





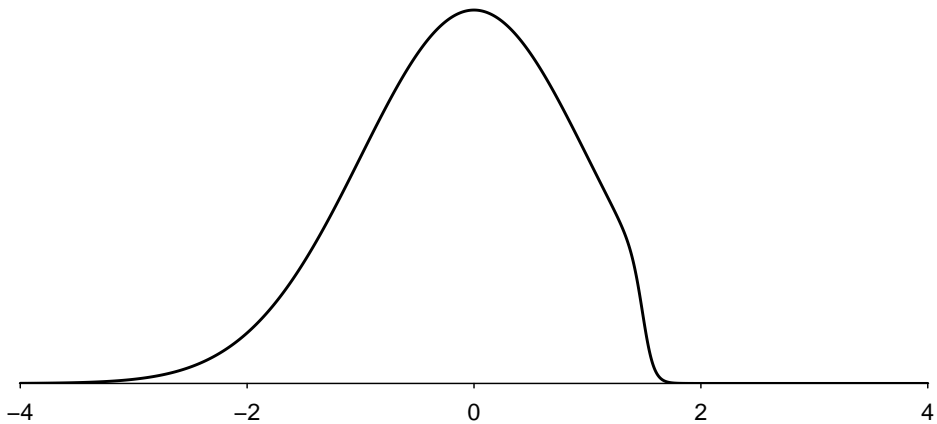




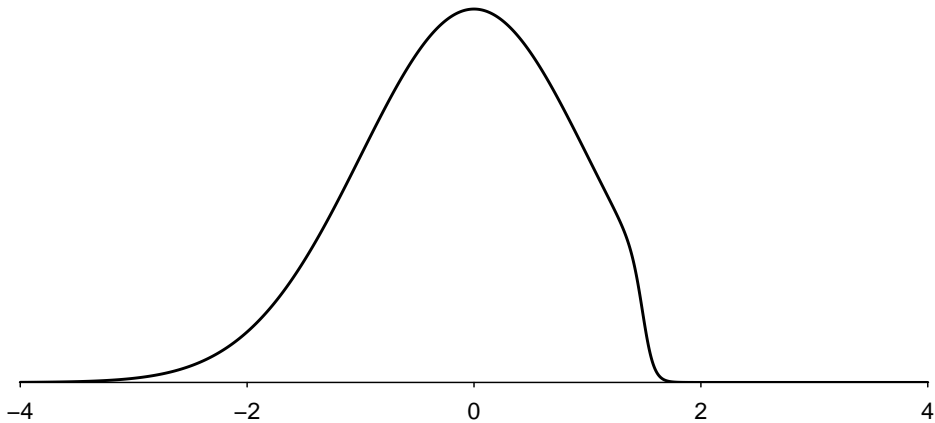




## Observed data

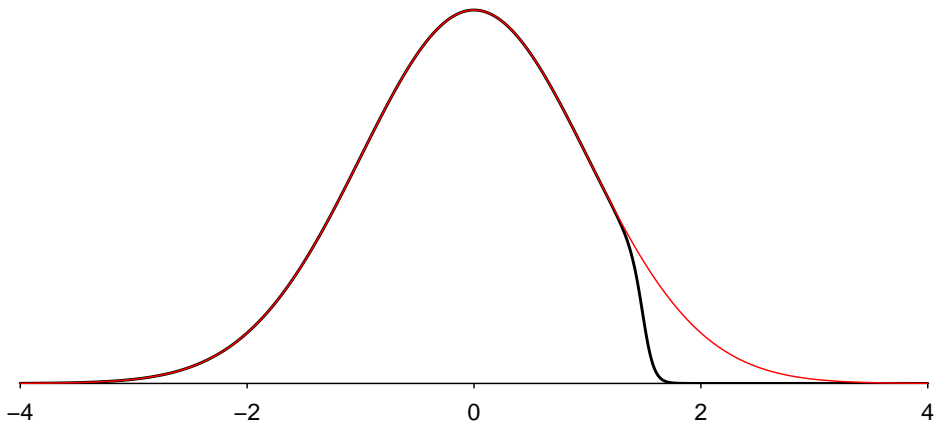


## Imputed data--Missing completely at random

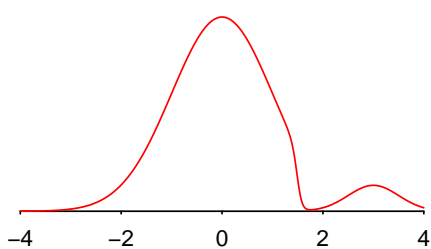
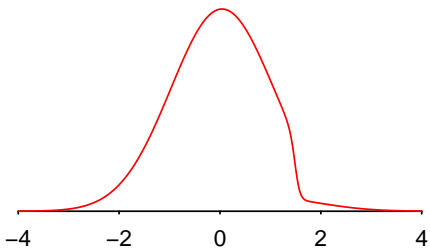
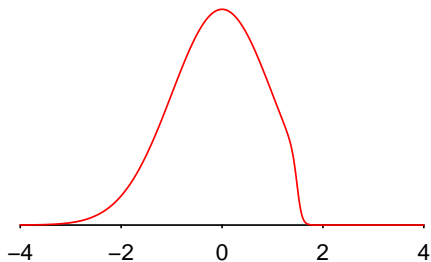
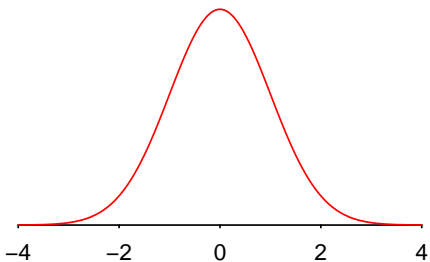


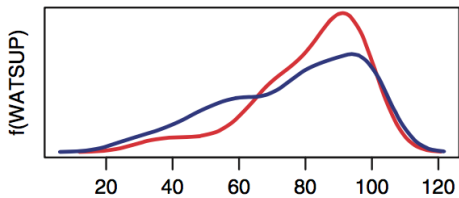
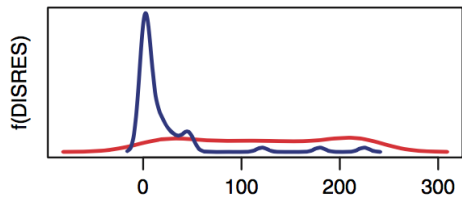
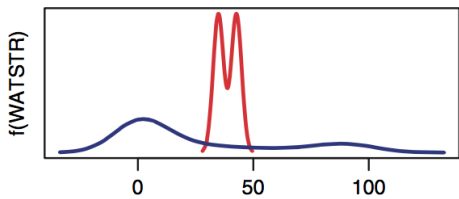
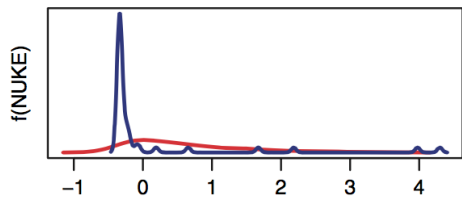
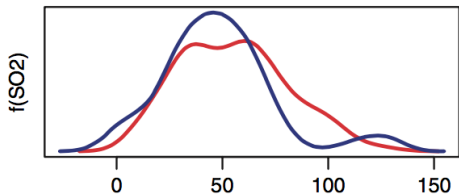
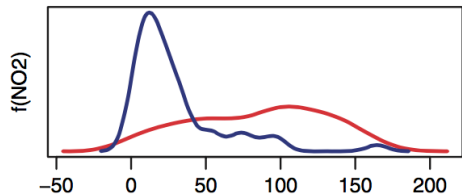


## Imputed data--Fitted normal model



## Imputed data--Various assumptions





# Easier Said Than Done: Open Problems in Multilevel Regression and Poststratification

Andrew Gelman

Department of Statistics and Department of Political Science  
Columbia University, New York

Conference in honor of Rod Little, 31 Oct 2015

From Wikipedia:

There is no satisfactory statistical solution to deal with missing data that may not be at random. Assuming an extreme bias in the responders is one suggested method of dealing with low survey response rates. Getting a high response rate (>80%) from a small, random sample is considered preferable to a low response rate from a large sample.<sup>[9]</sup>

Trend data were not readily available from the large number of private organizations who conduct survey research and polls. These organizations claim to be experiencing problems as the following excerpt from the 1973 Conference on Surveys of Human Populations (American Statistical Association, 1974) shows:

"...spokesmen for a number of private survey organizations, large and small, who were queried by one of the conference participants, all report that their completion rates on general population samples now average approximately 60 to 65 percent, in spite of three or four callbacks. This recent experience is in contrast to a completion figure of 80 to 85 percent for the same firms in the decade of the sixties."

