Survey Methodologies for Sensitive Questions

Kosuke Imai

Princeton University

October 30, 2014

Seminar at McGill University

Joint work with
Graeme Blair, Jason Lyall, Bryn Rosenfeld, Jacob Shapiro

Introduction

- Challenge of measuring sensitive attitudes and behaviors
  - social desirability bias
  - non-response bias

- Indirect methods becoming increasingly popular
  - list experiments
  - endorsement experiment
  - randomized response

- Development of statistical methods
  - multivariate regression for each survey technique (Bullock, Imai and Shapiro 2010; Imai 2011; Blair & Imai 2012; Blair, Imai & Zhou 2014)
  - using responses as predictors in outcome regression (Imai, Park & Greene 2014)

- Empirical validation studies
  - validation against ground truth (Rosenfeld, Imai & Shapiro 2014)
  - comparison of multiple measurements (Blair, Imai & Lyall 2014)
  - prediction of behavior (Hirose, Imai & Lyall 2014)

The Mississippi Validation Study (Rosenfeld, Imai & Shapiro 2014)

- Estimate voting on anti-abortion referendum using:
  - direct question
  - list experiment (item/unmatched count technique)
  - endorsement experiment
  - randomized response

- Validate estimates against official election outcome:
  - sample from voter history file
  - county-level voting recap reports for validation

- Case selection:
  - a poll conducted 24 hours before the election predicts 44% no votes
  - the amendment was defeated 58% to 42%

- Findings:
  - direct question ⇝ significant under-estimation though efficient
  - indirect methods ⇝ much less biased though more variable
  - endorsement and randomized response ⇝ least bias

Direct Question

Did you vote YES or NO on the Personhood Initiative, which appeared on the November 2011 Mississippi General Election ballot?

Voted Yes

Voted No

Did not vote

Don't know

Refused

List Experiment

Here is a list of four things that some people have done and some people have not. Please listen to them and then tell me HOW MANY of them you have done in the past two years. Do not tell me which you have and have not done. Just tell me how many:

Discussed politics with family or friends

Cast a ballot for Governor Phil Bryant

Paid dues to a union

Given money to a Tea Party candidate or organization

(treatment) Voted 'YES' on the 'Personhood' Initiative

How many of these things have you done in the past two years?

Endorsement Experiment

We'd like to get your overall opinion of some people in the news.
As I read each name, please say if you have a very favorable,
somewhat favorable, somewhat unfavorable, or very unfavorable
opinion of each person.

(control) Phil Bryant, Governor of Mississippi?

(treatment) Phil Bryant, Governor of Mississippi, who
campaigned in favor of the 'Personhood' Initiative on the
2011 Mississippi General Election ballot?

Randomized Response

To answer this question, you will need a coin. Once you have found one, please toss the coin two times and note the results of those tosses (heads or tails) one after the other on a sheet of paper. Do not reveal to me whether your coin lands on heads or tails. After you have recorded the results of your two coin tosses, just tell me you are ready and we will begin.

Now, please answer 'yes' if either your second coin toss came up heads or you voted 'YES' on the Personhood Initiative, which appeared on the November 2011 Mississippi General Election ballot.

Yes

No

Don't know

Refused

Method for List Experiment (Imai 2011; Blair & Imai 2012)

- Setup:
  - Y_i: observed response
  - X_i: observed covariates
  - Z_i^*: latent response to the sensitive item
  - Y_i^*: latent response to the control items
  - T_i: treatment such that $Y_i = Y_i^* + T_i Z_i^*$

- Assumptions: (1) no design effect, (2) no liar

- A total of $(2 \times (J+1))$ "types" (Y_i^*, Z_i^*)
- Example: three control items $(J = 3)$

Y_i	Treatment group		Control group	
4	(3,1)			
3	(2,1)	(3,0)	(3,1)	(3,0)
2	(1,1)	(2,0)	(2,1)	(2,0)
1	(0,1)	(1,0)	(1,1)	(1,0)
0	(0,0)		(0,1)	(0,0)

Method for List Experiment (Imai 2011; Blair & Imai 2012)

