

THEORIES AND METHODOLOGIES OF THE SOCIAL SCIENCES

Columbia University
GR4010, Fall 2017
Thurs, 6:10PM-8:00PM
602 NW CORNER BUILDING

Instructor: Marco Morales
Email: mam2519@columbia.edu
Office: 270 International Affairs Building
Office Hours: Thurs 8-9PM, and by appointment

TA: Ummugul Bezirhan
Email: ub2126@tc.columbia.edu

I. Overview

This course — one of the two foundational courses in the QMSS curriculum — is designed as an in-depth introduction to the social sciences and its methodologies. It is intended to give a broad overview so students can intelligently combine ideas in solving real-world problems.

We will focus on the logic and design of social research, beginning with some concepts and topics common to research across the social sciences. We will later move on to understanding the principles behind an array of methodologies used in the social sciences: causal inference, experimentation, observational studies, formal models, surveys, and applied machine-learning techniques. We will analyse their applications using cases drawn from the research literature.

The focus of this course is not on the techniques themselves — you will have ample opportunity to do that in other courses — but in understanding the logic behind the use of these tools to extract meaningful answers from their applications.

Prerequisites: it is assumed that you have had at least one semester of graduate-level statistics involving linear regression and analysis of variance. Some basic mathematics and algebra will also be assumed.

II. Course Materials

Two texts are intended to be resources for the crafting of your thesis proposal:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008
- Anthony Weston. *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition, 2009

All mandatory readings will be available on Canvas.

Required Readings are mandatory and should be completed before each class.

Topic Readings are also mandatory as they will serve as the foundation to class discussions.

Thesis Readings are intended to help you with the drafting and crafting of the Research Proposal, but will not be discussed in class. Read them in the suggested order/dates.

Complementary Readings are intended to serve as further (and future) references if you ever want to delve deeper on a particular topic, but are not a requirement to this class.

III. Course Dynamics

The class will be a combination of **lectures** and **focused discussion**. We will devote the **first half of the class** to a lecture on the topic assigned for the week, and during the **second half of the class** students will give a 5-minute presentation of one of the **Topic** readings, followed by a class discussion to analyze in detail these applications of the methods we are studying.

Each one of these **Topic Presentations** should not take more than 5 minutes, and should address the following questions: *What is the research question? What is the specific hypothesis under investigation? Can it be falsified? Why was a particular method chosen? What inferences can be drawn from its results? What are its limitations?* Each presentation should be accompanied by a **one-page summary** to be emailed to the entire class at 6PM on the day prior to class.

Before each class, you should have read, thought about and be prepared to discuss all assigned Required and Topic Readings.

Over the course of the semester, you will be required to complete two (2) **Assignments** where you will apply concepts and methods that you have learned in class. Make sure to read the instructions carefully and address all questions. Assignments should be **submitted on Canvas by 6PM on the indicated dates**.

At the end of the course, you will turn in a 12-15 page **Thesis Proposal**. To ensure that you produce a fully fledged product, you will be required to hand in **Thesis Proposal Deliverables** every two weeks. These will constitute the foundational pieces for your proposal. Feedback will be provided on these deliverables where required. Deliverables should be **submitted on Canvas by 6PM on the indicated dates**.

IV. Course requirements

Attendance is expected and reading assignments are to be completed before each session. All written work must be original and produced exclusively for this class. You are expected to follow the University's guidelines for the submission of written work.

The final grade of the course will be based of your **fulfillment** of each of the following requirements:

Assignments (20%): Students must complete a series of assignments where methods from the course are applied. Make sure to submit each one of them by 6PM on the indicated dates.

Class participation (20%): Students are expected to have read all the required readings before class and actively participate in class discussion. Note that you will not obtain this 20% unless you actively participate in class.

Topic presentations (10%): Students will regularly be assigned to present an overview of a "Topics" reading. Make sure to send your one-page summary to the full class by 6PM on the day prior to the class when you are presenting.

