Economics 270D: Research Transparency Methods in the Social Sciences (Ph.D. course)
Professor Edward Miguel (emiguel@berkeley.edu)
SPRING 2015 SYLLABUS

Description: This course covers a range of approaches that aim to enhance the transparency
and reproducibility of social science research. It is appropriate for Ph.D. students in social
science disciplines and related fields.

Prerequisites: Graduate coursework in econometrics and/or statistical methods, and some
experience with statistical software.

Lectures: Tuesday 2-4pm, Moffitt 145

Office hours: Please email Elisa Cascardi (ecascardi@berkeley.edu) for an appointment.

Extra discussion sections and office hours: Garret Christensen (garret@berkeley.edu) and
Felipe Gonzalez (fgonzalez@econ.berkeley.edu) will lead occasional discussion sections
related to class assignments, including on the statistical and software tools used in the course.

Additional materials: For further readings and presentation materials on related topics, refer
to the BITSS 2014 Summer Institute page (http://bitss.org/training/), the slides from the BITSS
Annual Meetings from 2014 (https://osf.io/view/bitss2014/), 2013 (http://bitss.org/annual-
meeting/2013-2/), and 2012 (http://bitss.org/annual-meeting/2012-2/).

The following books may also be useful reference materials:
-- Cooper, H., L. V. Hedges and J. C. Valentine. (eds.). (2009). The handbook of research
Cambridge University Press.

Assignments and Grading:
(Note: assignments should be posted on bCourses by 2pm on the due date.)
(1) Referee reports on working papers (30 percent):
-- Each referee report should be no more than 3 pages double-spaced. The report should start
off with a one paragraph summary of the main argument of the article. You should describe
your main 3-4 points in detail as if you were writing directly to the author. Conclude the report
with more minor comments. A good referee report not only clearly states the shortcomings of
the work, but also lays out constructive, detailed and realistic suggestions for improvement.
-- These assignments will allow you to read and critique papers on the research frontier on
research transparency topics.
-- Report 1 on Abel Brodeur, Mathias Le, Marc Sangnier, and Yanos Zylberberg, “Star Wars: The empirics strike back”, due Tuesday 2/3 [https://sites.google.com/site/abelbrodeur/papers]


* Extra credit option for each referee report: please critically assess the presentation of data or other forms of information in each of the referee report papers. You can receive one point of extra credit for each report if you also include a detailed discussion of the shortcomings of one particular table / figure in the paper, and the concrete modifications you would make to improve the clarity of presentation. (This does not count toward the 3 page limit for the referee report itself.)

(2) Three problem sets (30 percent):
Problem sets require manipulation and analysis of data. I ask you to use either STATA or R, to make the solutions more comparable and to facilitate grading. I will ask you to apply some of the methods and approaches from the course to real data, and ask you to interpret the results. These assignments will provide good practice in applying statistical methods to data, something most of you will need to do in your own research. Leading sources for publicly available social science datasets are the Harvard Dataverse Network [http://thedata.harvard.edu/dvn/] and the University of Michigan Inter-university Consortium for Political and Social Research (ICPSR) [http://www.icpsr.umich.edu/icpsrweb/ICPSR/], but there are many other active data repositories, and some journal websites also host data.

-- Problem set 1, due Tuesday 2/17: Mean effects estimation (based on Kling et al. 2007) and/or multiple hypothesis testing adjustments (based on Anderson 2008) using publicly available data from a published social science research paper.

-- Problem set 2, due Tuesday 3/17: Publication bias, P-curves (based on Simmons et al 2014), and/or meta-analysis (based on Gelman et al 2004) using publicly available data from an existing social science research literature.

-- Problem set 3, due Tuesday 4/14: Apply simple differential privacy techniques (based on Dwork and Smith 2010) to a real dataset (the Netflix Prize data), and identify implications of these techniques for the reliability of standard statistics in the dataset.

(3) One research proposal, 8-9 pages (30 percent):
The research proposal should briefly (3-4 pages) survey the existing literature on a topic related to research transparency and reproducibility, and then describe a planned original research
project (5-6 pages). Proposals should be in 12 point font, double-spaced with 1 inch margins. Proposals exceeding 9 pages in length will lose credit. This assignment will encourage you to generate, refine and receive feedback on a research idea that might form a research article and/or part of your dissertation.
-- The research proposal is due Friday 5/1 at 2pm (uploaded on bCourses).

(4) Class participation (10 percent): In borderline cases, attendance at lecture and the quality of classroom comments may be a factor in assigning grades. Active class participation enriches the course, benefiting yourself and other students (and me), and it is highly encouraged.

(5) There are no exams.

**Special Accommodations:** If you need disability-related accommodations in this class, if you have emergency medical information you wish to share with me, or if you need special arrangements in case the building must be evacuated, please inform me immediately by email or in office hours.

**Honor Code:** The student community at U.C. Berkeley has adopted the following Honor Code: "As a member of the U.C. Berkeley community, I act with honesty, integrity, and respect for others." The hope and expectation is that you will adhere to this code.
COURSE SYLLABUS

I. Introduction to issues of research transparency and reproducibility

Lecture 1: Understanding the problem (1/20)


Lectures 2-3: Publication bias and the file-drawer problem (1/27, 2/3)


**II. Approaches to pre-registration**

**Lectures 4-5: Using pre-analysis plans (2/10, 2/17)**


Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. (2013). "Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration", *Political Analysis*, 21, 1-20.

Additional readings:


**Lecture 6: Transparency in non-experimental research (2/24)**


**Lecture 7: Data Adaptive Pre-specification Approaches (3/3)**


**III. Building scientific knowledge**

**Lecture 8: Approaches to the replication of research (3/10)**


Additional readings:


**Lectures 9-10: Meta-analysis techniques (3/17, 3/31)**


Additional readings:


[SPRING BREAK – 3/24]

IV. Open data and materials

**Lecture 11: What does open data do? (4/7)**


**Lecture 12: Differential privacy and the cost of openness (4/14)**


V. Looking forward

**Lecture 13: Presenting and visualizing data (4/21)**


**Lecture 14: Next steps in changing scientific research practices (4/28) – Guest lecturer, Garret Christensen (U.C. Berkeley)**

