Causality and Statistical Learning

Andrew Gelman
Department of Statistics and Department of Political Science,
Columbia University

27 Mar 2013

1. Different questions, different approaches

- Forward causal inference:
  - What might happen if we do X?
  - Effects of smoking on health
  - Effects of schooling on knowledge
  - Effects of campaigns on election outcomes
- Reverse causal inference:
  - What causes Y?
  - Why do more attractive people earn more money?
  - Why do poor people in India turn out to vote at a higher rate than the middle class and rich?
  - Why are health care costs going up so fast?

Different perspectives on causal inference

- Humans: reverse causal reasoning
- Macro: state-space models
- Applied micro: forward casual inference
- Statisticians: fitting models
- Political scientists: no single dominant framework
- Computer scientists: modeling everyday reasoning (traveling salesman story)

Spectrum of attitudes toward causal reasoning

- (Most conservative) Heckman and Deaton: experiments are no gold standard, you need a substantive model
- Angrist and Pischke, labor economics: identification is all
- Epidemiologists: causal inference from observational data using statistical models
- Social psychologists: structural equation models
- (most permissive) Cognitive scientists: causal *structure* can be estimated from purely observational data

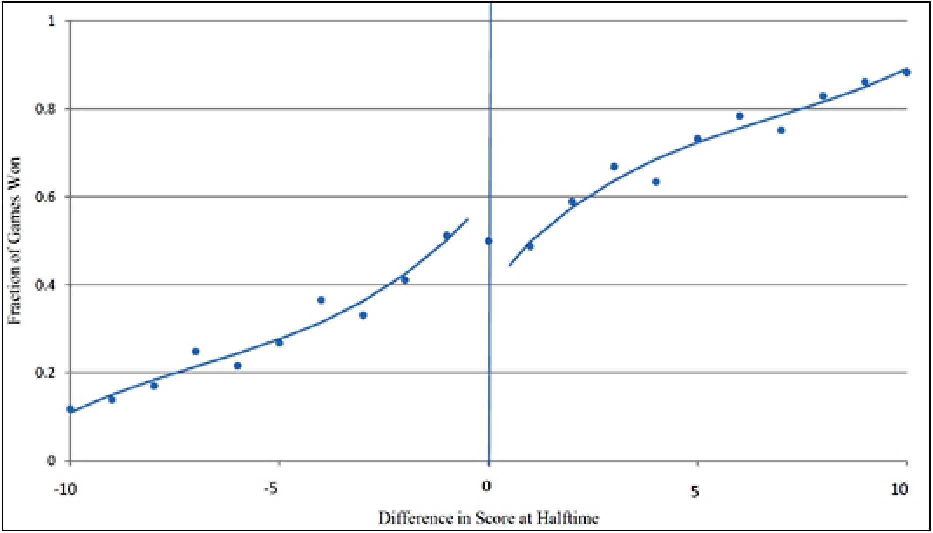
2. Understanding natural experiments

- Example: Levitt study of policing and crime rates
- In cities with mayoral election years:
  - More cops on the street
  - Crime rate goes down
- Can interpret the joint outcome without worrying about instrumental-variables assumptions

Halftime motivation in basketball

- Economists Jonah Berger and Devin Pope:
  "Analysis of over 6,000 collegiate basketball games illustrates that being slightly behind increases a team's chance of winning. Teams behind by a point at halftime, for example, actually win more often than teams ahead by one. This increase is between 5.5 and 7.7 percentage points . . ."
- But . . . in their data, teams that were behind at halftime by 1 point won 51.3% of the time
- Approx 600 such games; thus, std. error is $0.5/\sqrt{600} = 0.02$
- Estimate ± 1 se is $[0.513 \pm 0.02] = [0.49, 0.53]$
- So where did they get "5.5 and 7.7 percentage points"??

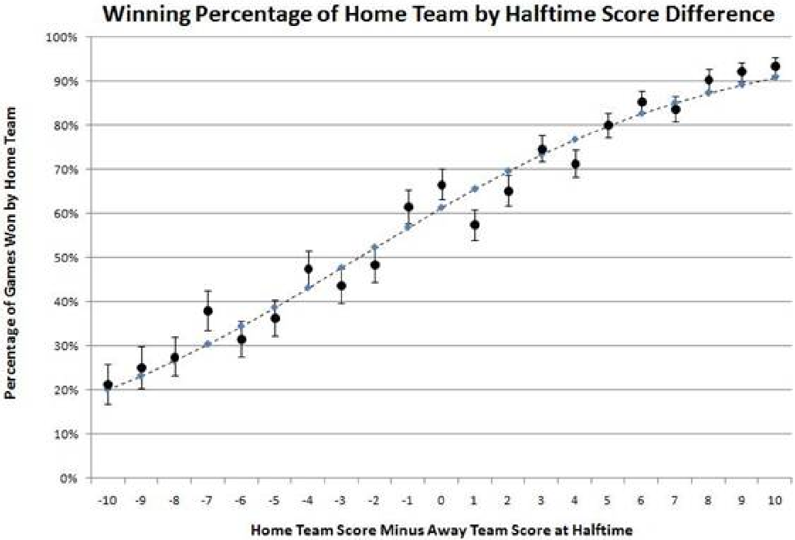
Halftime motivation in basketball: the data



