# Of beauty, sex, and power: Statistical challenges in estimating small effects

Andrew Gelman[1]
Department of Statistics and Department of Political Science
Columbia University

Columbia University Business School, 16 Apr 2013

[1]Collaborators on these projects include David Weakliem, David Park, Boris Shor, Yu-Sung Su, and Daniel Lee

# Millions of scientific papers published every year . . .

# Each paper has its own story . . .

## Journal's Paper on ESP Expected to Prompt Outrage

By BENEDICT CAREY
Published: January 5, 2011

One of psychology's most respected journals has agreed to publish a paper presenting what its author describes as strong evidence for extrasensory perception, the ability to sense future events.

Enlarge This Image

Heather Ainsworth for The New York Times

Work by Daryl J. Bem on extrasensory perception is scheduled to be published this year.

The decision may delight believers in so-called paranormal events, but it is already mortifying scientists. Advance copies of the paper, to be published this year in The Journal of Personality and Social Psychology, have circulated widely among psychological researchers in recent weeks and have generated a mixture of amusement and scorn.

The paper describes nine unusual lab experiments performed over the past decade by its author, Daryl J. Bem, an emeritus professor at Cornell, testing the ability of college students to accurately sense random events,

# Under pressure . . .

Enlarge This Image

Joris Buijs/Pve

The psychologist, Diederik Stapel
committed academic fraud in "se
papers, many accepted in respect
in the news media, according to a
Monday by the three Dutch institu
worked: the University of Groning
Amsterdam, and Tilburg. The jou
published one of Dr. Stapel's pap
"editorial expression of concern"
online on Tuesday.

# Under pressure . . .

## Marc Hauser Resigns From Harvard

*By Tom Bartlett*

Marc D. Hauser, the Harvard psych
found responsible for eight counts
misconduct by the university, has r
ending speculation about whether
embattled professor would return t
this fall.

In a letter dated July 7, Mr. Hauser
Michael D. Smith, Harvard's dean o

# It's not just the silly stuff . . .

SPECIAL ARTICLE

## The Spread of Obesity in a Large Social Network over 32 Years

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

N Engl J Med 2007; 357:370-379 | July 26, 2007

Abstract | Article | References | Citing Articles (405) | Glossary | Letters

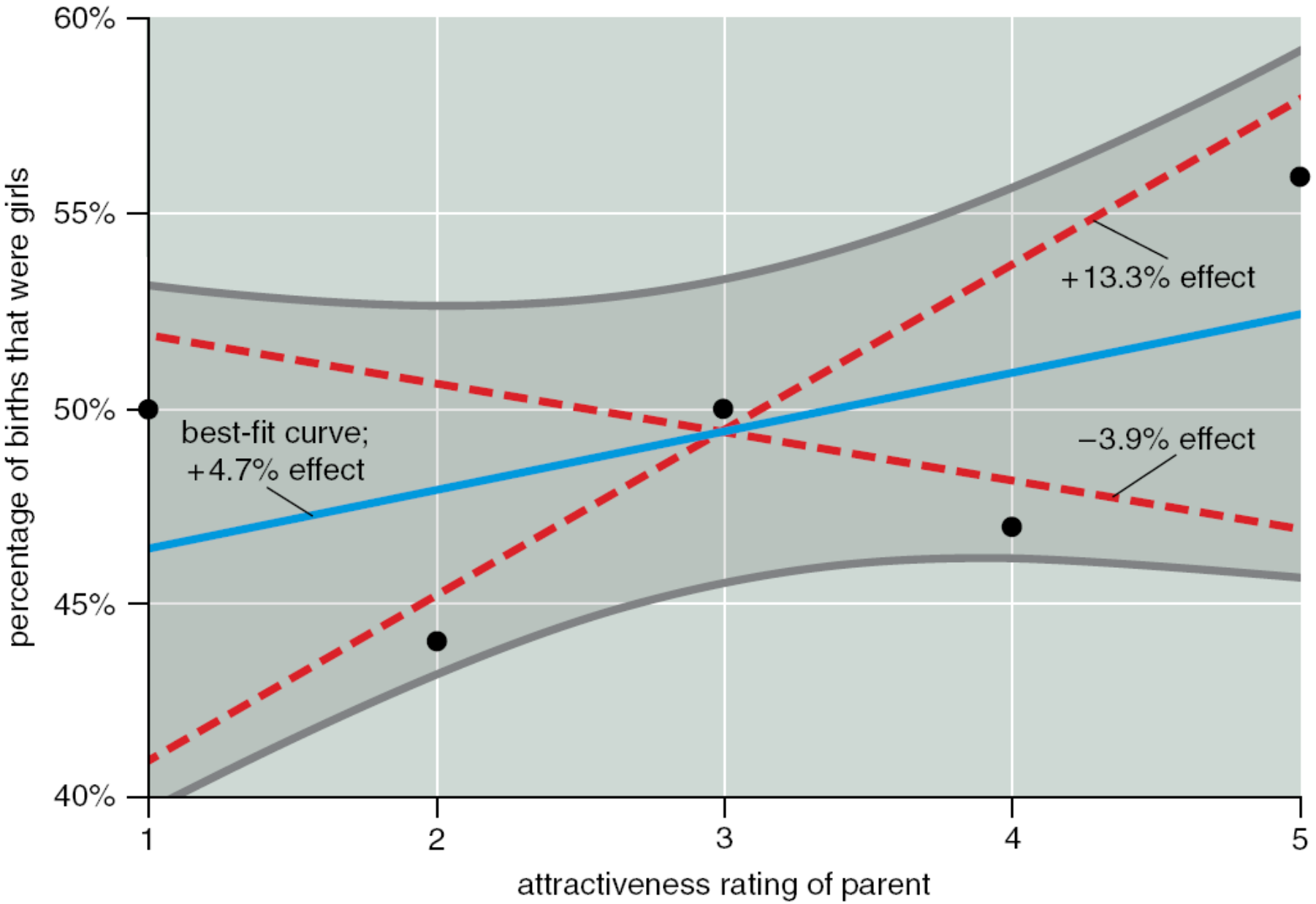
### BACKGROUND

The prevalence of obesity has increased substantially over the past 30 years. We performed a quantitative analysis of the nature and extent of the person-to-person spread of obesity as a possible factor contributing to the obesity epidemic.

# Beautiful parents have more daughters?

- S. Kanazawa (2007). Beautiful parents have more daughters: a further implication of the generalized Trivers-Willard hypothesis. *Journal of Theoretical Biology*.
- Attractiveness was measured on a 1–5 scale ("very unattractive" to "very attractive")
  - 56% of children of parents in category 5 were girls
  - 48% of children of parents in categories 1–4 were girls
- Statistically significant (2.44 s.e.'s from zero, $p = 1.5\%$)
- But the simple regression of sex ratio on attractiveness is not significant (estimate is 1.5 with s.e. of 1.4)
- Multiple comparisons problem: 5 natural comparisons $\times$ 4 possible time summaries!