- Setup:
  - Y_i: observed response
  - X_i: observed covariates
  - Z_i^*: latent response to the sensitive item
  - Y_i^*: latent response to the control items
  - T_i: treatment such that $Y_i = Y_i^* + T_i Z_i^*$

- Assumptions: (1) no design effect, (2) no liar

- A total of $(2 \times (J + 1))$ "types" (Y_i^*, Z_i^*)
- Example: three control items $(J = 3)$

Y_i	Treatment group		Control group	
4	(3,1)			
3	(2,1)	(3,0)	(3,1)	(3,0)
2	(1,1)	(2,0)	(2,1)	(2,0)
1	(0,1)	(1,0)	(1,1)	(1,0)
0	(0,0)		(0,1)	(0,0)

Statistical Model for List Experiment

▶ Submodel for the response to the sensitive item Z_i^*:

$$\Pr(Z_i^* = 1 \mid X_i) = \text{logit}^{-1}(\alpha + \beta^\top X_i)$$

▶ Submodel for the responses to the control items Y_i^*:

$$\Pr(Y_i^* = y \mid X_i, Z_i^*) = \text{Binomial}(J, \text{logit}^{-1}(\gamma + \delta^\top X_i + \zeta Z_i^*))$$

▶ Combine them under the likelihood framework

Model for Endorsement Experiment (Bullock, Imai & Shapiro 2011)

- Setup:
  - T_i: treatment
  - Y_i: observed (ordinal) response
  - Y_i^*: latent (continuous) response
  - X_i: observed covariates
  - V_i^*: latent ideological position
  - Z_i^*: latent additional support inducted by the endorsement

- Latent measurement model:

$$Y_i^* \overset{\text{indep.}}{\sim} \mathcal{N}(\beta(V_i^* + T_i Z_i^*) - \alpha,\ 1)$$

  with appropriate cut-points

- Hierarchical model V_i^* and Z_i^*:

$$V_i^* \overset{\text{indep.}}{\sim} \mathcal{N}(\delta^\top X_i, 1)$$

$$Z_i^* \overset{\text{indep.}}{\sim} \mathcal{N}(\lambda^\top X_i, \omega^2)$$

- Probability of positive support: $\Pr(Z_i^* > 0 \mid X_i)$

Model for Randomized Response (Blair, Imai & Zhou 2014)

- Setup:
  - Y_i: observed response
  - Z_i^*: latent response to the sensitive item
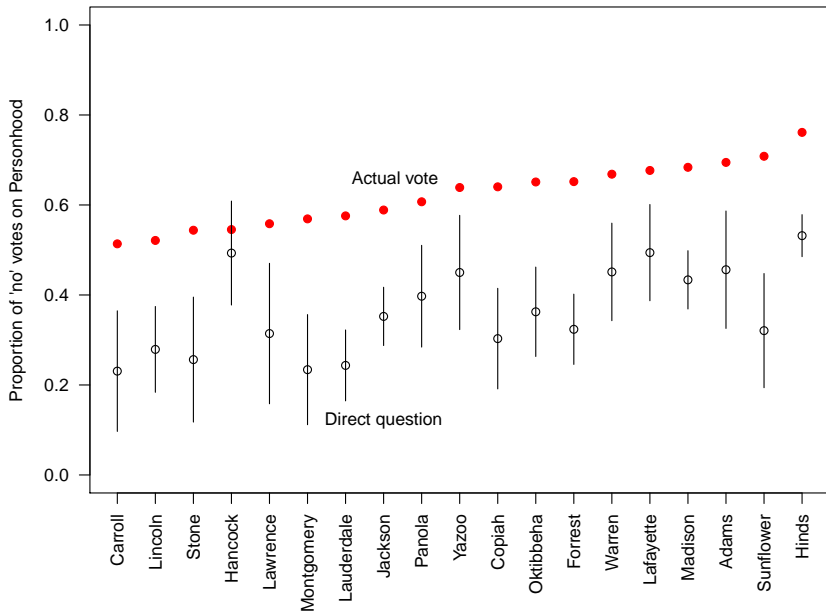 - X_i: covariates