---

## Surveys Face Growing Difficulty Reaching, Persuading Potential Respondents

|  | 1997 | 2000 | 2003 | 2006 | 2009 | 2012 |
|--|------|------|------|------|------|------|
|  | %    | %    | %    | %    | %    | %    |
| <b>Contact rate</b><br>(percent of households in which an adult was reached)           | 90   | 77   | 79   | 73   | 72   | 62   |
| <b>Cooperation rate</b><br>(percent of households contacted that yielded an interview) | 43   | 40   | 34   | 31   | 21   | 14   |
| <b>Response rate</b><br>(percent of households sampled that yielded an interview)      | 36   | 28   | 25   | 21   | 15   | 9    |

PEW RESEARCH CENTER 2012 Methodology Study. Rates computed according to American Association for Public Opinion Research (AAPOR) standard definitions for CON2, COOP3 and RR3. Rates are typical for surveys conducted in each year.

---

AD NO. \_\_\_\_\_  
DDC FILE COPY

AD A 053

6

Survey Response Rates: Some Trends, Causes and Correlates,

10

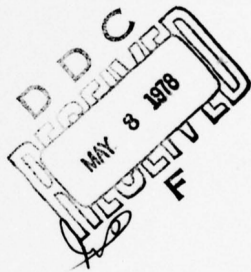
Kent H. Marquis

11

1 Apr 1977

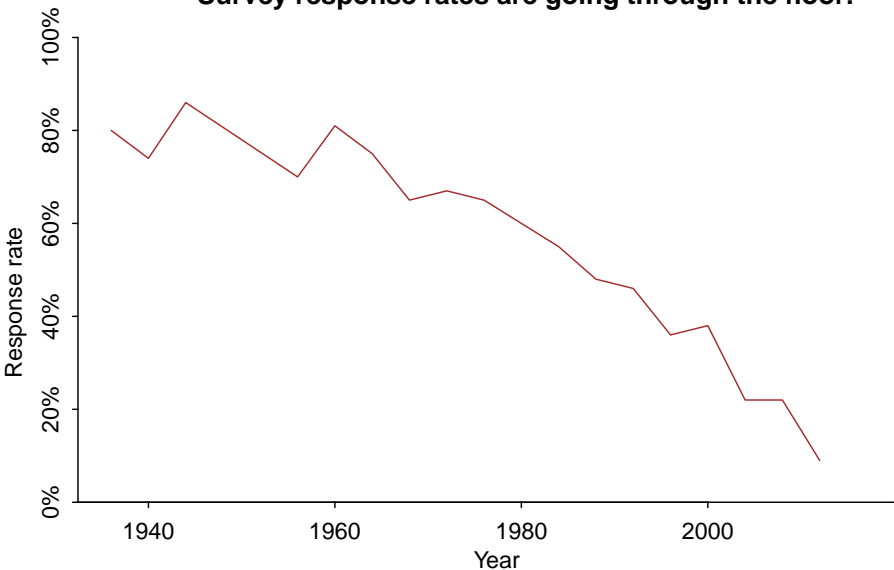
12

29p.