# The data and fitted regression line



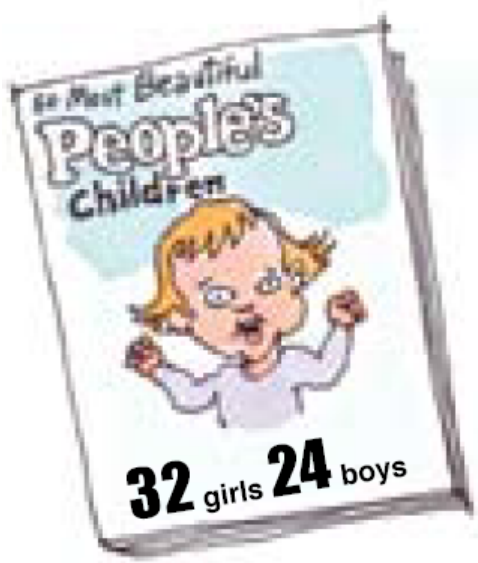
# Background on sex ratios

- Pr (boy birth) $\approx$ 51.5%
  - Boys die at a higher rate than girls
  - At age 20, the number of boys and girls is about the same
  - Evolutionary story
- What can affect Pr (boy births)?
  - Race, parental age, birth order, maternal weight, season of birth: effects of about 1% or less
  - Extreme poverty and famine: effects as high as 3%
- We expect any effects of beauty to be less than 1%

# Interpreting the Kanazawa study

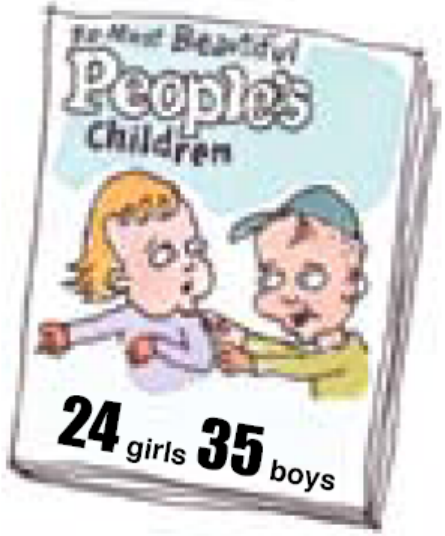
- Data are consistent with effects ranging from $-4\%$ to $+13.3\%$
- More plausibly, consistent with effects less than 0.5% (in either direction!)
- You can take the evolutionary argument in either direction:
  - Beauty is more useful for women than for men, selection pressure, ...
  - Assessed "beauty" is associated with wealthy, dominant ethnic groups who have more power, a trait that is more useful for men than for women, ...
- Results are "more 'vampirical' than 'empirical'—unable to be killed by mere evidence" (Freese, 2007)
- Bottom line
  - Beautiful parents *in this one survey* have more daughters
  - Can't say much about the general population

# Another try: data from *People* magazine

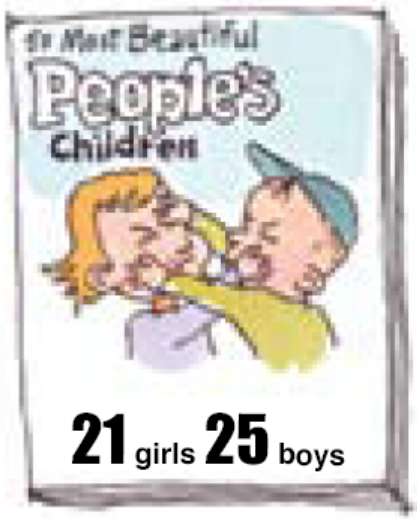



1995


1996

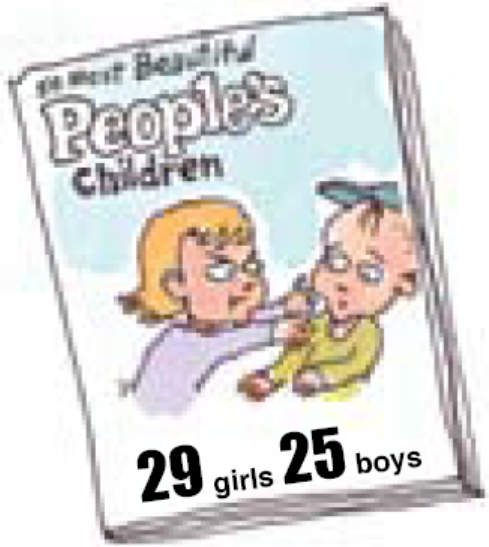



1997




1998


1999




2000

# The children of each year's "50 most beautiful people"

- We collected data from 1995–2000
- 1995: 32 girls and 24 boys: 57.1% girls (standard error 8.6)
- 1996: 45 girls and 35 boys: $56.2\% \pm 7.8\%$
- 1995 + 1996: $56.6\% \pm 4.3\%$: almost statistically significant!
- 1997: 24 girls and 35 boys, . . .
- Pooling 1995–2000: $47.7\% \pm 2.8\%$: not statistically significantly different from 48.5%

# Statistical inference for small effects

- Estimated effect of 4.7 percentage points (with standard error of 4.3):
  - 95% confidence interval is $[-4\%, 13\%]$
  - Given that true effect is most likely below 1%, the study provides essentially *no information*
- Theoretical analysis
  - Suppose the true effect was 0.3% and we gather data on 3000 people
  - 3% probability of a statistically-significant positive result
  - 2% probability of a statistically-significant *negative* result