- The model is,

$$\Pr(Z_i^* = 1 \mid X_i) \ \ = \ \ \mathrm{logit}^{-1}(\alpha + \beta^\top X_i)$$
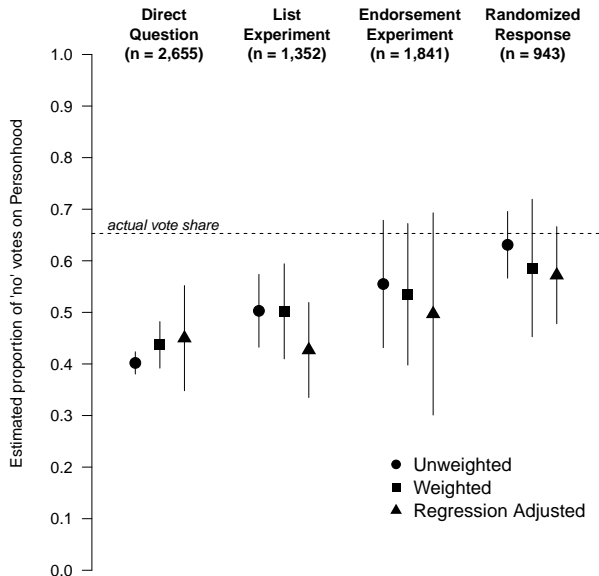
- The likelihood function is given by,

$$\prod_{i=1}^{n} \left\{ \frac{1}{2} \cdot \mathrm{logit}^{-1}(\alpha + \beta^\top X_i) \right\}^{Y_i} \left\{ 1 - \frac{1}{2} \cdot \mathrm{logit}^{-1}(\alpha + \beta^\top X_i) \right\}^{1 - Y_i}$$

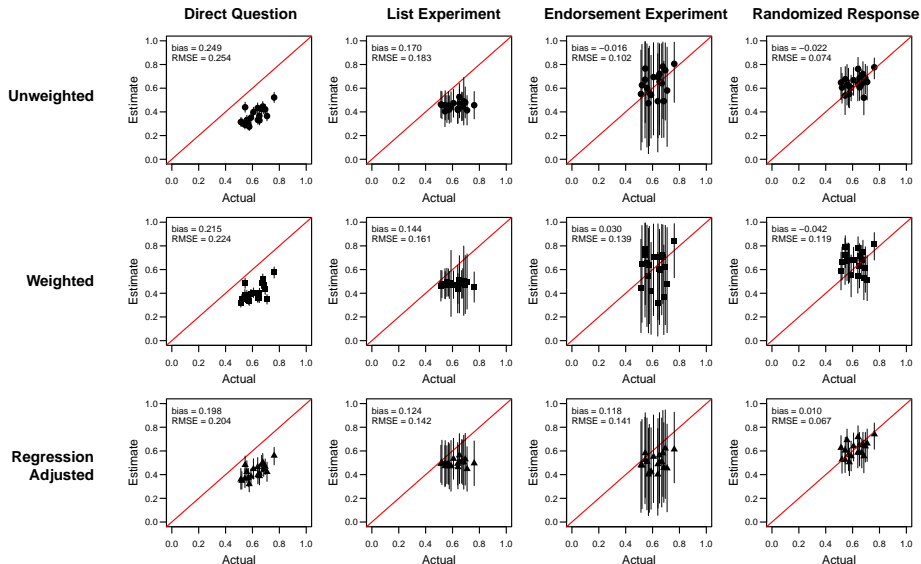- Many other designs and accompanying methods are available

Bias of the Direct Question

Pooled Analysis

County-level Analysis

Hearts and Minds in Afghanistan (Blair, Imai & Lyall, 2014)

- ▶ How do we measure civilian attitudes in a conflict setting?

- ▶ Current efforts in Afghanistan rely on direct questions:
    1. USAID (TCAPF): "Who do you believe can solve your problems?"
    2. ISAF (ANQAR): "Over the past 6 months, do you think the Taliban have grown stronger, grown weaker, or remained the same?"