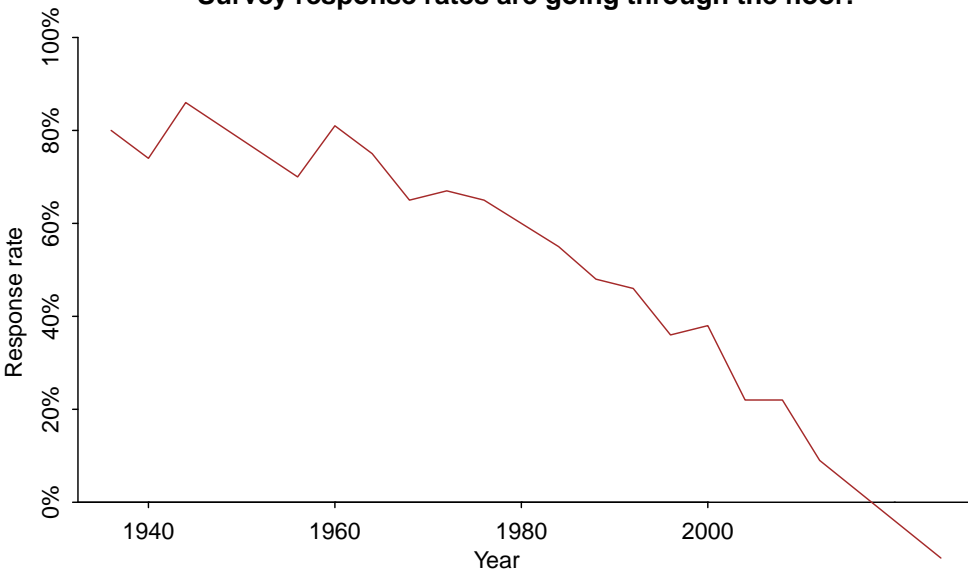




## Survey response rates are going through the floor!



## Survey response rates are going through the floor!



## The poststratification identity

$$\theta = \frac{\sum_{j=1}^J N_j \theta_j}{\sum_{j=1}^J N_j}$$

## The poststratified estimate

$$\hat{\theta} = \frac{\sum_{j=1}^J N_j \hat{\theta}_j}{\sum_{j=1}^J N_j}$$



me social tv video games music apps

A side-by-side portrait of Joe Biden on the left and Mitt Romney on the right, both in suits and ties, against a blue background.


Join the VP Debate  
We're polling LIVE, 9 PM ET

**UFC** 153

Press Conf

A stylized, glowing image of a person's head and shoulders, likely representing a Kinect game or application.

Kinect Cent

A close-up image of a person's hand holding a motorcycle handlebar, with a motorcycle wheel visible in the background.

New Trials

A close-up portrait of a man with a beard and mustache, looking directly at the camera.

ADVERTISEMENT

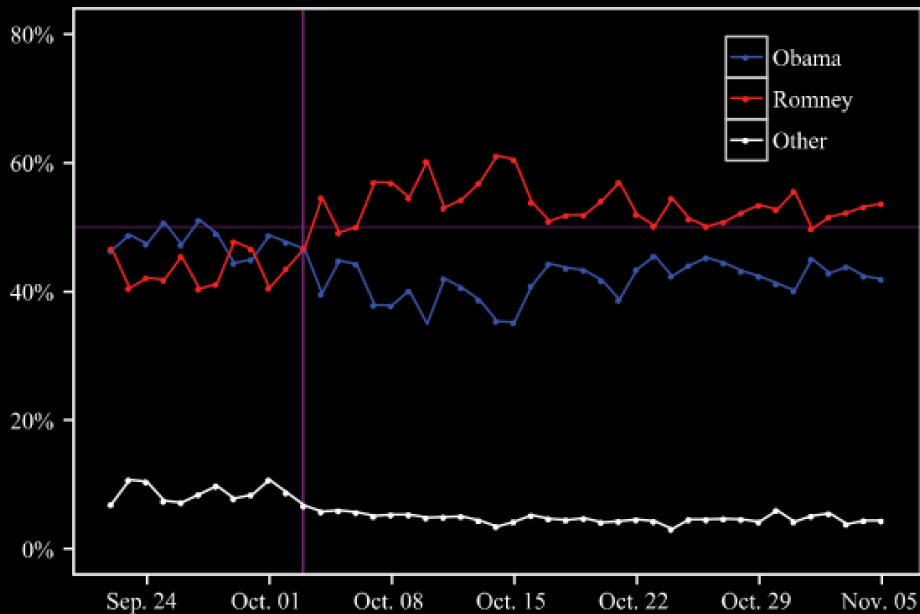
If the election were held today, who would you vote for?

