982. "Discrete-time Methods for the analysis of event histories," Pp. 61-98 in S. Leinhardt (ed.), *Sociological Methodology*, 1982, San Francisco: Jossey-Bass.

Diggle, Petter J., Patrick Heagerty, Kung-Yee Liang, and Scott L. Zeger. 2002. *Analysis of Longitudinal Data*, Second Edition. Oxford: Oxford University Press.

Freedman, D. A., A. Thorton, D., Camburn, D., Alwin, D., and L. YoungDeMarco. 1988. "The Life History Calendar." Pp. 37-67 in C. Clogg, (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Griffin, L. and L. W. Isaac. 1992. "Recursive regression and the historical use of 'time' in time-series analyses of historical processes." *Historical Methods* 25(4):166-179.

Luke, Douglas A.. 1993. "Charting the process of change: a primer on survival analysis." *American Journal of Community Psychology* 21, 2: 203-246.

McArdle, J.J., and Edward Anderson. 1990. "Latent variable growth models for research on aging." Pp. 21-44 in *Handbook of the Psychology of Aging* (3rd edition). J. E. Birren, and K. W. Schaie (eds.). San Diego: Academic Press.

McArdle, J.J. and David Epstein. 1987. "Latent growth curves within developmental structural equation models." *Child Development* 58: 110133.

Ostrom, Charles W., Jr. 1990. *Time-Series Analysis: Regression Techniques*. 2nd edition. Beverly Hills: Sage.

Singer, Judith D. and John Willett. 1991. "Modeling the Days of our Lives: Using Survival Analysis When Designing and Analyzing Longitudinal Studies of Duration and Timing of Events," *Psychological Bulletin* 110 (2): 268-290.

Singer, Judith D. and John B. Willett. 1993. "Its About Time: Using Discrete Time Survival Analysis to Study Duration and the Timing of Events," *The Journal of Educational Statistics* 18 (2): 155-195.

Willett, John B., Judith Singer, and Nina Martin. 1998. "The Design and Analysis of Longitudinal Studies of Development and Psychopathology in Context: Statistical Models and Methodological Recommendations." *Development and Psychopathology* 10: 395-426.

Willett, John B. and A.G. Sayer. 1994. "Using covariance structure analysis to detect correlates and predictors of individual change over time." *Psychological Bulletin* 116 (2): 363-381.

6. Sampling and Inference Issues (20 units)

6(a) Sampling and Design Effects (5 units)

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). Survey methodology. Hoboken, NJ: Wiley. (read chapter 3: Target populations, sampling frames, and coverage error) (1)

Groves, R. M., & Couper, M. P. (1998). Nonresponse in household interview surveys. New York: Wiley. (read pp. 15-46) (1)

Hambleton, R.K., Swaminthan, H. and Rogers, H.J. 1991. Fundamentals of Item Response Theory. Newbury Park: Sage. Pp. 1-6. (1)

Korn, Edward L. and Barry Grabard. 1991. "Epidemiological Studies Utilizing Surveys: Accounting for the Sampling Design." American Journal of Public Health, 81:1166-1173. (1)

Presser, S., Couper, M. P., Lessler, J. T., Martin, E., Rothgeb, J. M., & Singer, E. 2004. Methods for testing and evaluating survey questions. . Public Opinion Quarterly, 68, 109-130. (1)

6(b) Inference and Model Selection (4 units)

Raftery, Adrian E. 1995. "Judging the Meaning and Significance of Statistical Results: Bayesian Model Selection in Social Research," Sociological Methodology, Vol. 25. pp. 111-163. (1)

Raftery, Adrian E. 1999. "Bayes Factors and BIC: Comment on 'A Critique of the Bayesian Information Criterion for Model Selection,'" Sociological Methods and Research, 27, (3): 411-427. (1)

Weakliem, David L.1999. "A Critique of the Bayesian Information Criterion for Model Selection," Sociological Methods and Research, 27 (3): 359-397. (1)

Western, Bruce. 1996. "Vague Theory and Model Uncertainty in Macrosociology." Sociological Methodology, Vol. 26. pp. 165-192. (1)

6(c) Other Issues with Inference (11 units)

Allison, P. (2002). Missing data. Thousand Oaks, CA: Sage.(read pp. 1-12: Introduction; Assumptions; Conventional methods) (1)

Allison, Paul D. 2000. Multiple Imputation for Missing Data: A Cautionary Tale. Sociological Methods and Research, 28 (3): 301-309. (1)

Berk, Richard A., Western, Bruce, and Robert E. Weiss. 1995. "Statistical Inference for Apparent Populations." Sociological Methodology, Vol. 25. pp. 421-458, (1)

Davison, A.C. and D.V. Hinkley. 1997. Bootstrap Methods and Their Application. Cambridge University Press. (chapters 1-5) (3).

Enders, Craig K. 2010. Applied Missing Data Analysis. 1st ed. New York: The Guilford Press. (5)

Supplemental Readings

Abbott, Andrew. 1998. The Causal Devolution. *Sociological Methods and Research*, 127(2): 148-181. (1)

Berk, Richard. 1991. "Toward a Methodology for Mere Mortals." *Sociological Methodology*. 21, 315-324.

Freedman, D.A. "From association to causation: Some remarks on the history of statistics." *Statistical Science*, vol. 14 (1999) pp. 243–58. Reprinted in *Journal de la Société Francaise de Statistique*, vol. 140 (1999) pp. 5–32 and in *Stochastic Musings: Perspectives from the Pioneers of the Late 20th Century.* Lawrence Erlbaum Associates (2003) pp. 45–71. J. Panaretos, ed. (1)

Lieberson, Stanley. 1985. *Making it Count: The Improvement of Social Research and Theory.* Berkeley: University of California Press.

Marini, Margaret Mooney; Singer, Burton, "Causality in the Social Sciences".

Sociological Methodology, 1988, 18, 347-409

Schafer, J.L. 1991. *Analysis of Incomplete Multivariate Data*. Chapman & Hall/CRC.

Taleb, Nassim Nicholas. 2007. *The Black Swan: the Impact of the Highly Improbable.* New York: Random House.

7. Big Data, Online Data and Sampling (15 units)

Baker, Reg et al. 2013. Report of the AAPOR Task Force on Non-Probability Sampling. (5)

O'Neil, Cathy and Rachel Schutt. 2014. Doing Data Science. Cambridge: O'Reilly. (5)

Salganik, Matthew J. 2018. Bit by Bit: Social Research in the Digital Age. Princeton, N.J.: Princeton University Press. (5)