- ▶ Why are direct questions a bad idea?
    1. Threats to enumerators and respondents
    2. Nonresponse, social desirability bias
    3. Interviews are public
    4. Danger of selection bias in sampling locations (role of gatekeepers)

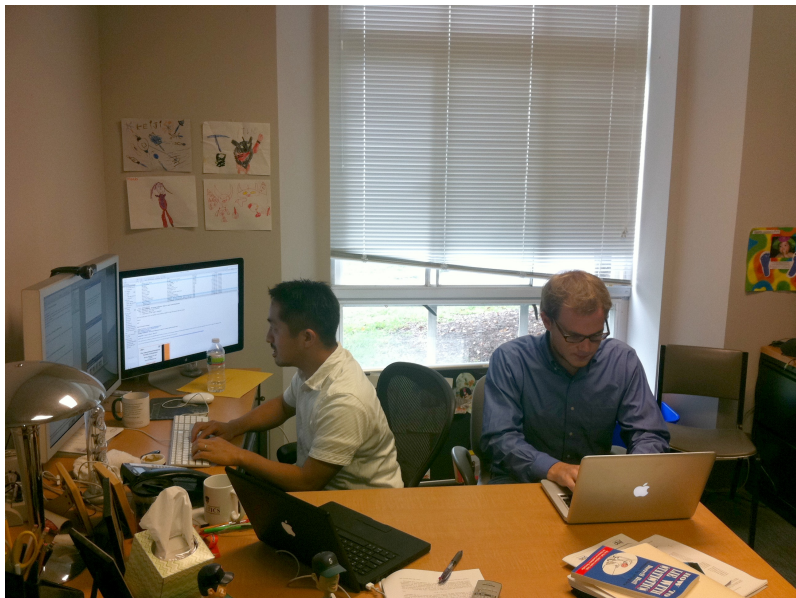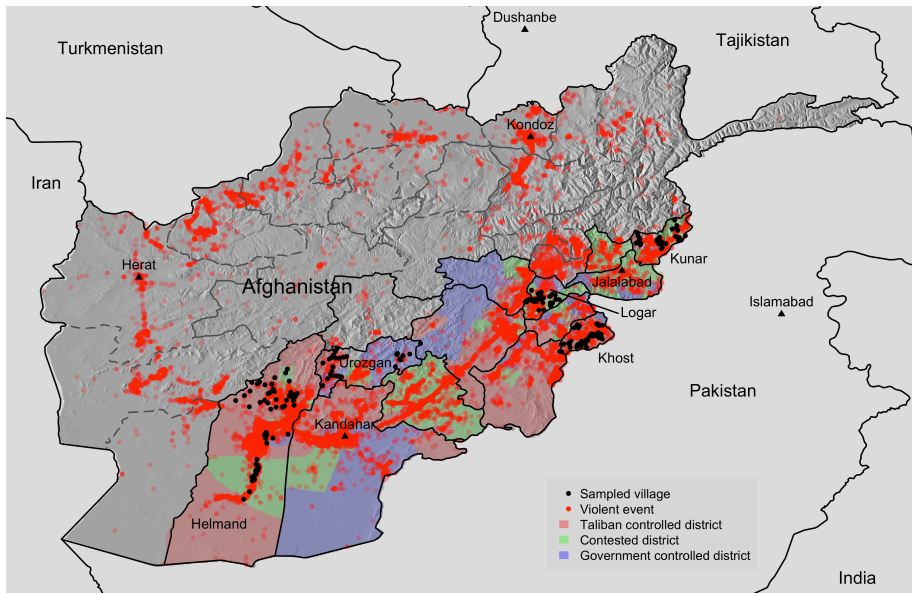- ▶ ANQAR (November-December 2011): 50% refusal rate

Negotiated Access

Princeton Battlefield

Surveying in the Heartland of Insurgency

Endorsement Experiments

- Script for the control group:

  ```
  A recent proposal calls for the sweeping reform of the
  Afghan prison system, including the construction of new
  prisons in every district to help alleviate
  overcrowding in existing facilities.  Though expensive,
  new programs for inmates would also be offered, and new
  judges and prosecutors would be trained.  How do you
  feel about this proposal?

  Strongly agree;   Agree;   Indifferent;   Disagree;
  Strongly disagree;   Don't Know;   Refuse to answer
  ```

Endorsement Experiments

▶ Script for the treatment group:

A recent proposal **by ISAF** calls for the sweeping reform
of the Afghan prison system, including the construction
of new prisons in every district to help alleviate
overcrowding in existing facilities. Though expensive,
new programs for inmates would also be offered, and new
judges and prosecutors would be trained. How do you
feel about this proposal?