# Which headline sells more papers?

The Daily W

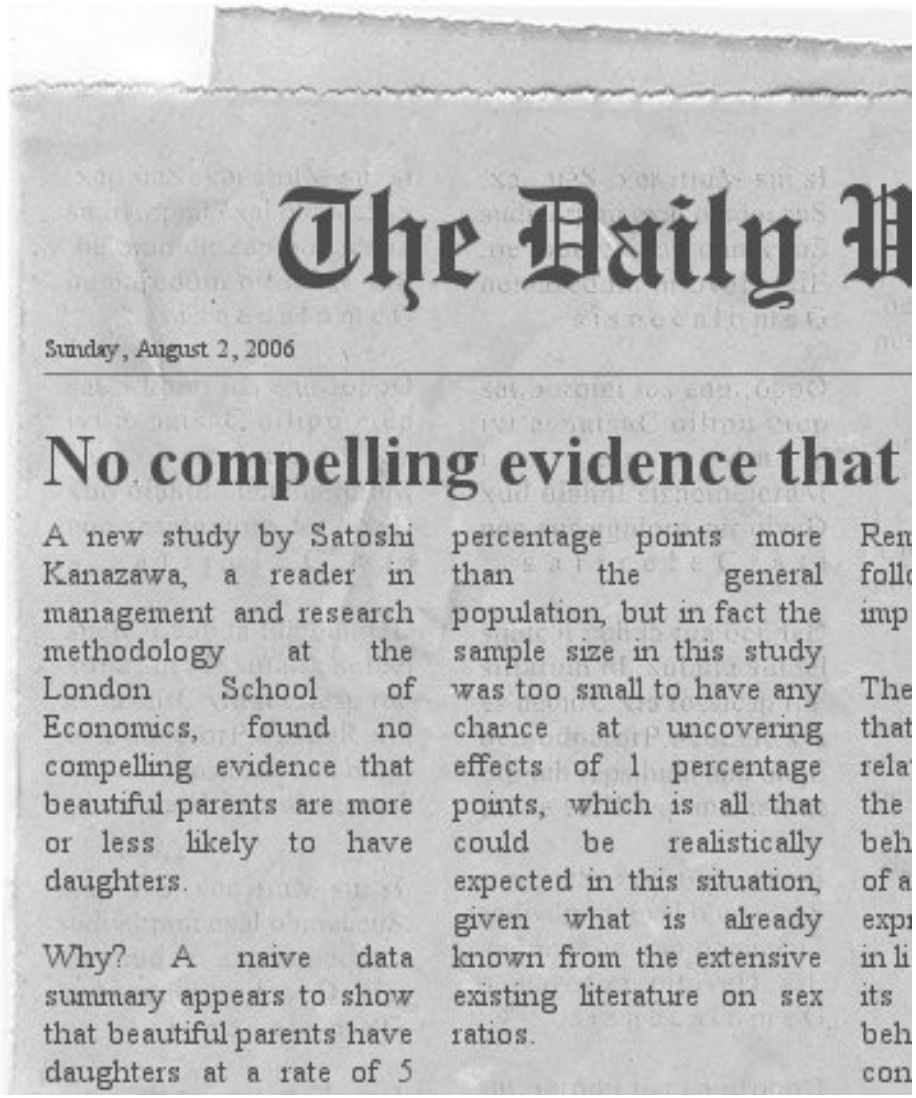
Sunday, August 2, 2006

## Beautiful parents have more

A new study by Satoshi Kanazawa, an evolutionary psychologist at the London School of Economics, suggests there are more beautiful women in the world than there are handsome men.

Why? Kanazawa argues it¥s because good-looking parents are 36 percent more likely to have a baby daughter as their first child than a baby sonswhich suggests, evolutionarily speaking, that beauty is a trait more valuable for women than for men.

The study was conducted with data from 3,000 Americans, derived from the National Longitudinal Study of Adolescent Health, and was published in the Journal of Theoretical Biology.



The Daily W

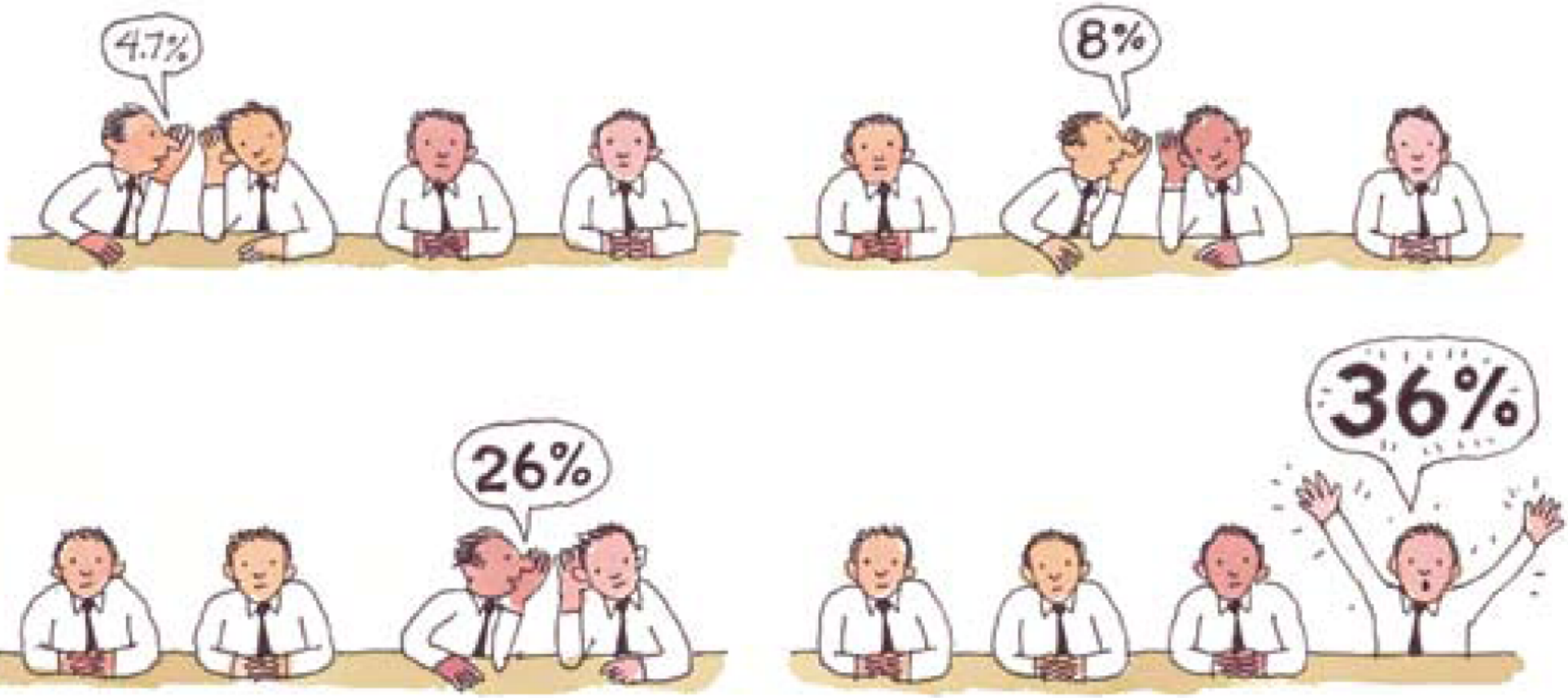
Sunday, August 2, 2006

## No compelling evidence that

A new study by Satoshi Kanazawa, a reader in management and research methodology at the London School of Economics, found no compelling evidence that beautiful parents are more or less likely to have daughters.

Why? A naive data summary appears to show that beautiful parents have daughters at a rate of 5 percentage points more than the general population, but in fact the sample size in this study was too small to have any chance at uncovering effects of 1 percentage points, which is all that could be realistically expected in this situation, given what is already known from the extensive existing literature on sex ratios.

# Communication of the findings



# How to evaluate such claims?

- From the *Freakonomics* blog:
  - "A new study by Satoshi Kanazawa, an evolutionary psychologist at the London School of Economics, suggests . . . there are more beautiful women in the world than there are handsome men. Why? Kanazawa argues it's because good-looking parents are 36 percent more likely to have a baby daughter as their first child than a baby son—which suggests, evolutionarily speaking, that beauty is a trait more valuable for women than for men. The study was conducted with data from 3,000 Americans, derived from the National Longitudinal Study of Adolescent Health, and was published in the Journal of Theoretical Biology."
- If Steven Levitt can't get this right, who can??

