

Spring 2023

PSCI 4991 - Global Development: Intermediate Topics in Politics, Policy and Data

Instructor: Diego Romero

E-mail: djrom18@sas.upenn.edu

Office: PCPSE 350

Meeting time: Tue-Thu, 3:30 to 4:59 PM

Classroom: PCPSE 225

Office hours: By appointment ([sign up here](#)).

Course Description

New sources of data (big data, small data and everything in between) raise the exciting possibility that such data could benefit the world's poor. This course is designed to provide students with an intermediate-level review of recent research that deploys new data for insights on development and hands-on analysis of different kinds of datasets. Students will investigate 3 key themes in development: corruption and accountability, migration, and Authoritarian Backsliding.

As we explore these topics, students will develop data analytical skills that are useful across a wide range of research and real-world applications. Students will learn how to implement four of the most common research designs in the social sciences: observational studies, randomized experiments (in particular, survey experiments), regression discontinuity designs, and difference-in-difference designs. Students will have the opportunity to work with a wide variety of data types, from survey and government-produced administrative data to text as data. As such, students will come face-to-face with the opportunities and challenges for data-intensive approaches to international development.

Course Structure

The course is divided in four parts as follows. Part 1 presents a general introduction to the course and a brief introduction to correlation, causality and statistical inference. The remaining four parts are structured around three key topics in development: migration, corruption and authoritarian backsliding. In every case, we will begin with a brief overview of the literature. Then, we will focus on one particular research question and use it as a guide to learn how to implement four of the most common research designs in the social sciences: observational studies, randomized experiments (in particular, survey experiments), regression discontinuity designs, and difference-in-difference designs. In every case, you will get hands on experience with the kind of data that can best help answer the research question. You will learn how to work with public procurement and elections data, and get an overview of the design, implementation and analysis of a survey. When we move on to working with text as data you will learn basic natural language processing tools to process, visualize and classify newspaper articles.

Learning Objectives

By the end of the semester, students should be able to:

- Apply the techniques for data cleaning, visualization and analysis most appropriate for each type of data.
- Apply basic econometric and statistical tools to implement four of the most common research designs in the social sciences.
- Productively participate in discussions about about the three substantive topics covered in this course.

Prerequisites

1. Students should have taken PSCI 1800 or an equivalent course.
2. Students should be ready to engage with data assignments in the programming language R. Note that, while students who are more comfortable with Python or Stata are also welcomed, all assignments must be completed using R. There are several online resources to learn and master R, such as [this free course](#) or [this overview of R packages for data science](#).

The Penn Program on Opinion Research and Election Studies (PORES) will offer R tutoring hours in case you need extra support. I will update the syllabus with their schedule once it is available.

3. Students will need to have a Dropbox account. Data for replication materials and in-class exercises will be shared through Dropbox.

Textbook

There is no required textbook for this class. All readings (both required as well as suggested) will be available in Canvas. However, here is a list of useful books to complement what you will learn in this class.

- Wickham, Hadley, and Garrett Grolemund. 2017. *R for Data Science: Visualize, Model, Transform, Tidy and Import Data*. O'Reilly Media.
- de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.
- Llaudet, Elena, and Kosuke Imai. 2022. *Data Analysis for Social Science: A Friendly and Practical Introduction*. Princeton University Press.

Expectations and Evaluation

Your performance in this class will be evaluated with 3 different tools: attendance (through 4 random quizzes), data assignments (a total of 3), and a final project. You are expected to do the readings and come to class prepared to comment on, and discuss the material.

- **Attendance** (10% of the grade): There will be 4 random quizzes of a single question each throughout the semester. Attendance will be graded based on your performance in each of these.
- **Data Assignments** (40% of the grade): There will be 3 data assignments throughout the semester. These assignments are designed to give you an opportunity to work directly with administrative data (from public procurement), survey data and newspaper articles. Tasks will vary depending on the type of data, but will always include: cleaning, summarizing and visualization of the data, as well as guided analysis that will enable you to learn how to implement some of the most commonly used research designs in the social sciences. These are individual assignments. While I encourage you to collaborate with your colleagues as you think through the tasks, you are expected to turn in *your own work*. Your code must be appropriately commented and reproducible.
- **Final Project** (50% of the grade): The final project consists on developing a research proposal. This is an individual assignment. In order to make sure everyone stays on track the first milestone for this assignment will consist of a 2 page research proposal draft outlining your research question. Your research proposal must include a hypothesis developed with reference to a concise literature review. This first draft is due by 11 PM on **March 2**. Afterwards, you will need to develop your research design, find the appropriate data to answer your question, and incorporate a research design section as well as a description of your data to your draft. You will present your finished research design in front of the classroom during the last two sessions. There will be a brief feedback session after each presentation. **The final draft of your research proposal will be due on April 28.** Your final draft will be graded for substance, clarity and the degree to which your proposed design can adequately help you answer your research question. Formatting requirements: no more than 6 pages, double spaced, with 1 inch margins and 12-point Times New Roman font.