Methodological conservatism

- What about that 5th-degree polynomial?
- Berger and Pope write:
  "While the regression discontinuity methods we use in the paper (including the 5th degree polynomial) are standard in economics (see for example the 2009 working paper on R&D implementation by David Lee and Thomas Lemiuex) we respect that your may find a different approach to the problem to be more useful. . . ."

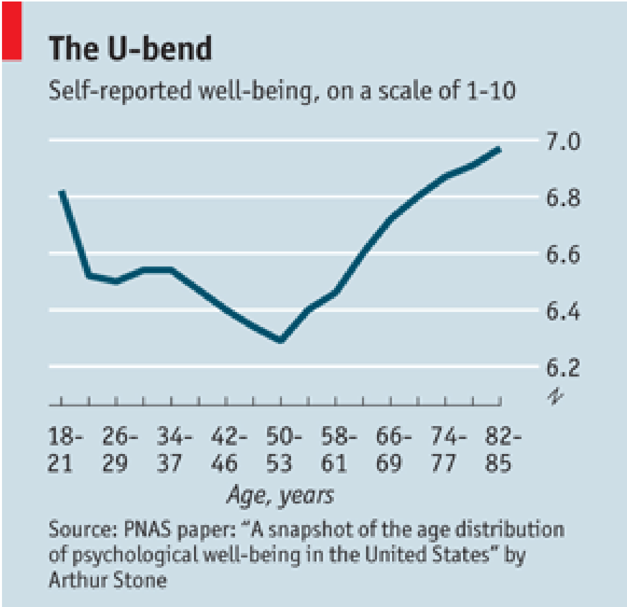
The data without the 5th-degree polynomial



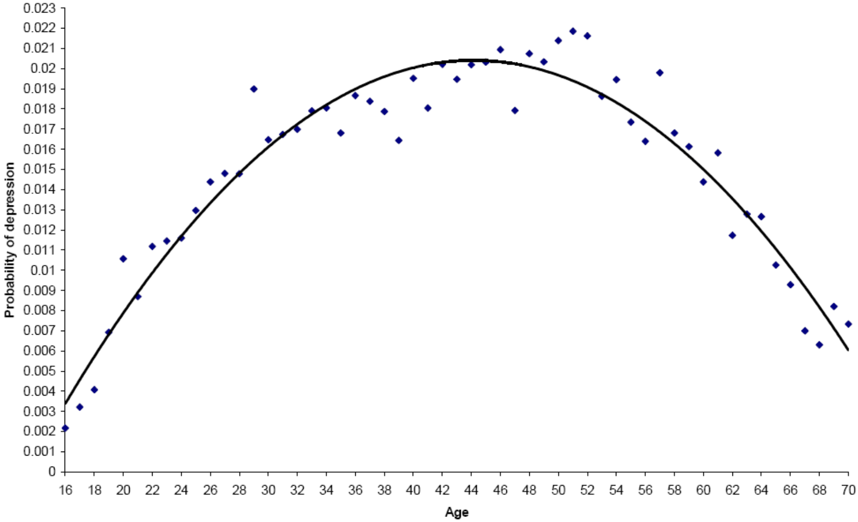
3. The importance of data and measurement

- A key principle in applied statistics is that you should be able to connect between data, model, methods, and conclusions