Thesis Proposal Deliverables (20%): Students will submit pieces of their proposal every two weeks. Make sure to submit each one of them by 6PM on the indicated dates.

Thesis proposal (30%): Throughout the curse, you will work on a research proposal. This 12-15 page paper will be turned in at the end of the semester.

V. Course Outline

WEEK 1: INTRODUCTION TO THE COURSE

What this course is (and what it is not). Course overview. What is so unique about quantitative methods applied to the social sciences? Why do we need models to understand the world? Why is it useful to have statistical models in the social sciences?

WEEK 2: THE “SCIENCE” OF SOCIAL SCIENCE

What is so scientific about social sciences? A look at “the method”. Inductive v deductive perspectives. Theories, hypothesis and falsifiability. Links to quantitative methods. Mechanisms.

Required Readings:

- Andrew Gelman. Induction and deduction in bayesian data analysis. *Rationality, Markets and Morals*, (2):67–78, 2011
- John Elster. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge University Press, Cambridge, MA, 2007, Ch. 1-2

Thesis Readings:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008, Ch 3
- Antohony Weston. *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition, 2009, Ch I-II

Complementary Readings:

- Karl Popper. *The Logic of Scientific Discovery*. Routledge, New York, NY, 2002[1935]
- Thomas S. Kuhn. *The Structure of Scientific Revolution*. University of Chicago Press, Chicago, IL, fourth edition, 2012
- Gary King, Robert O. Keohane, and Sidney Verba. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ, 1995
- Andrew Gelman and Cosma Rohilla Shalizi. Philosophy and the practice of bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38, 2013
- Paul E. Meehl. Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34:103–115, 1967

WEEKS 3 | 4: CAUSALITY AND CAUSAL INFERENCE (I & II)

Causes of effects or effects of causes? The search for causes: from Aristotle to Fisher. The fundamental problem of causal inference. The Neyman-Rubin model.

Required Readings:

- Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1980
- Donald B. Rubin. Causal inference using potential outcomes: Design, modelling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005

Thesis Readings:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008, Ch 4
- Anthony Weston. *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition, 2009, Ch VI-VII

Complementary Readings:

- Guido Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, MA, 2015
- Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge, MA, second edition, 2009
- Jasjeet S Sekhon. Quality meets quantity: Case studies, conditional probability, and counterfactuals. *Perspectives on Politics*, 2(2):281–293, 2004
- A.P. Dawid. Causal inference without counterfactuals. *Journal of the American Statistical Association*, 95(450):407–448, 2000
- Scott Page. Path dependence. *Quarterly Journal of Political Science*, 1(1):87–115, 2006

[9/21] - Thesis Proposal Deliverable #1 (research topic) due.

[9/28] - Students will receive Assignment #1.

WEEK 5: EXPERIMENTS AND RANDOMIZATION

Theoretical Foundations of Experiments. Statistical foundations of experiments. Taxonomy of randomized experiments. Randomized experiments as the golden standard for causal inference. Inference from randomized experiments.

Required Readings:

- Donald Green and Alan Gerber. The under provision of experiments in political science. *Annals of the American Academy of Political and Social Science*, 589:94–112, 2003
- Joshua D. Angrist and Pischke Jörn-Steffen. *Mastering ‘Metrics’: The Path from Cause to Effect*. Princeton University Press, Princeton, NJ, 2015, Ch 1

Topic: Voter Turnout

- Ana De la O. Do conditional cash transfers affect electoral behavior? evidence from a randomized experiment in Mexico. *American Journal of Political Science*, 57(1): 1–14, 2013
- Leonard Wantchekon. Clientelism and voting behavior: Evidence from a field experiment in Benin. *World Politics*, 55(3):399–422, 2003

Thesis Readings:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008, Ch 15

Complementary Readings:

- Fernando Martel García and Leonard Wantchekon. Theory, external validity, and experimental inference: Some conjectures. *Annals of the American Academy of Political and Social Science*, (628):132–147, 2010
- Joshua D. Angrist and Pischke Jörn-Steffen. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press, Princeton, NJ, 2009
- Guido Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, MA, 2015
- Rebecca Morton and Kenneth Williams. *Experimental Political Science and the Study of Causality*. Cambridge University Press, Cambridge, MA, 2010

[10/5]- Thesis Proposal Deliverable #2 (hypotheses) due.