Grading for this assignment will be as follows:

- First draft (due on March 2): 20 points
- Presentation: 20 points
- Participation in the feedback sessions: 10 points
- Final draft (due on April 28): 50 points

Policies

- *Late submission of assignments*: 2 points will be deducted for every day late, except in documented cases of serious illness or family tragedy.

- *Academic integrity*: I follow the University of Pennsylvania's guidelines on plagiarism: <https://catalog.upenn.edu/pennbook/code-of-academic-integrity/>
- *Laptops in class*: They are a necessary tool for this course.

Important Deadlines

- **February 14** (by 11:00PM): First problem set due.
- **March 2** (by 11:00PM): First draft of the research proposal due.
- **March 17** (by 11:00PM): Second problem set due.
- **April 14** (by 11:00PM): Third problem set due.
- **April 20** and **April 25** (during class): Presentation of the final research proposal.
- **April 28** (by 11:00PM): Research proposal due.

Schedule and Reading Assignments

Part 1 - Introduction

Meeting 1 (January 12): Introduction to the Course

- No required readings.

Meeting 2 (January 17): Correlation and Causation

- **REQUIRED.** (Chapters 2 and 3) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.

Meeting 3 (January 19): Statistical Inference

- **REQUIRED.** (Chapter 6) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.

Part 2 - Migration (Theme 1)

Meeting 4 (January 24): An Overview of Migration I

- **REQUIRED.** Alesina, Alberto, and Marco Tabellini. 2022. “The Political Effects of Immigration: Culture or Economics?.” *Journal of Economic Literature*, Forthcoming.
- **REQUIRED.** Rose, Sarah, Reva Resstack, Helen Dempster, Elisa Cascardi, and Jeremy Weinstein. 2021. “Addressing the “Root Causes” of Irregular Migration from Central America: An Evidence Agenda for USAID.”
- **Suggested.** Fitzgerald, Jennifer, David Leblang, and Jessica C. Teets. 2014. “Defying the law of gravity: The political economy of international migration.” *World Politics*, 66(3): 406-445.
- **Suggested.** Hainmueller, Jens, and Daniel J. Hopkins. 2014. “Public attitudes toward immigration.” *Annual Review of Political Science*, 17: 225–249.

Meeting 5 (January 26): An Overview of Migration II

- **REQUIRED.** (Introduction) Castles, Stephen and Mark J. Miller. 1998. *The age of migration: International population movements in the modern world*. Bloomsbury Publishing.
- **REQUIRED.** Denny, Elaine, David A. Dow, Wayne Pitts, Diego Romero, Juan Tellez, Mateo Villamizar Chaparro, Erik Wibbels, Pamela Zabala. 2021. “The Human Impact of Deportation.” Working paper.

Meeting 6 (January 31): Reasons to Migrate and Observational Studies

- **REQUIRED.** Takenaka, Ayumi, and Karen A. Pren. 2010. Determinants of emigration: Comparing migrants’ selectivity from Peru and Mexico.” *The ANNALS of the American Academy of Political and Social Science*, 630(1): 178-193.
- **REQUIRED.** (Chapter 10) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.
- **Suggested.** Ha, Shang E. 2010. “The consequences of multiracial contexts on public attitudes toward immigration.” *Political Research Quarterly* 63(1): 29-42.
- **Suggested.** Nalepa, Monika. 2021. “Transitional justice and authoritarian backsliding.” *Constitutional Political Economy*, 32(3): 278-300.

Meeting 7 (February 2): Survey Design, Implementation and Analysis - Workshop

- No readings required.

