Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan

Kosuke Imai
Princeton University
Joint work with G. Blair and J. Lyall

March 1, 2013
University of Michigan Political Science Department
Survey is used widely in social sciences
Validity of survey depends on the accuracy of self-reports

Sensitive questions $\implies$ social desirability, privacy concerns
Racial prejudice, corruption, support for political actors
Lies and nonresponses $\implies$ potential bias

Survey “experiments” as a solution

1. Randomization: Randomized response method
2. Aggregation: List experiment (item count technique)
3. Cueing: Endorsement experiment

Two problems of indirect measures and proposed solutions:

1. Measurement error $\implies$ comparing two measures
2. Statistical inefficiency $\implies$ combining two measures
Theoretical and Substantive Motivation

- How do we measure “hearts and minds” 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
A Battlefield in Princeton, New Jersey

Kosuke Imai (Princeton)
List and Endorsement Experiments
March 1, 2013
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
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
The Data from the List Experiment

<table>
<thead>
<tr>
<th>response value</th>
<th>Control Group frequency</th>
<th>Control Group proportion</th>
<th>ISAF Treatment Group frequency</th>
<th>ISAF Treatment Group proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>188</td>
<td>20.5%</td>
<td>174</td>
<td>19.0%</td>
</tr>
<tr>
<td>1</td>
<td>265</td>
<td>28.9</td>
<td>278</td>
<td>30.3</td>
</tr>
<tr>
<td>2</td>
<td>265</td>
<td>28.9</td>
<td>260</td>
<td>28.3</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>21.8</td>
<td>182</td>
<td>19.8</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td></td>
<td></td>
<td>2.6</td>
</tr>
<tr>
<td>Total</td>
<td>918</td>
<td></td>
<td>918</td>
<td></td>
</tr>
</tbody>
</table>
Identification Assumptions

1. **No Design Effect**: The inclusion of the sensitive item does not affect answers to control items.

2. **No Liars**: Answers about the sensitive item are truthful.
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
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

<table>
<thead>
<tr>
<th></th>
<th>List Experiment</th>
<th>Endorsement Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct Elections</td>
<td>Prison Reform</td>
</tr>
<tr>
<td></td>
<td>Independent Election Commission</td>
<td>Anti-Corruption Reform</td>
</tr>
</tbody>
</table>

#### Overall (N = 1836)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Helmand (N = 570)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Khost (N = 420)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Kunar (N = 264)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Logar (N = 324)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Urozgan (N = 258)

<table>
<thead>
<tr>
<th></th>
<th>ISAF</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- 0 items
- 1
- 2
- 3
- 4
- Strongly disagree
- Disagree
- Indifferent
- Agree
- Strongly agree
- Don't know
- Refused

---

Kosuke Imai (Princeton)  List and Endorsement Experiments  March 1, 2013

14 / 32
Assumption and Interpretation

- More indirect than list experiments
- Easier to implement but harder to interpret

- When can we interpret endorsement effects as support (or affinity) for endorser?
  - Endorsements have no influence on respondents’ interpretation of policy questions. No learning
  - I’m a hardcore Democrat but don’t know much about this traditionally democratic policy. You now tell me even a Republican supports it and so the policy must be really good

- Some considerations when designing endorsement experiments:
  1. Policies must belong to the same policy dimension
  2. Endorsements must be credible
  3. Few respondents with extreme views
A statistical test: $H_0: \rho_0 = \rho_1$ and $H_1: \rho_0 < \rho_1$ with bootstrap
Descriptive Comparison: Question by Question

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Treatment Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prison Reform (p &lt; .01)</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>rho = 0.44</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>Direct Elections (p = 0.26)</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>rho = 0.12</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>Election Commission (p &lt; .01)</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>rho = 0.44</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>Corruption Reform (p &lt; .01)</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>rho = 0.50</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>Endorsement Experiment</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
<tr>
<td>rho = 0.04</td>
<td>ISAF</td>
<td>ISAF</td>
</tr>
</tbody>
</table>

Kosuke Imai (Princeton)
Descriptive Comparison: Violence & Territorial Control

District Violence Level

Low (p < .01)

High (p < .01)

District Territory Control

Taliban (p < .01)

Contested (p < .01)

Endorsement Experiment

List Experiment

List and Endorsement Experiments

Kosuke Imai (Princeton)
Models for List and Endorsement Experiments

- **LIST EXPERIMENTS** (Imai 2011, *JASA*; Blair & Imai 2012, *PA*):
  1. Likelihood framework with missing data
  2. Assumptions: no design effect, no liar
  3. Latent variable modeling for support

- **ENDORSEMENT EXPERIMENTS** (Bullock, Imai & Shapiro 2011, *PA*):
  1. Item response theory to combine multiple questions
  2. Assumptions: single policy dimension, no learning
  3. Latent variable modeling for support

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!
List Experiments Framework

- $N$ respondents
- $J$ control items
- $T_i$: binary treatment indicator ($1 = \text{treatment}, 0 = \text{control}$)
- $V_i$: pre-treatment covariates
- $Y_i$: outcome variable

Define a type of each respondent by
- total number of yes for $J$ control items $Y_i(0)$
- truthful answer to the sensitive item $Z_i^*$: $Y_i(1) = Z_i^* + Y_i(0)$
- A total of $(2 \times (J + 1))$ types

