How to Teach Quantitative Methods to Social Science Students: The Princeton Experience

Kosuke Imai

Department of Politics
Center for Statistics and Machine Learning
Princeton University

Annual Meeting of the American Political Science Association

September 1, 2017

Motivation

- Massive technological changes $\rightsquigarrow$ Internet and computing revolution
- **Past**: only statisticians and methodologists analyzed data
- **Today**: EVERYONE is analyzing data

  In God we trust. All others must bring data. — William Deming

- **Past**: government data, national survey data
- **Today**: more of old types of data and lots of new data
  - surveys
  - experiments
  - administrative records
  - social media data
  - GIS data
  - text, images, sounds, videos

- "Big (Social Science) Data" revolution inside and outside the academia
- We must teach students how to analyze data

How Well Are We Teaching? Let's Look at Some Data

- Non-politics introductory quantitative methods courses in social sciences:
  - 5 year average: 2008/09 – 2013/14
  - Economics, Psychology, Sociology, Public Policy

	Lectures	Assignments	Readings	Labs	Overall
Statistics	3.2	3.3	3.1	3.6	3.1
All courses	3.8	3.7	3.7	4.0	3.9

- Politics introductory quantitative methods courses:

	Lectures	Assignments	Readings	Labs	Overall
POL 245	4.4	3.9	3.5	3.9	4.3
POL 345	4.0	3.8	3.7	4.2	4.1

  - Increase in enrollment from 40 to more than 200
  - Enrollment in the third course (POL 346) increased from a single digit to 45
  - More undergraduate students in graduate courses
  - Now offered jointly with the sociology department

Why is Teaching Quantitative Methods Courses So Hard?

1. Students are NOT interested in statistics:

	Professor	Distribution Requirement	Departmental	Certificate Program	General Interest
Statistics	0%	20%	71%	3%	6%
All PU courses	6%	12%	32%	7%	42%

"Professor Imai tried hard to make statistics interesting. But, statistics is boring."

2. Students have weak mathematical and programming background

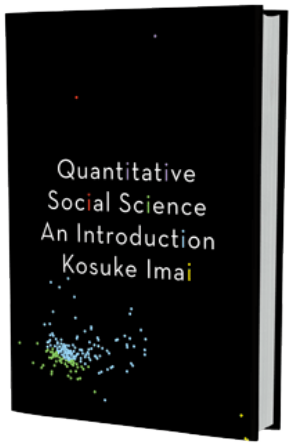
"as a person not naturally inclined towards statistics and probability, I don't feel at all qualified to pass judgement on how the course might have been improved."

New Teaching Strategies

1. Motivating students
   - Data analysis as a necessary tool for social science research
   - Data analysis as a useful skill for post-graduate career
2. Helping students learn efficiently
   - Short but frequent assignments
   - Hands-on instruction in computer labs
   - Outside-of-classroom assistance: online or in-person

Traditional		**New**
paper-and-pencil statistics	$\rightsquigarrow$	data analysis
probability $\longrightarrow$ statistics $\longrightarrow$ data	$\rightsquigarrow$	data $\longrightarrow$ probability $\longrightarrow$ statistics
general $\longrightarrow$ application	$\rightsquigarrow$	application $\longrightarrow$ general $\longrightarrow$ application
toy examples	$\rightsquigarrow$	data from published research
lectures	$\rightsquigarrow$	computer labs
exams	$\rightsquigarrow$	projects

The Textbook — Quantitative Social Science: An Introduction



- Combines three essential components:
  1. social science research
  2. methodological concepts
  3. computer programming (using **R** and **RStudio**)

- Teaches data analysis before statistics:
  1. Introduction
  2. Causality
  3. Measurement
  4. Prediction
  5. Discovery
  6. Probability
  7. Uncertainty
  8. Next

- Contains about 50 data sets from published social science research
  1. Effects of raising minimum wage
  2. Hearts and minds in Afghanistan
  3. Forecasting election outcomes
  4. Who wrote the Federalist papers?
  5. Predicting race from surname
  6. Return to political office

- Additional exercises including **swirl** lessons available

How Much Did Students Learn?

- Final project
  - group project (3 students)
  - start from data collection to data analysis
  - short write-up with 3 figures and 750 words
  - 5 minute presentation followed by Q&A
- Take-home exam
  - students must complete it within a week
  - open book, no collaboration, no assistance
- Electoral effects of Fox News (published in *Quarterly Journal of Economics*)
  1. examining balance of pre-treatment covariates
  2. examining balance using k-means algorithm
  3. recoding of a key variable, before-and-after comparison
  4. difference-in-differences
  5. placebo tests
- Emphasis on interpretation: semi open-ended questions

Other Measures of Success

- High numerical evaluation
- Students' feedback:

  "The course was a lot of fun and really interesting and I plan on taking the next level of the course."

  "I felt it gave me a very true sense of what to expect at Princeton."

- Diverse students in the next level of the course
- Increasing enrollment (over 5 years):
  - introductory course: 40 $\rightsquigarrow$ 230
  - advanced course: 5 $\rightsquigarrow$ 45
  - enrollment in graduate statistics courses
- Increasing use of quantitative methods in junior papers and senior theses
- Research assistantships, top PhD programs

Concluding Remarks

- Technological changes $\rightsquigarrow$ everyone must analyze data!
- paper-and-pencil statistics $\rightsquigarrow$ practical data analysis
- Goal: teach how exciting quantitative social science research is
- Key: use of published research

- *Quantitative Social Science: An Introduction*
  - brings together materials accumulated over years
  - published from Princeton Univ. Press in March 2017
  - early users: American U., Columbia, Dartmouth, Stanford, UCSD, etc.
  - book website with many more applications and contributions from instructors
    `http://qss.princeton.press`