# My reaction

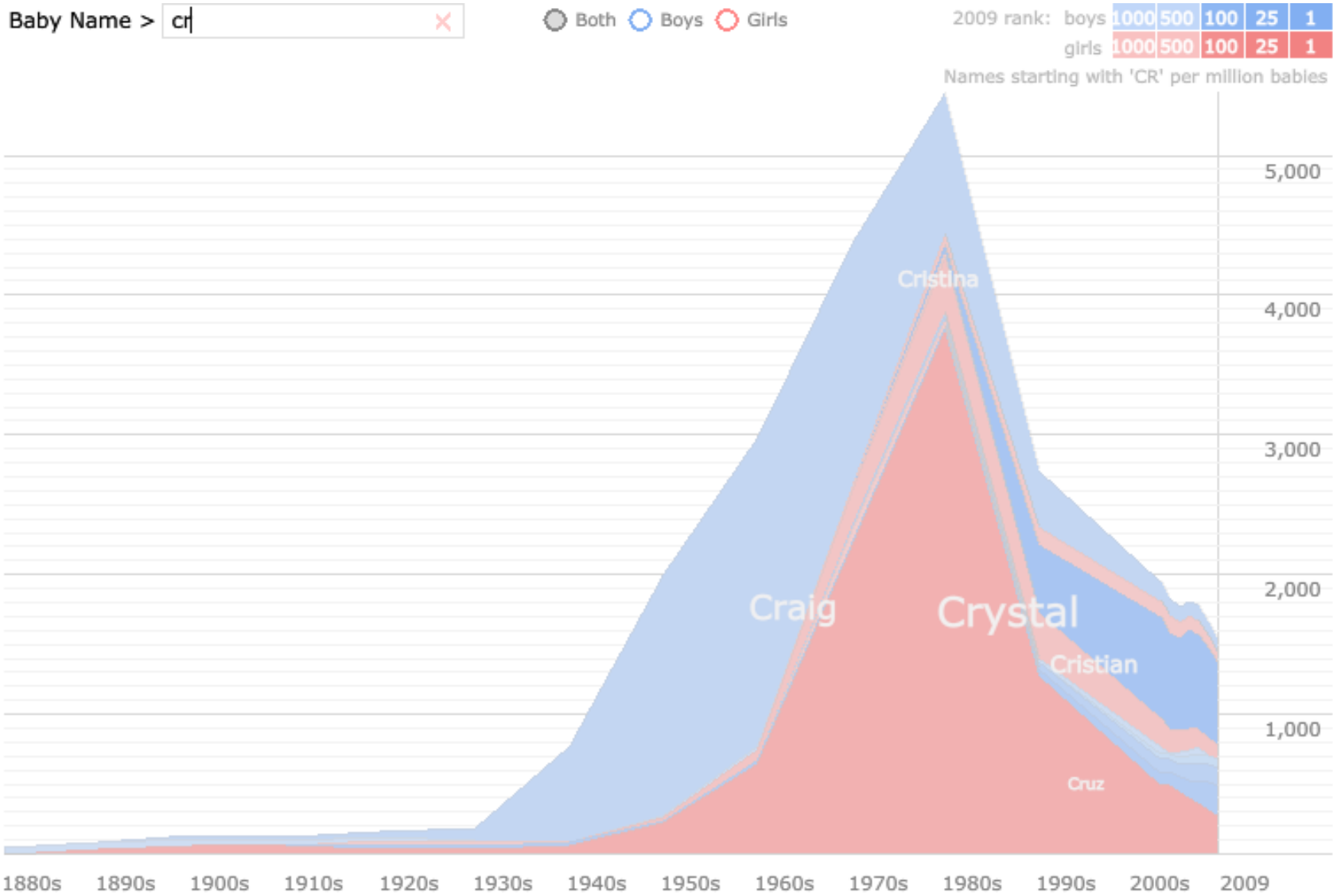
- The claim of "36%" raised suspicion
  - 10 to 100 times larger than reported sex-ratio effects in the literature
- An avoidable error:
  - Small sample size . . .
  - Standard error of 4.3 percentage points . . .
  - To be "statistically significant," the estimate must be at least 2 standard errors away from 0 . . .
  - Any statistically significant finding is *necessarily* a huge overestimate!

# Why is this not obvious?

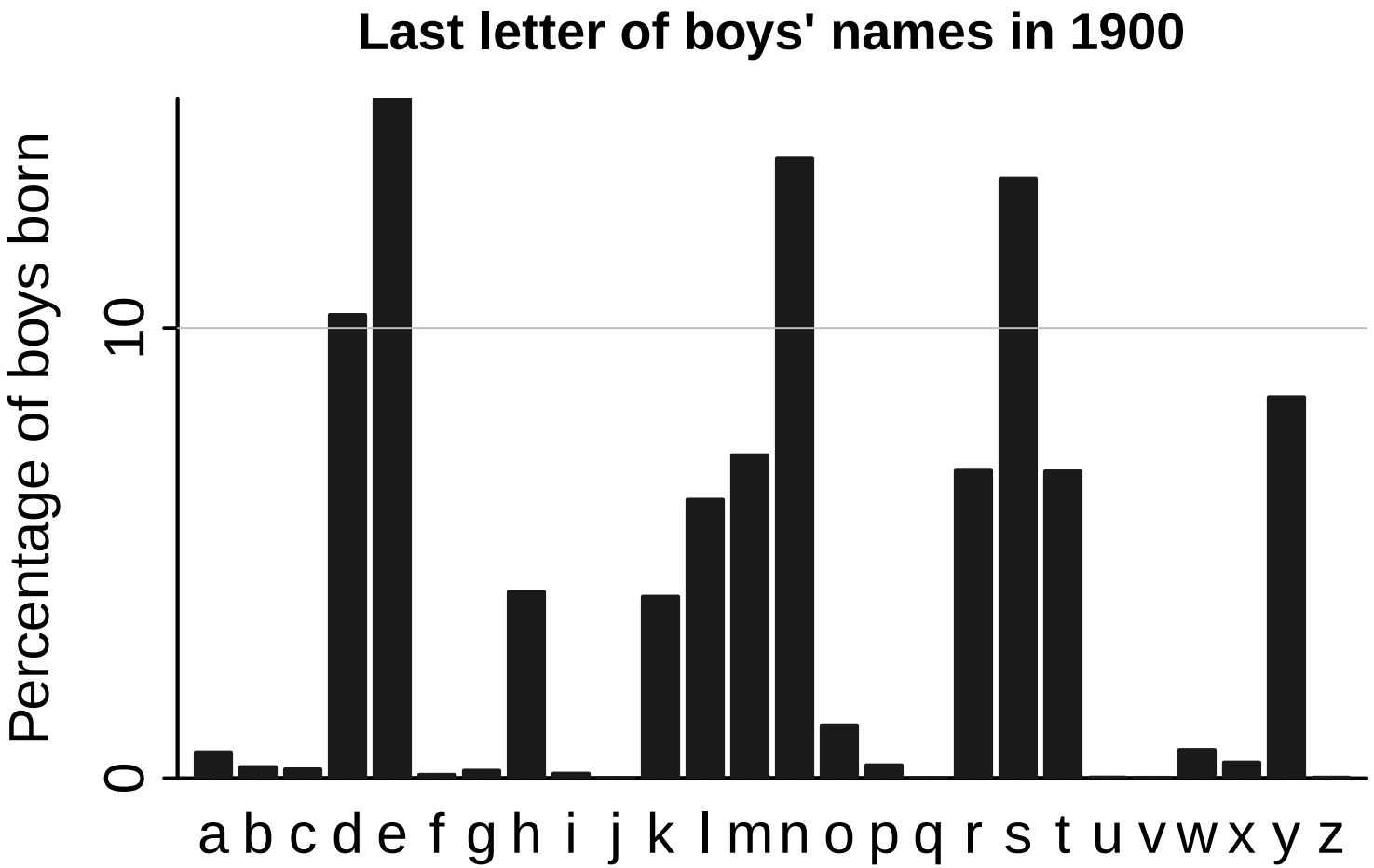
- Statistical theory and education are focused on estimating one effect at a time
- "Statistical significance" is a useful idea, but it doesn't work when studying very small effects
- Methods exist for including prior knowledge of effect sizes, but these methods are not well integrated into statistical practice