Meeting 8 (February 7): Reasons to Migrate and Randomized Experiments

- **REQUIRED.** (Chapter 11) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.
- **REQUIRED.** Bah, Tijan L; Batista, Catia, Gubert, Flore, and McKenzie, David J. 2022. “Can Information and Alternatives to Irregular Migration Reduce ‘Backway’ Migration from The Gambia?” Policy Research working paper No. 10146; Impact Evaluation series Washington, D.C. : World Bank Group.
- **Suggested.** Meghir, Costas, A. Mushfiq Mobarak, Corina Mommaerts, and Melanie Morten. 2022. “Migration and informal insurance: evidence from a randomized controlled trial and a structural model.” *The Review of Economic Studies* 89(1): 452-480.

Meeting 9 (February 9): Reasons to Migrate and Survey Experiments I

- **REQUIRED.** McKenzie, David, and Melissa Siegel. 2013. “Eliciting illegal migration rates through list randomization.” *Migration Studies* 1(3): 276-291.
- **REQUIRED.** Grady, Christopher. “10 Things to Know About Survey Experiments.” <https://egap.org/resource/10-things-to-know-about-survey-experiments/> (*Focus on Section 2.*)
- **Suggested.** Blair, Graeme, and Kosuke Imai. 2012. “Statistical analysis of list experiments.” *Political Analysis*, 20(1): 47-77.
- **Suggested.** Blair, Graeme, Kosuke Imai, and Jason Lyall. 2014. “Comparing and combining list and endorsement experiments: Evidence from Afghanistan.” *American Journal of Political Science*, 58(4): 1043-1063.
- **Suggested.** Imai, Kosuke. 2011. “Multivariate Regression Analysis for the Item Count Technique.” *Journal of the American Statistical Association*, 106(494): 407-416.
- **Suggested.** Aronow, Peter M., Alexander Coppock, Forrest W. Crawford, and Donald P. Green. 2015. “Combining list experiment and direct question estimates of sensitive behavior prevalence.” *Journal of Survey Statistics and Methodology*, 3(1): 43-66.
- **Suggested.** Creighton, Mathew J., Peter Schmidt, and Diana Zavala-Rojas. 2019. “Race, Wealth and the Masking of Opposition to Immigrants in the Netherlands.” *International Migration* 57(1): 245-263.

Meeting 10 (February 14): Reasons to Migrate and Survey Experiments II

- **REQUIRED.** Detlefsen, Lena, Tobias Heidland, and Claas Schneiderheinze. 2022. “What Explains People’s Migration Aspirations? Experimental Evidence from Sub-Saharan Africa.” Working paper.

- **REQUIRED.** Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2019. “Conjoint Survey Experiments.” In: Druckman, James N., and Donald P. Green, eds. *Cambridge Handbook of Advances in Experimental Political Science*, New York: Cambridge University Press.
- **Suggested.** Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. “Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments.” *Political Analysis*, 22(1): 1–30.
- **Suggested.** Hainmueller, Jens, and Daniel J. Hopkins. 2015. “The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants.” *American Journal of Political Science*, 59(3): 529-548.
- **Suggested.** Brierley, Sarah. 2020. “Unprincipled principals: Co-opted bureaucrats and corruption in Ghana.” *American Journal of Political Science*, 64(2): 209-222.
- **Suggested.** Klačnja, Marko and Lupu, Noam and Tucker, Joshua A. 2021. “When do voters sanction corrupt politicians?” *Journal of Experimental Political Science*, 8(2): 161-171.
- **Suggested.** Breitenstein, Sofia. 2019. “Choosing the crook: A conjoint experiment on voting for corrupt politicians.” *Research & Politics* 6(1): 2053168019832230.
- **Suggested.** Graham, Matthew H., and Milan W. Svolik. 2020. “Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States.” *American Political Science Review*, 114(2): 392-409.
- **Suggested.** Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. “Validating vignette and conjoint survey experiments against real-world behavior.” *Proceedings of the National Academy of Sciences* 112(8): 2395-2400.

Meeting 11 (February 16): Data Assignment 1 Review

- No readings required.

