PPHA 35577: Big Data and Development
Winter 2023

Professor Austin Wright (austinlw@uchicago.edu)
Section Times and Professor Office Hours: See Canvas

Description: *Big Data and Development* is a seminar course focused on the use of innovative data capture and analysis techniques to investigate topics related to economic and political development. Microlevel data is increasingly used to target and evaluate development interventions. In this course, students will engage with cutting-edge theoretical and quantitative research, drawing on readings in economics, political science, and data science. The course is organized around a set of core topics, including political and economic development, community-driven aid interventions, causes and consequences of conflict, and climate change. Course assessments will include three short research briefs and a final research design.

Lectures: This course consists of live lectures of 1 hour and 20 minutes 2 times per week. All lectures will be in person. Lectures will focus on a set of readings each week, typically 2-3 papers (roughly 100 pages), and reading prior to lecture is expected. If a student requires remote accommodations, only Student Disability Services (not your instructor) can approve students to take their courses on a fully remote basis. Please contact Dean of Students Kate Biddle (kbiddle@uchicago.edu) for more information.

Assignments and Grading: The final grade for the course will be a function of three short research briefs and a final research design. The research brief is equivalent to a referee report, where the student writes approximately 500 words evaluating papers along a number of dimensions. Students will sign up for readings to review in the first week, ensuring approximate balance across weeks. The research design will be approximately 4000 words, including a statement of the research question and theoretical argument, brief review of the relevant literature, and a research design, which includes an identification strategy, known and prospective data (including innovative measurement and collection protocols), and preliminary results to assess the feasibility of the identification strategy. **Important** The authors of the top three research designs will receive a research grant to pursue their proposal. Once you have read this note, send me an email with the subject line "Read the syllabus Prof Wright!".

Late assignments/requests for extensions Late assignments will be accepted within two business days with a 25% penalty; other assignments submitted later than two business days will not be accepted. Plan your commitments accordingly.

Stata and R Support: Harris offers free tutoring support to students in need of one-on-one help with their core courses as well as coding in Stata, R, and Python. Tutoring opens on Monday of Week 3 each quarter and students can utilize up 10 hours total of tutoring per quarter. If you would like to learn more about the tutoring program or book an appointment here: `https://canvas.uchicago.edu/courses/36319/pages/core-tutoring-program?module_item_id=1415052`.

Recommended Textbooks: These textbooks are a recommended resource for understanding the identification strategies used in various papers we will study. You likely have these manuscripts from prior courses at Harris. Other course readings, made available via Canvas, will serve as the primary texts.

- *Mostly Harmless Econometrics* by Joshua D. Angrist and Jorn-Steffen Pischke

- *Introductory Econometrics: A Modern Approach* (7th Ed.) by Jeffrey M. Wooldridge

Ethical Academic Conduct: The University's Academic Policies and Procedures and guidance regarding Civil Conduct apply to all activity in our course. If you need to review the University's policies, please see: `https://studentmanual.uchicago.edu/Policies#Honesty` and `https://studentmanual.uchicago.edu/university`. By taking this course, you explicitly pledge your honor that you will not cheat (or help others to cheat) in any way on the assignments. I adhere to the official Harris School protocol for ethical violations: *Harris Procedures for Allegations of Plagiarism, Cheating, and Academic Dishonesty.*

—First Violation

If a student is accused by an instructor or teaching assistant of plagiarism, cheating, or any other form of academic dishonesty, the student will be summoned to meet with the Dean of Students and the instructor. In the meeting, the student and instructor both present information about the situation. If it is determined by the instructor and the Dean of Students that the student has, in fact, plagiarized or cheated, the following sanctions will be imposed for the first violation:

- The student will generally receive a grade of 0 on the assignment or exam in question (subject to the discretion of the instructor). They may be penalized in other ways, up to and including failing the class.

- The student may be asked to re-do the assignment or retake the exam (without credit) to ensure that the student has learned how to properly cite sources or demonstrate that he or she has command of material covered.

- A formal letter of finding is sent to the student stating that the student has been found in violation of the code of academic honesty and what the sanctions were. The letter, along with any evidence presented, is archived in Harris Student Affairs records until the student graduates if the student has no other violations.

—Second Violation

If a student who has already been found in violation academic dishonesty is again accused of academic dishonesty, the case will be sent to the Harris Area Disciplinary Committee. Details about the Area Disciplinary Committee procedures can be found in the University Student Manual (`https://studentmanual.uchicago.edu/area`). If the student is found in violation of academic honesty a second time, the Area Disciplinary Committee can assign sanctions including suspension or expulsion from the University.

Copyrights and Course Content (Use of Course Hero and similar websites): This course is a work of original authorship. All course materials (including, but not limited to, class lectures and discussions, handouts, examinations, study guides and web materials) and the intellectual content of the course itself are protected by United States Federal Copyright Law. Students are permitted to make notes solely for their own private educational use. Students and all other persons are expressly forbidden from recording lectures or discussions and from distributing or selling lectures notes and all other course materials without the prior written permission of the instructors. Because the instructors own the copyright to the classroom presentations and all course materials, any notes taken during those presentations and subsequently sold or distributed to others would constitute an unauthorized derivative work and expose the person or persons involved to individual copyright infringement actions by the instructors.

Course Calendar: The following calendar is meant as a **rough guide**. We will do our best to keep the homework, midterm and final dates unchanged. In terms of lecture material, this is the order of the material, but we expect some content to take longer than one lecture, so the dates may change. Additional readings will be posted on Canvas.

1. **Week of** January 2

    Topic: Using big data to study political and economic development

2. **Week of** January 9

    Topic: Big data and politics: fighting corruption

3. **Week of** January 16

    Topic: Big data and politics: political engagement

4. **Week of** January 23

    Topic: Big data and economics: economic growth

5. Week of January 30

Topic: Big data and economics: targeting aid

6. Week of February 6

Topic: Big data and conflict: countering conflict

7. Week of February 13

Topic: Big data and conflict: population displacement

8. Week of February 20

Topic: Big data and climate change: costs of carbon

9. Week of Feburary 27

Topic: Big data and climate change: political economy of the environment

>**9. Week of** March 6

Final research design submission. Date TBD.