# Not all effects are small!

Laura and Martin Wattenberg's Baby Name Wizard:

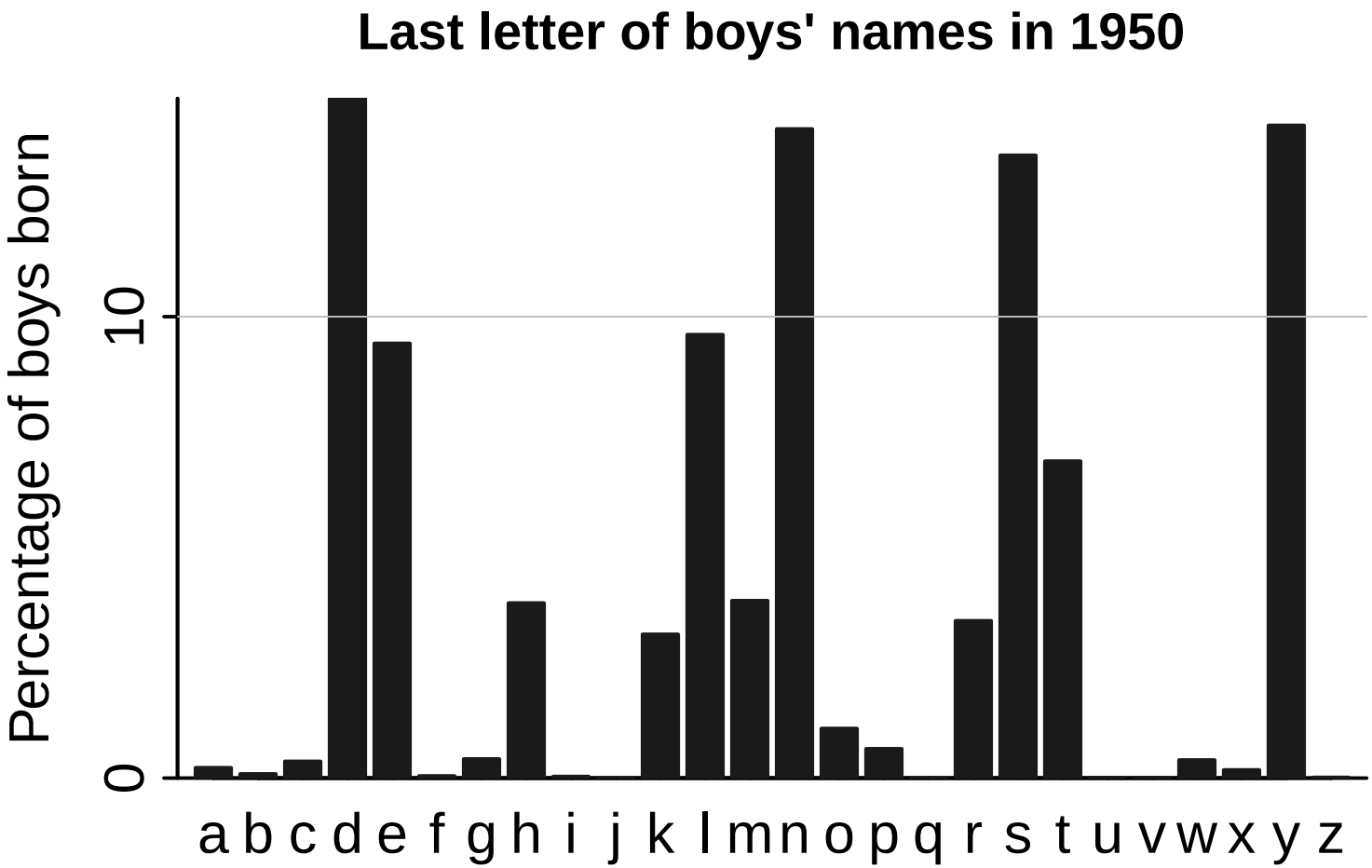


# Last letters of boys' names, 1900



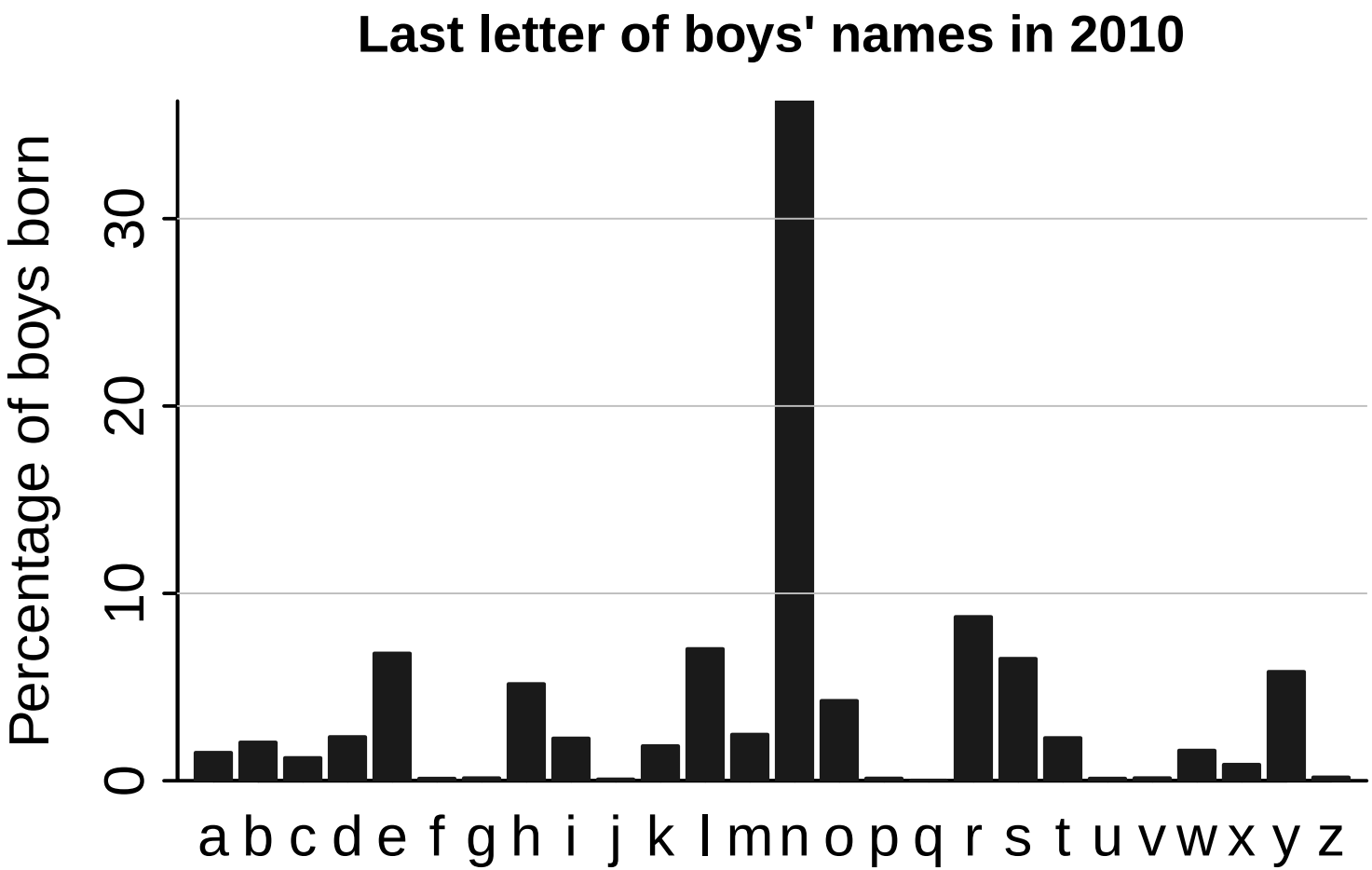

John, James, George, Charles, Edward, ...

# Last letters of boys' names, 1950




Michael, David, Thomas, Larry, ...