Part 3 - Corruption (Theme 2)

Meeting 12 (February 21): An Overview of Corruption I

- **REQUIRED.** Olken, Benjamin A., and Rohini Pande. 2012. “Corruption in developing countries.” *Annu. Rev. Econ.*, 4(1): 479-509.
- **REQUIRED.** Dimant, Eugen, and Guglielmo Tosato. 2018. “Causes and effects of corruption: what has past decade’s empirical research taught us? A survey.” *Journal of Economic Surveys*, 32(2): 335-356.

Meeting 13 (February 23): An Overview of Corruption II

- **REQUIRED.** (pp. 1-37) Horn, Péter., Czibik, Ágnes., Fazekas, Mihály., and Tóth, Bence. 2021. *Analyzing Public Procurement Risks: Training manual*. Budapest: R2G4P / Government Transparency Institute.
- **REQUIRED.** Fazekas, Mihály; Blum, Jurgen Rene. 2021. “Improving Public Procurement Outcomes : Review of Tools and the State of the Evidence Base.” Policy Research working paper No. WPS 9690 Washington, D.C. : World Bank Group.

Meeting 14 (February 28): Political Favoritism in Public Procurement and Regression Discontinuity Designs (RDD)

- **REQUIRED.** (Chapter 12) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.
- **Suggested.** Titiunik, Rocio, and Matias D. Cattaneo. 2022. “Regression Discontinuity Designs”. *Annual Review of Economics*, 14: 821-851.

Meeting 15 (March 2): Working with Procurement Data - Workshop

- No reading required.

Meeting 16 (March 7)

- No class.

Meeting 17 (March 9)

- No class.

Meeting 18 (March 14): Political Favoritism in Public Procurement - Application I

- **REQUIRED.** Lehne, Jonathan, Jacob N Shapiro and Oliver Vanden Eynde. 2018. “Building connections: Political corruption and road construction in India.” *Journal of Development Economics*, 131:62-78.

Meeting 19 (March 16): Political Favoritism in Public Procurement - Application II

- **REQUIRED.** Gerardino, Maria Paula, Stephan Litschig, and Dina Pomeranz. 2017. “Can audits backfire? Evidence from public procurement in Chile.” NBER Working Papers 23978.

Meeting 20 (March 21): Data Assignment 2 Review

- No readings required.

Part 4 - Corruption (Theme 3)

Meeting 21 (March 23): An Overview of Authoritarian Backsliding I

- **REQUIRED.** Waldner, David, and Ellen Lust. 2018. “Unwelcome change: Coming to terms with democratic backsliding.” *Annual Review of Political Science*, 21(1): 93-113.
- **REQUIRED.** Haggard, Stephan, and Robert Kaufman. 2021. “The anatomy of democratic backsliding.” *Journal of Democracy*, 32(4): 27-41.

Meeting 22 (March 28): An Overview of Authoritarian Backsliding II

- **REQUIRED.** Petrova, Maria. 2011. “Newspapers and parties: How advertising revenues created an independent press.” *American Political Science Review*, 105(4): 790-808.
- **REQUIRED.** Salazar, Grisel. 2019. “Strategic allies and the survival of critical media under repressive conditions: An empirical analysis of local Mexican press.” *The International Journal of Press/Politics*, 24(3): 341-362.

Meeting 23 (March 30): Media Responses to Government Repression and Difference-in-Differences Designs

- **REQUIRED.** (Chapter 13) de Mesquita, Ethan Bueno, and Anthony Fowler. 2021. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press.

Meeting 24 (April 4): Introduction to Working with Text as Data

- **REQUIRED.** (pages 603-611) Chowdhary, K. R. 2020. *Fundamentals of artificial intelligence*. New Delhi: Springer India.

Meeting 25 (April 6): Basic Tasks in Natural Language Processing Workshop (R & Python examples)

- No required readings.

Meeting 26 (April 11): Text Classification (R & Python examples)

- No required readings.

Meeting 27 (April 13): Working with Text as Data

- **REQUIRED.** Adiguzel, Fatih Serkant, Diego Romero, and Erik Wibbels. 2022. “Democratic Backsliding and Media Responses to Government Repression: Machine Learning Evidence from Tanzania.” Working paper.

Meeting 28 (April 18): Data Assignment 3 Review

- No readings required.

Meeting 29 (April 20): Day 1 of Research Proposal Presentations

Meeting 30 (April 25): Day 2 of Research Proposal Presentations

- No required readings.