[10/5] - Assignment #1 due

WEEK 6: OBSERVATIONAL STUDIES

How are they different from experiments? Understanding their limitations for causal inference. Statistical tools to cope with non-random assignment of treatments. Inference from observational studies.

Required Readings:

- Donald B. Rubin. For objective casual inference, design trumps analysis. *Annals of Applied Statistics*, 2(3):808–840, 2008
- Adam Przeworski. Is the science of comparative politics possible? In Carles Boix and Susan C. Stokes, editors, *Oxford Handbook of Comparative Politics*. Oxford University Press, New York, NY, 2007

Topic: Selected statistical applications to observational studies

- Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, 2010
- Robert Erikson and Rocío Titiunik. Using regression discontinuity to uncover the personal incumbency advantage. *Quarterly Journal of Political Science*, 10:101–119, 2015

Complementary Readings:

- William G Cochran. Observational studies. *Observational Studies*, 1(1):126–136, 2015[1972]
- Donald B. Rubin. *Matched Sampling for Causal Effects*. Cambridge University Press, Cambridge, MA, 2006
- Paul R. Rosenbaum. *Observational Studies*. Springer, New York, NY, second edition, 2002
- Kosuke Imai, Gary King, and Elizabeth A. Stuart. Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 171(2):481–502, 2008
- Christopher Winship and Stephen L. Morgan. The estimation of causal effects from observational data. *Annual Review of Sociology*, 25:659–706, 1999
- Jasjeet S Sekhon and Rocío Titiunik. When natural experiments are neither natural nor experiments. *The American Political Science Review*, 106(1):35–57, 2012

WEEK 7: CONCEPTS, MEASUREMENT, AND MEASUREMENT ERROR

Research design and the research question. Measurements as a function of concepts. Theoretical consequences of measurement error. Statistical consequences of measurement error.

Required Readings:

- Barbara Geddes. How the cases you choose affect the answers you get: Selection bias in comparative politics". *Political Analysis*, 2(1):131–152, 1990
- Jerry Hausman. Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(1):57–67, 2001

Topic: Economic Perceptions

- Kristin Michelitch, Marco A. Morales, Joshua Tucker, and Andrew Owen. Looking to the future: Prospective economic voting in 2008 presidential elections. *Electoral Studies*, 31(4):838–851, 2012

Thesis Readings:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008, Ch 7-9, 12

Complementary Readings:

- Duane F. Alvin. *Margins of Error. A Study in Reliability in Survey Measurement*. John Wiley & Sons, Hoboken, NJ, 2007
- Gary Goertz. Concepts, theories, and numbers: A checklist for constructing, evaluating, and using concepts or quantitative measures. In Janet M. Box-Steffensmeier, Henry E. Brady, David Collier, and Gary Goertz, editors, *Oxford Handbook of Political Methodology*. Oxford University Press, New York, NY, 2008
- Simon Jackman. Measurement. In Janet M. Box-Steffensmeier, Henry E. Brady, David Collier, and Gary Goertz, editors, *Oxford Handbook of Political Methodology*. Oxford University Press, New York, NY, 2008

[10/19] - Thesis Proposal Deliverable #3 (data description) due.

WEEK 8: IDENTIFICATION

Graph models and identification. Recursive causation. The empirical - and theoretical - problems of endogeneity. Quantitative methods to address identification. Instrumental variables.