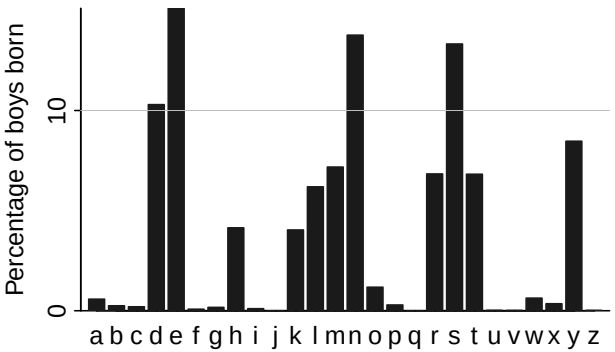
# Last letters of boys' names, now



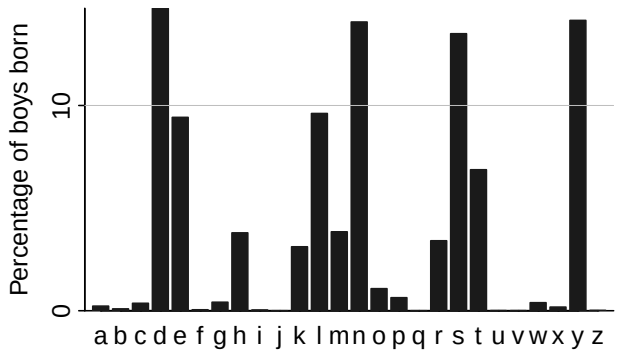

Ethan (#2), Jayden (4), Aiden (9), Mason (12), Logan (17), Benjamin (22), Ryan (23), Jackson (25), John (26), Nathan (27), Jonathan (28), Christian (29), 31, 32, 36, 37, 40, ...