<table>
<thead>
<tr>
<th>$Y_i$</th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>(3,1)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(2,1) (3,0)</td>
<td>(3,1) (3,0)</td>
</tr>
<tr>
<td>2</td>
<td>(1,1) (2,0)</td>
<td>(2,1) (2,0)</td>
</tr>
<tr>
<td>1</td>
<td>(0,1) (1,0)</td>
<td>(1,1) (1,0)</td>
</tr>
<tr>
<td>0</td>
<td>(0,0)</td>
<td>(0,1) (0,0)</td>
</tr>
</tbody>
</table>
List Experiments Framework

- \( N \) respondents
- \( J \) control items
- \( T_i \): binary treatment indicator (1 = treatment, 0 = control)
- \( X_i \): pre-treatment covariates
- \( Y_i \): outcome variable

Define a **type** of each respondent by

- total number of yes for \( J \) control items \( Y_i(0) \)
- truthful answer to the sensitive item \( Z_i^* \): \( Y_i(1) = Z_i^* + Y_i(0) \)
- A total of \( 2 \times (J + 1) \) types

<table>
<thead>
<tr>
<th>( Y_i )</th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>(3,1)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(2,1) (3,0)</td>
<td>(3,1) (3,0)</td>
</tr>
<tr>
<td>2</td>
<td>(1,1) (2,0)</td>
<td>(2,1) (2,0)</td>
</tr>
<tr>
<td>1</td>
<td>(0,1) (1,0)</td>
<td>(1,1) (1,0)</td>
</tr>
<tr>
<td>0</td>
<td>(0,0)</td>
<td>(0,1) (0,0)</td>
</tr>
</tbody>
</table>

**Joint distribution** of \((Y_i(0), Z_i^*)\) is identified
Model for sensitive item: e.g., probit regression

\[ \Pr(Z_i^* = 1 \mid V_i) = \Phi(V_i^\top \delta) \]

Model for control items given the response to sensitive item: e.g., binomial or beta-binomial probit regression

\[ \Pr(Y_i(0) = y \mid V_i, Z_i^* = z) = J \times \Phi(V_i^\top \psi_z) \]

Maximum likelihood with the EM algorithm or Bayes with MCMC
- $N$ respondents
- $J$ policy questions
- $Y_{ij} \in \{0, 1\}$: response of respondent $i$ to policy $j$ (can be ordinal)
- $T_{ij} \in \{0, 1\}$: random endorsement of policy $j$ for respondent $i$
- For the Afghan experiment, an individual receives the same treatment across policies $T_i = T_{ij}$
- $V_i$: Covariates measured prior to the treatment
Multiple questions $\implies$ item response theory

$$\Pr(Y_{ij} = 1 \mid T_i = t) = \Phi(\alpha_j + \beta_j(x_i + ts_{ij}^*))$$

- $\alpha_j$: average popularity of policy $j$
- $\beta_j$: how much policy $j$ differentiates pro- and anti-reform respondents
- $x_i$: “ideal point” = how pro-reform respondent $i$ is
- $s_{ij}^*$: endorsement effect

Support level:

$$s_{ij} = \begin{cases} 
  s_{ij}^* & \text{if } \beta_j \geq 0 \\
  -s_{ij}^* & \text{otherwise}
\end{cases}$$

such that $\frac{\partial}{\partial s_{ij}} \Pr(Y_{ij} = 1 \mid T_{ij} = 1) > 0$

Hierarchical model of support:

$$s_{ij} \overset{\text{indep.}}{\sim} \mathcal{N}(V_i^\top \lambda, \omega^2)$$
Comparing and Combining the Two Models

- Key quantity: Probability of being a supporter
- List experiments:

\[ \Pr(Z_i^* = 1 \mid V_i) = \Phi(V_i^\top \gamma) \]

- Endorsement experiments:

\[ \Pr(s_{ij} > 0 \mid V_i) = \Phi(V_i^\top \lambda/\omega) \]

- Compare the coefficients: \( \gamma \) and \( \lambda/\omega \)
- Combine the two models: \( \gamma = \lambda/\omega \)
Effects of Taliban and ISAF Victimization

Victimization

by Taliban

by ISAF

Effects on Probability of Supporting ISAF

Endorse

List

Combined

List

Combined

Kosuke Imai (Princeton)  List and Endorsement Experiments  March 1, 2013
Effects of Taliban/ISAF Post-Harm Mitigation Efforts

Approach after Victimization

by Taliban

by ISAF

List
Endorse
Combined

List
Endorse
Combined

Kosuke Imai (Princeton)
Effects of CERP Aid Spending

Proportion of ISAF Supporters vs. CERP Aid Spending (hundred thousands)

- Endorsement Experiment
- List Experiment
Concluding Remarks

- Challenges of eliciting truthful responses to sensitive questions
- List and endorsement experiments: indirect questioning methods

- Need for validation \(\iff\) multiple measurement strategy
- Statistical methods for comparing and combining list and endorsement experiments
- Open-source software list and endorse for implementation

- Practical suggestions:
  1. Randomize the treatment across, not within, respondents
  2. List experiments are more prone to social desirability bias than endorsement experiments
  3. Multiple pre-tests and focus groups
The project website for papers and software:

http://imai.princeton.edu/projects/sensitive.html

Email for comments and suggestions:

kimai@princeton.edu