Required Readings:

- Joshua D. Angrist and Pischke Jörn-Steffen. *Mastering 'Metrics': The Path from Cause to Effect*. Princeton University Press, Princeton, NJ, 2015, Ch 3
- Stephen L. Morgan and Christopher Winship. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, Cambridge, MA, 2007, Ch. 3

Topic: Selected statistical applications to address identification

- Edward Miguel, Shanker Satyanath, and Ernest Sergenti. Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4): 725–753, 2004
- Daron Acemoglu, Simon Johnson, and James A. Robinson. The colonial origins of comparative development: An empirical investigation. *The American Economic Review*, 91(5):1369–1401, 2001

Complementary Readings:

- Charles F. Manski. *Identification Problems in the Social Sciences*. Harvard University Press, Boston, MA, 1999
- Joshua D. Angrist and Alan B. Krueger. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4):69–85, 2001
- John Antonakis, Samuel Bendahan, Philippe Jacquart, and Rafael Lalive. Causality and endogeneity: Problems and solutions. In David V Day, editor, *Oxford Handbook of Leadership and Organizations*. Oxford University Press, New York, NY, 2014
- Scott Page. Path dependence. *Quarterly Journal of Political Science*, 1(1):87–115, 2006

WEEKS 9 | 10: SURVEY RESEARCH AND SURVEY METHODOLOGY (I & II)

What is wrong (and what is right) with polls? The Total Survey Error Paradigm. Respondent Selection Issues. Response Accuracy Issues. Survey Administration Issues.

Required Readings:

- René Bautista. An overlooked approach in survey research. total survey error. In Lior Gideon, editor, *Handbook of Survey Methodology for the Social Sciences*. Springer, New York, NY, 2012
- Courtney Kennedy, Mark Blumenthal, Scott Clement, Joshua D. Clinton, Claire Durand, Charles Franklin, Kiley McGeeney, Lee Miringoff, Kristen Olson, Doug Rivers, Lydia Saad, Evans Witt, and Chris Wlezien. An evaluation of 2016 election polls in the united states, 2017. URL <http://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>. AAPOR Ad Hoc Committee on 2016 Election Polling

Topic: Question Wording and Response Scales

- Lundmark. Sebastian, Mikael Gilljam, and Stefan Dahlberg. Measuring generalized trust. an examination of question wording and the number of scale points. *Public Opinion Quarterly*, 80(1):26–43, 2016

Topic: Non-response and Data Quality

- Scott Fricker and Roger Tourangeau. Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public O*, 74(4):934–955, 2010

Topic: Mode of data collection

- Joseph W. Sakshaug, Ting Yan, and Roger Tourangeau. Nonresponse error, measurement error, and mode of data collection. *Public Opinion Quarterly*, 74(5):907–933, 2010

Topic: Probability v Non-probability samples

- David S. Yaeger, Jon A. Krosnick, Linchiat Chang, Harold S. Javitz, Matthew S. Levendusky, Alberto Simpser, and Rui Wang. Comparing the accuracy of rdd telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75(4):709–747, 2011

Thesis Readings:

- Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. *The Craft of Research*. University of Chicago Press, Chicago, IL, third edition, 2008, Ch 13-14, 16-17
- Anthony Weston. *A Rulebook for Arguments*. Hackett Publishing Co, Indianapolis, IN, fourth edition, 2009, Ch VIII

Complementary Readings:

- Herbert Weisberg. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. University of Chicago Press, Chicago, IL, 2005
- Roger Tourangeau, Lance J Rips, and Kenneth Rasinksi. *The Psychology of Survey Response*. Cambridge University Press, Cambridge, MA, 2000
- Ian Foster, Rayid Ghani, Ron S. Jarmin, Frauke Kreuter, and Julia Lane, editors. *Big Data and Social Science: A Practical Guide to Methods and Tools*. Taylor & Francis Group, New York, NY, 2016
- Edith D. Leeuw, Joop J. Hox, and Don A. Dillman. *International Handbook of Survey Methodology*. Routledge, New York, NY, 2008

[11/2] - Thesis Proposal Deliverable #4 (literature review) due.
 [11/9] - Students will receive Assignment #2.

WEEKS 11 | 13: RATIONAL CHOICE AND FORMAL MODELS (I & II)

Formal models of social behavior. The rational choice paradigm. What is rationality? What are its limitations? Can rational choice help explain collective action?