# The trend in last letters of boys' names

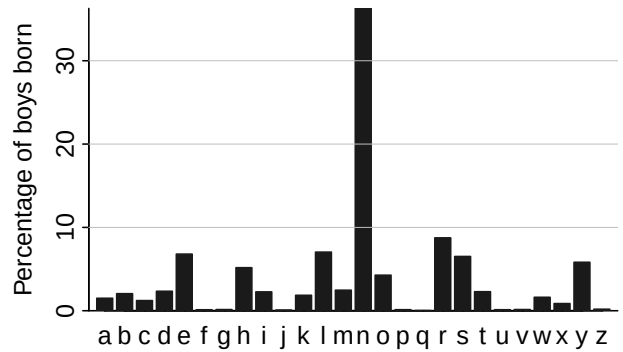
1900




1950



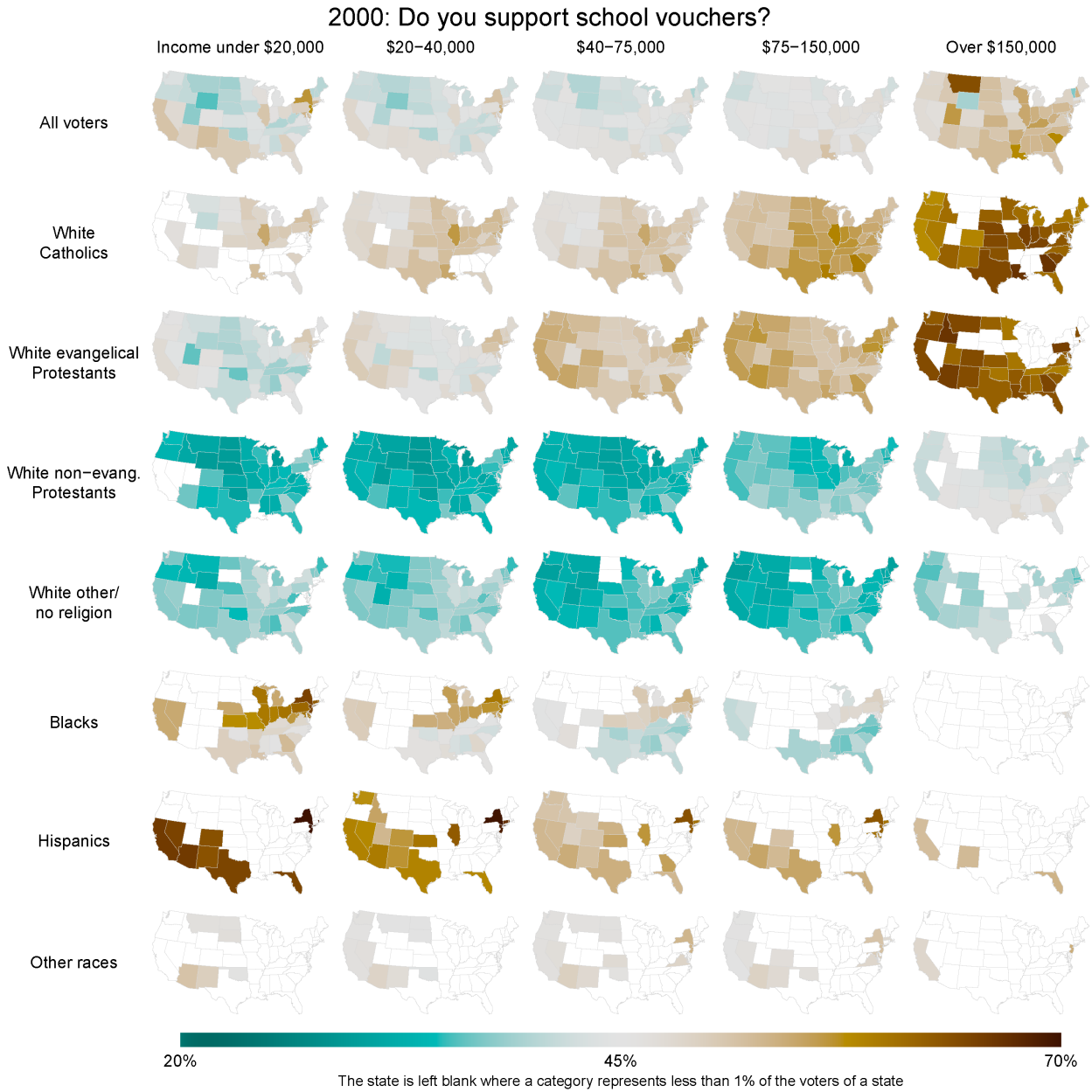

2010



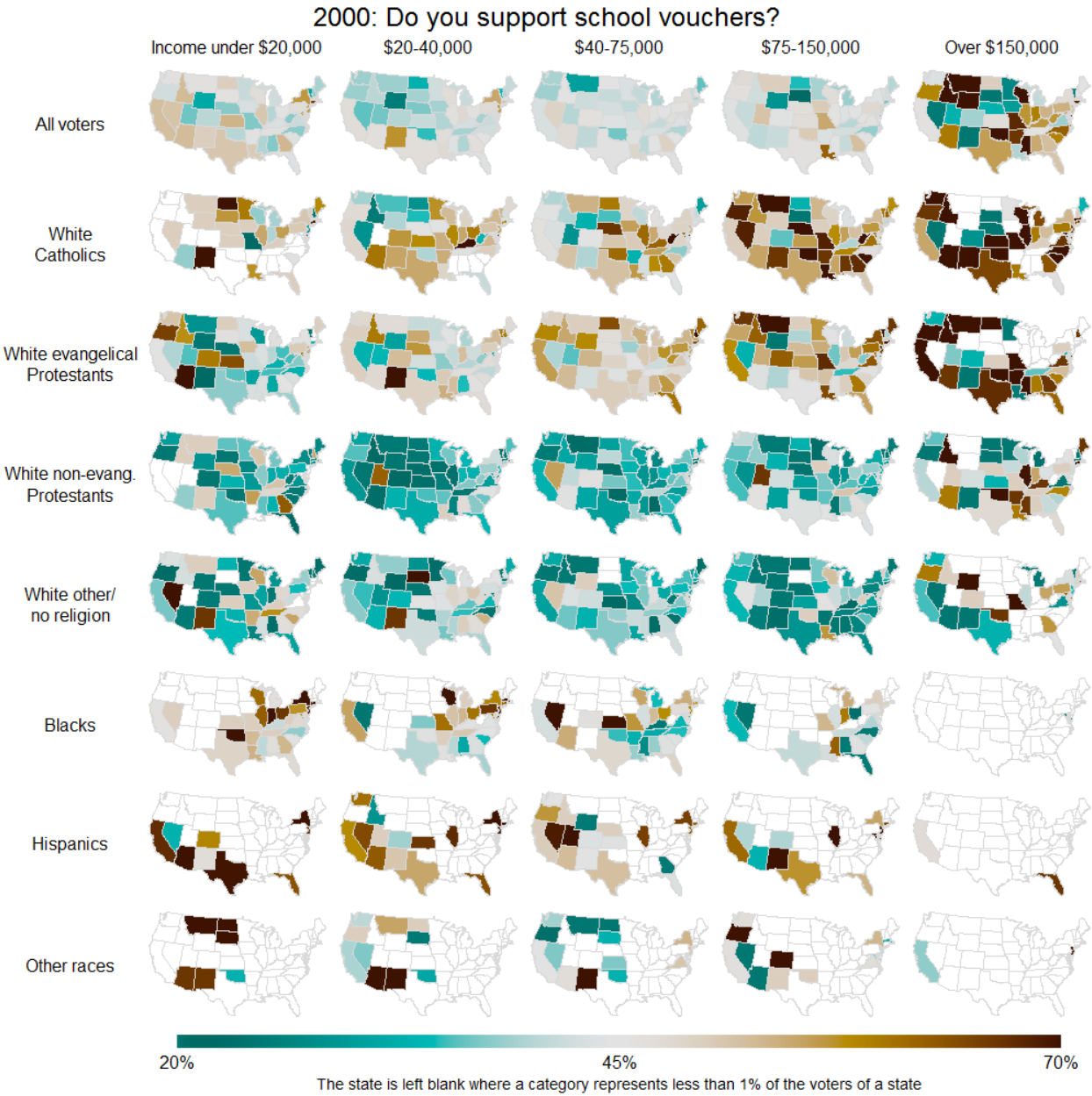

The long tail . . .