Barack Obama

Mitt Romney

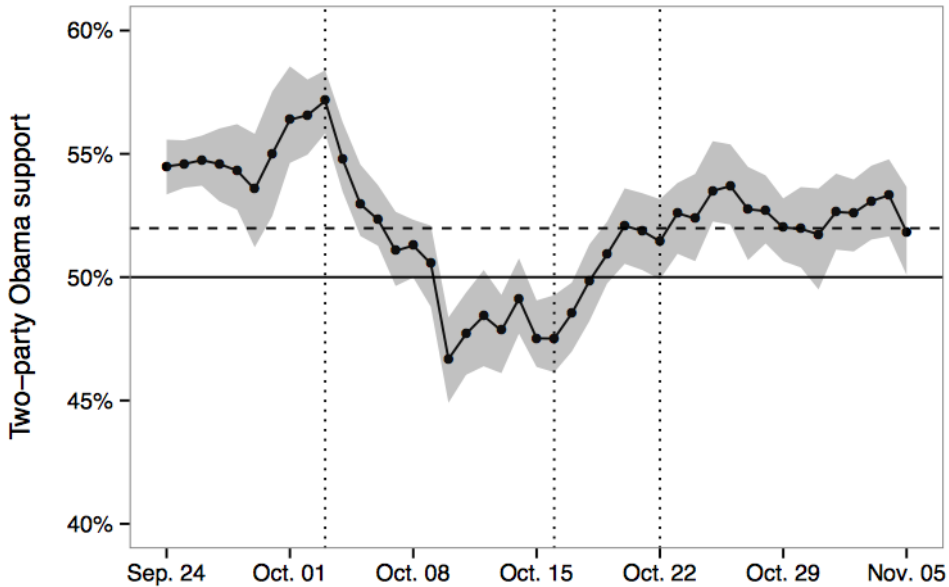
Other

Not sure

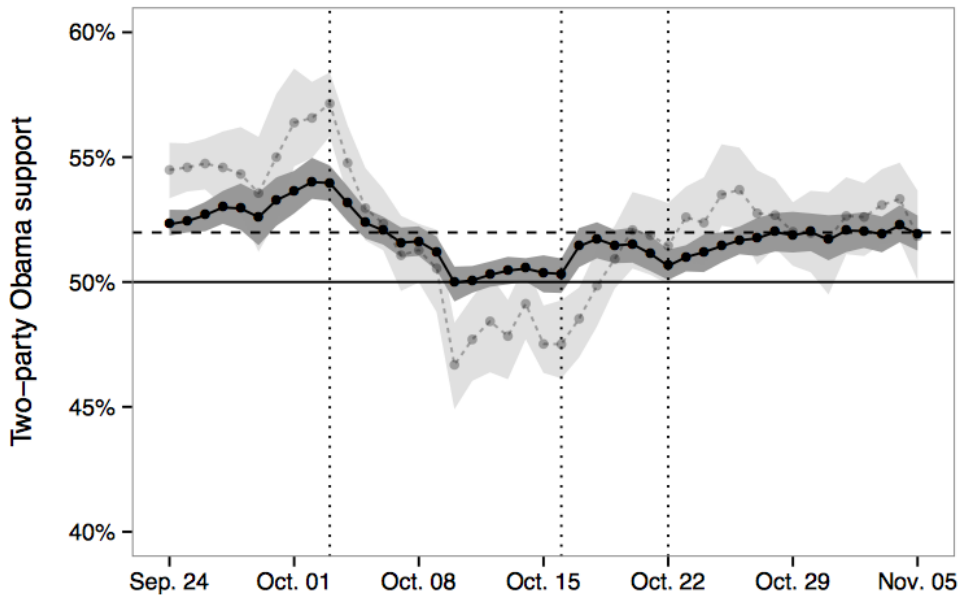




Xbox estimates, adjusting for demographics:



Xbox estimates, adjusting for demographics and partisanship:



# Why multilevel regression?



# Open problems in MRP

- ▶ Deep interactions
- ▶ Non-census variables
- ▶ Survey weights
- ▶ Cluster sampling
- ▶ Estimating regression coefficients
- ▶ Building trust in results