Strongly agree;   Agree;   Indifferent;   Disagree;
Strongly disagree;   Don't Know;   Refuse to answer

Data from the Endorsement Experiments

List Experiments

▶ Script for the control group:

```
I'm going to read you a list with the names of
different groups and individuals on it. After I read
the entire list, I'd like you to tell me how many of
these groups and individuals you broadly support,
meaning that you generally agree with the goals and
policies of the group or individual. Please don't tell
me which ones you generally agree with; only tell me
how many groups or individuals you broadly support.

Karzai Government; National Solidarity Program; Local
Farmers
```

List Experiments

▶ Script for the treatment group:

I'm going to read you a list with the names of different groups and individuals on it. After I read the entire list, I'd like you to tell me how many of these groups and individuals you broadly support, meaning that you generally agree with the goals and policies of the group or individual. Please don't tell me which ones you generally agree with; only tell me how many groups or individuals you broadly support.

Karzai Government; National Solidarity Program; Local Farmers; **ISAF**

Comparing and Combining List and Endorse Experiments

- Need for validation $\implies$ Multiple measurement strategy
- Two measures should give similar results

- What is the probability of supporting ISAF?
  1. List: prob. of saying yes to the sensitive item
  2. Endorsement: prob. of endorsement having a positive effect on support for policy
- These probabilities should be similar!
- They can be estimated with a new multivariate regression method
- Endorsement and list experiments can even be combined for a joint analysis

Statistical Analysis of Endorsement Experiments

- Item response theory to combine questions:

$$\Pr(Y_{ij} = 1 \mid T_i = k) = \Phi(\alpha_j + \beta_j(x_i + s_{ijk}))$$

  - α_j: average popularity of policy j
  - β_j: how much policy j differentiates pro- and anti-reform respondents
  - x_i: "ideal point" = how pro-reform respondent i is
  - s_{ijk}: support level for combatant k in policy j

- Multi-stage sampling $\rightsquigarrow$ Multi-level modeling

$$s_{ijk} \overset{\text{indep.}}{\sim} \mathcal{N}(\lambda_{k,\text{village}[i]} + Z_i^\top \lambda_k^Z, \ \omega_{k,\text{village}}^2)$$
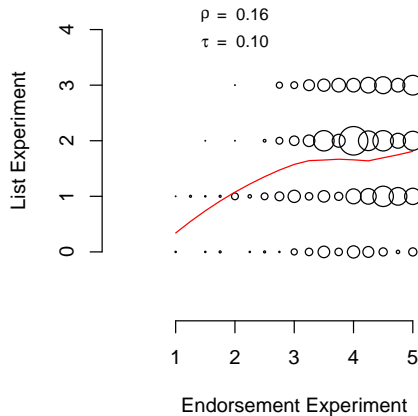
$$\lambda_{k,\text{village}[i]} \overset{\text{indep.}}{\sim} \mathcal{N}(\lambda_{k,\text{district}[i]} + V_{\text{village}[i]}^\top \lambda_k^V, \ \omega_{k,\text{district}}^2)$$