. . . and the paradox of freedom

# Ethnicity/religion, income, and school vouchers



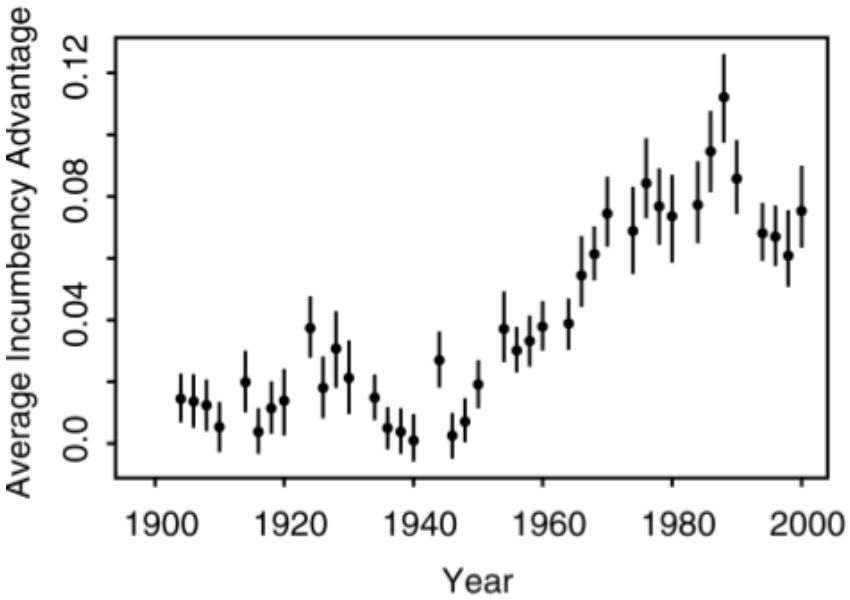
# The raw data


2000: Do you support school vouchers?
Income under $20,000
$20-40,000
$40-75,000
$75-150,000
Over $150,000
All voters
White Catholics
White evangelical Protestants
White non-evang. Protestants
White other/ no religion
Blacks
Hispanics
Other races
20%
45%
70%
The state is left blank where a category represents less than 1% of the voters of a state

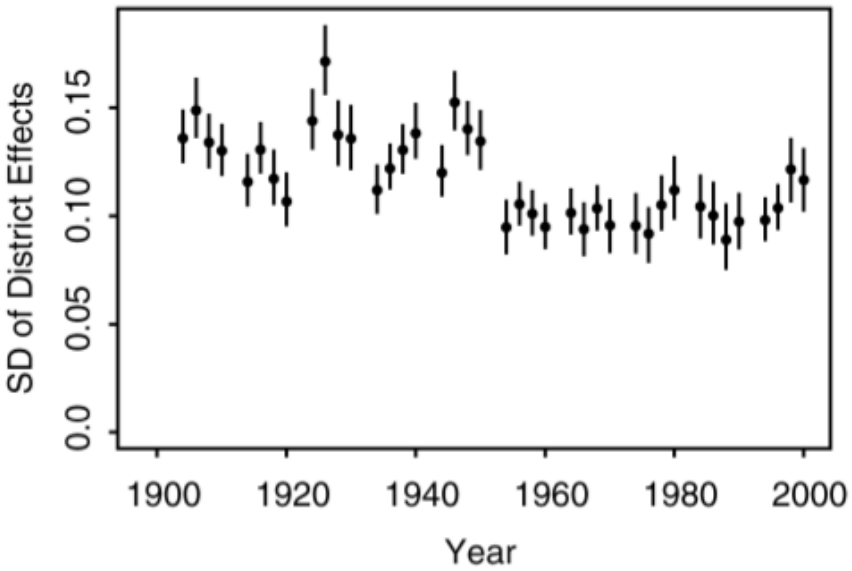
# Doing it without the technology

- Display data or simple estimates in a grid of graphs
- Implicit multilevel modeling by eye

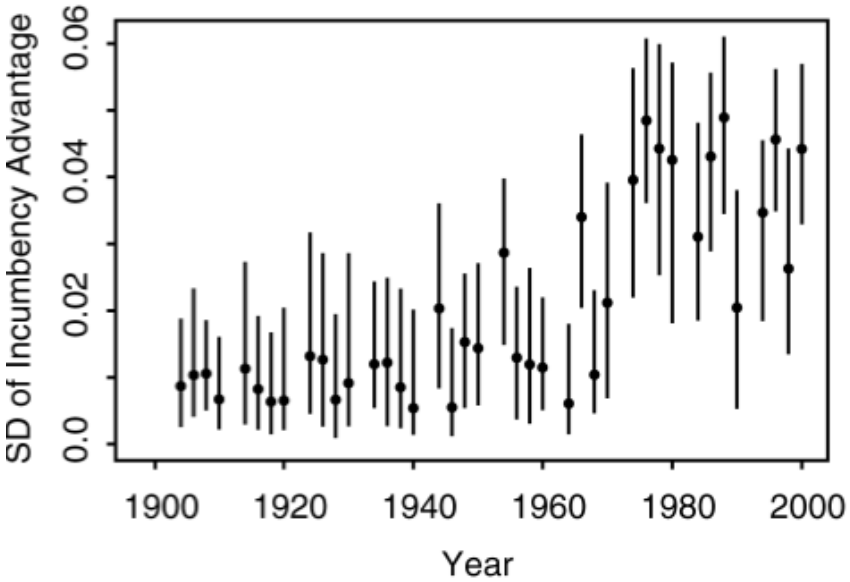
# Example: incumbency advantage over time

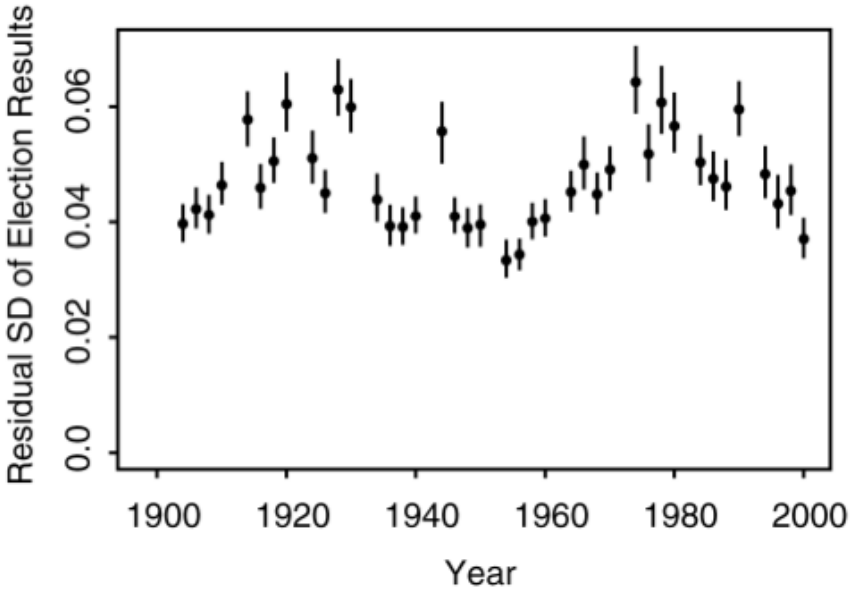



(c)




(d)



# Take-home points

- When using small samples to study small effects, any statistically significant finding is *necessarily* a huge overestimate
- Don't study factors (e.g., beauty) in isolation
- Place them in a larger model
- Statistical modeling as an exploratory tool
- Discussion: what is "modern statistics" and what do you need to know about it?