Age and happiness

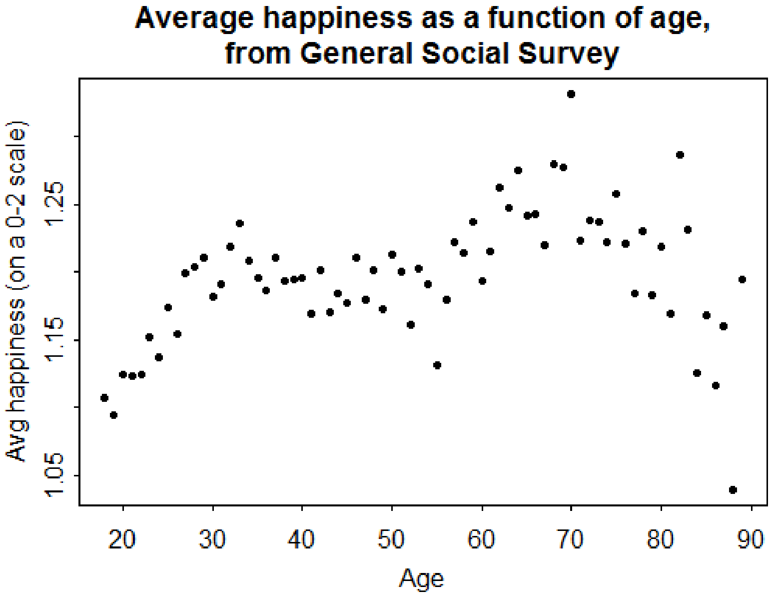


Data!



Probability of depression
0.023
0.022
0.021
0.02
0.019
0.018
0.017
0.016
0.015
0.014
0.013
0.012
0.011
0.01
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70
Age

More data



Average happiness as a function of age,
from General Social Survey
Avg happiness (on a 0-2 scale)
1.05
1.15
1.25
20
30
40
50
60
70
80
90
Age

The perils of pooling

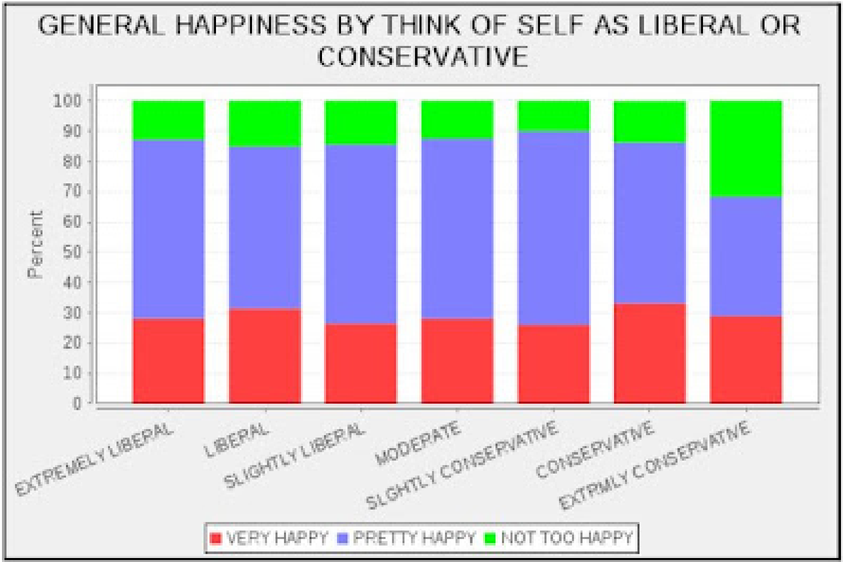
Arthur "not David" Brooks in the *New York Times*:

> *"People at the extremes are happier than political moderates. . . . none, it seems, are happier than the Tea Partiers . . . "*

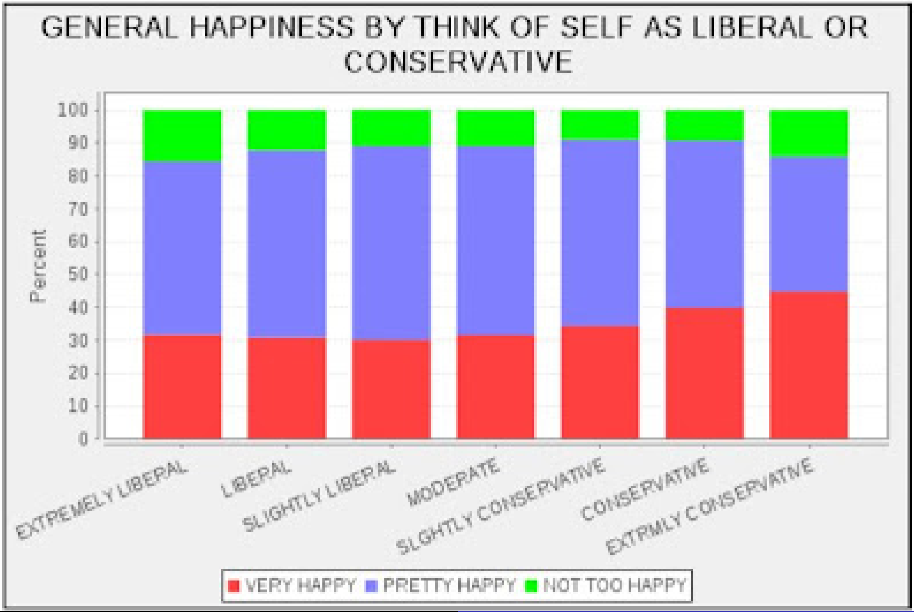
Jay Livingston (sociology, Montclair State University) looks up the data in the General Social Survey . . .