$$\lambda_{k,\text{district}[i]} \overset{\text{indep.}}{\sim} \mathcal{N}(\lambda_{k,\text{province}[i]} + W_{\text{district}[i]}^\top \lambda_k^W, \ \omega_{k,\text{province}}^2)$$
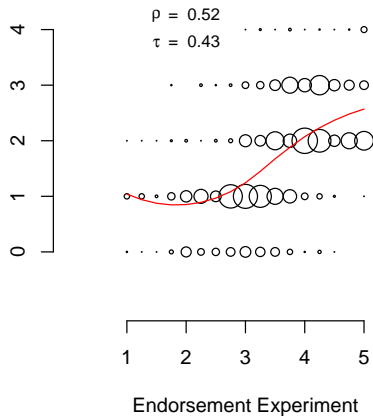
- Same multi-level structure for ideal points x_i
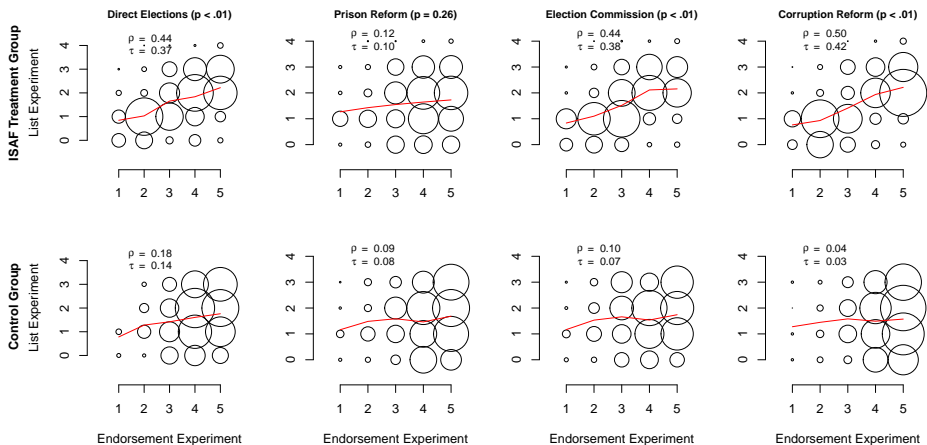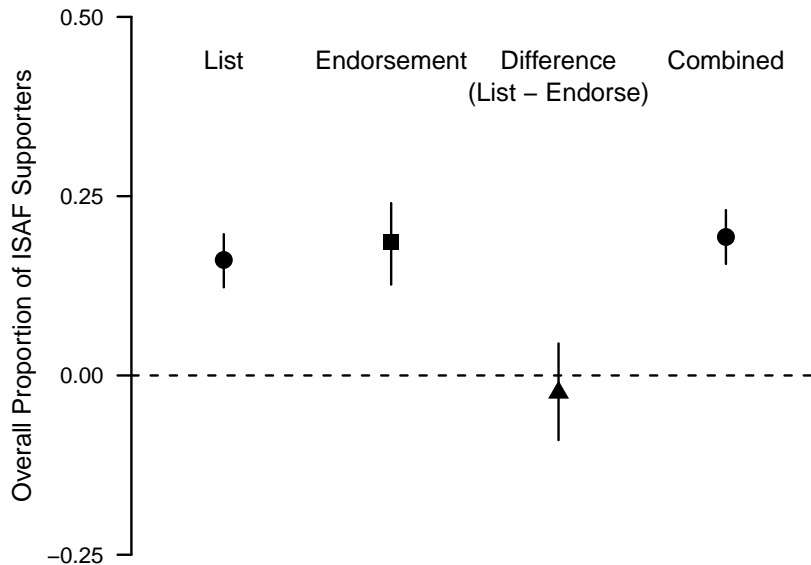
Descriptive Comparison: Overall

Descriptive Comparison: Question by Question

Overall Proportion of ISAF Supporters

Concluding Remarks

- Direct question is severely biased

- All indirect methods reduce bias:
    - Endorsement and randomized response ⤳ least bias
    - List experiment ⤳ ceiling/floor effects, design effects
    - Ease of implementation: list > endorse > randomized response

- Use of multiple-measurement strategies when truth is not available

- Future research directions:
    - How to balance bias, precision, and cost
    - Use aggregate-level truth to improve individual-level estimates

- Open-source software:
    - `list` for list experiment (Blair, Imai & Park)
    - `endorse` for endorsement experiment (Shiraito & Imai)
    - `rr` for randomized response (Blair, Imai & Zhou)

Project Reference

- Development of new methods:
  - "Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan." *Political Analysis*
  - "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association*
  - "Statistical Analysis of List Experiments." *Political Analysis*
  - "Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan." *American Journal of Political Science*
  - "Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models." *Political Analysis*
  - "Design and Analysis of the Randomized Response Technique."

- Empirical applications and validations:
  - "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review*
  - "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions."
  - "Can Civilian Attitudes Predict Civil War Violence?"