Required Readings:

- William H. Riker. The political psychology of rational choice theory. *Political Psychology*, 1(16):23–44, 1995
- Kenneth Arrow. Methodological individualism and social knowledge. *American Economic Review*, 84(1):1–9, 1994

Topic: Institutionalism

- Adam Przeworski. Democracy as an equilibrium. *Public Choice*, 123:253–273, 2005

Topic: Partisanship and the transmission of partisan attachments

- Christopher Achen. Parental socialization and rational party identification. *Political Behavior*, 24(4):151–170, 2002

Topic: Legislative Behavior

- Keith Krehbiel. *Pivotal Politics: A Theory of U.S. Lawmaking*. University of Chicago Press, Chicago, IL, 1998, Ch 2

Topic: Rational Choice and its critics

- Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979

Complementary Readings:

- Milton Friedman. The methodology of positive economics. In Daniel J. Hausman, editor, *The Philosophy of Economics*. Cambridge University Press, Cambridge, MA, third edition, 2008
- Gary Becker. *The Economic Approach to Human Behavior*. University of Chicago Press, Chicago, IL, 1976
- Amartya Sen. The possibility of social choice. *American Economic Review*, 89(3): 349–378, 1999
- Rebecca Morton. *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science*. Cam, Cambridge, MA, 1999
- Martin Osborne. *An Introduction to Game Theory*. Oxford University Press, New York, NY, 2009
- Daniel Kahneman. *Thinking Fast and Slow*. Farrar, Strauss and Giroux, New York, NY, 2013
- Michael Chwe. *Jane Austen: Game Theorist*. Princeton University Press, Princeton, NJ, 2013

[11/16] - Thesis Proposal Deliverable #5 (research strategy) due.

[11/16] - Assignment #2 due.

[11/30] - Thesis Proposal Deliverable #6 (research proposal draft) due.

WEEK 12: ACADEMIC HOLIDAY

WEEK 14: TEXT AS DATA

Text analysis beyond Natural Language Processing. Statistical and algorithmic analysis of (text) data. Big Data, Machine Learning and Causal Inference.

Required Readings:

- Justin Grimmer and Brandon Stewart. Text as data: The promise and pitfalls of automatic content analysis methods for political documents. *Political Analysis*, 21(3):267–297, 2013

Topic: Estimating Ideology from Texts

- Michael Laver, Kenneth Benoit, and John Garry. Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2):311–331, 2003
- Pablo Barberá. Birds of the same feather tweet together. bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1):76–91, 2015

Complementary Readings:

- Christopher Manning, Prabhakar Raghavan, and Hinrich Schütze. *An Introduction to Information Retrieval*. Cambridge University Press, Cambridge, MA, 2009

[12/15] - FINAL PAPER DUE

Statement on Academic Integrity

Columbia's intellectual community relies on academic integrity and responsibility as the cornerstone of its work. Graduate students are expected to exhibit the highest level of personal and academic honesty as they engage in scholarly discourse and research. In practical terms, you must be responsible for the full and accurate attribution of the ideas of others in all of your research papers and projects; you must be honest when taking your examinations; you must always submit your own work and not that of another student, scholar, or internet source. Graduate students are responsible for knowing and correctly utilizing referencing and bibliographical guidelines. When in doubt, consult your professor. Citation and plagiarism-prevention resources can be found at the GSAS page on Academic Integrity and Responsible Conduct of Research.

Failure to observe these rules of conduct will have serious academic consequences, up to and including dismissal from the university. If a faculty member suspects a breach of academic honesty, appropriate investigative and disciplinary action will be taken following Dean's Discipline procedures.

Statement on Disability Accommodations

If you have been certified by Disability Services (DS) to receive accommodations, please either bring your accommodation letter from DS to your professor's office hours to confirm your accommodation needs, or ask your liaison in GSAS to consult with your professor. If you believe that you may have a disability that requires accommodation, please contact **Disability Services** at 212-854-2388 or disability@columbia.edu.

Important: To request and receive an accommodation you must be certified by DS.