"None, it seems, are happier than the Tea Partiers . . ." ??

Chart for YEAR = 4(2009-2010)



Pooling, 1972–2010



4. Difficulties with the research program of learning causal structure

- For example: income, religion, religious attendance, and vote choice in different regions of the country
- No true zeros
- I respect that some social scientists find it useful to frame their research in terms of conditional independence and the testing of null effects, but I don't generally find this approach helpful—and I certainly dont believe it *necessary* to think in terms of conditional independence in order to study causality

Challenges of causal reasoning are not going away

From a recent book by a cognitive scientist:

If two of the variables are dependent, say, intelligence and socioeconomic status, but conditionally independent given the third variable [beer consumption], then either they are related by one of two chains:

(Intelligence → Amount of beer consumed → Socioeconomic status)

(Socioeconomic status → Amount of beer consumed → Intelligence)

or by a fork:

$$\begin{tikzcd}[column sep=small, row sep=tiny] & \text{Socioeconomic status} \\ \text{Amount of beer consumed} \arrow[ur] \arrow[dr] & \\ & \text{Intelligence} \end{tikzcd}$$

and then we must use some other means [other than observational data] to decide between these three possibilities. In some cases, common sense may be sufficient, but we can also, if necessary, run an experiment. If we intervene and vary the amount of beer consumed and see that we affect intelligence, that implies that the second or third model is possible; the first one is not. Or course, all this assumes that there aren't other variables mediating between the ones shown that provide alternative explanations of the depen-

The problem understanding the world using "stylized facts"

- Problems with is-it-there-or-is-it-not models of correlations and effects
- Problems with the concept of "false positives"
- Accepting variation (as distinct from measurement error)
- Don't fool yourself!

Our brains can do causal inference, so why can't social scientists?

- Humans do (model-based) everyday causal inference all the time
- We rarely use experimental data, certainly not the double-blind stuff that is considered the gold standard
- But . . .
  - The sorts of inferences used as examples by the proponents of "everyday causal reasoning" look much less precise than the sorts of inferences we demand in science (or even social science).
  - Also, everyday causal reasoning is not purely observational
  - We use informal experimentation in our ordinary lives is to resolve some of the causal questions left open by models and observational inference

5. Story time

- When the data go to bed, the stories come out . . .
- Ole Rogeberg:

> *The puzzle that we try to explain is this frequent disconnect between high-quality, sophisticated work in some dimensions, and almost incompetently argued claims about the real world on the other*

"A Raise Won't Make You Work Harder"

- Economist Ray Fisman writing in *Slate*:
  - Students were employed in a six-hour data-entry job for \$12/hour. Half the students were actually paid this amount. The other half were paid \$20/hour.
  - At first, the \$20-per-hour employees were more productive than the \$12-an-hour employees. But by the end the two groups were working at the same pace.
- Conclusions:
  - "The goodwill of high wages took less than three hours to evaporate completely—hardly a prescription for boosting long-term productivity."
  - "A raise won't make you work harder."
- Conflict between internal and external validity:
  - "All participants were told that this was a one-time job—otherwise, the higher-paid group might work harder in hopes of securing another overpaying library gig."

6. A Bayesian view of forward and reverse causal inference

- Forward causal inference = predictions from a model
- Reverse causal inference = posterior predictive checking
- Forward causal inference supplies *answers*
- Reverse causal inference supplies *questions*

Summary 1: Perspectives

- Controlled experiments are the gold standard, but I never do them!
- (Some) computer scientists' view: we don't need controlled experiments; we can automatically learn from observational data
- Psychologists' view: each causal question requires its own experiment
- Observational scientist's version: each causal question requires its own data analysis

Summary 2: Working together

- Experimenters can learn from:
  - Sample surveys (for the problem of extending from sample to population)
  - Descriptive observational research (for the problem of modeling complex interactions and response surfaces)
- Observational researchers (i.e., most empirical social scientists, including me) should model our biases and connect our work to experimental research wherever possible