Addressing Incentives to Conceal
Studying Sensitive Questions with Surveys

Yang-Yang Zhou
Department of Political Science, University of British Columbia

Population Data for the 21st century: Advances in data collection methodologies
December 6, 2019

Sensitive questions are widely asked in survey research

- What proportion of people have racial bias?
  e.g. Kuklinski, Cobb, and Gilens (1997)
- Who votes for anti-abortion policies?
  Rosenfeld, Imai, and Shapiro (2015)
- Who supports and joins insurgent groups?
  e.g. Lyall, Blair, and Imai (2013); Lyall, Zhou, and Imai (2019)
- How much vote-buying occurs in an election?
  Gonzalez-Ocantos et al. (2011)
- What are the rates of risky sexual behavior among college students?
  e.g. LaBrie and Earleywine (2010)
- What are the rates of illegal hunting?
  Chang et al. (2017)

Sensitive questions are widely asked in survey research

Cannot ask direct questions when there are **incentives to conceal sensitive responses**

1. Privacy concerns
2. Social desirability
3. Physical retaliation
4. Legal jeopardy

> Ethical and Empirical (refusal to participate, deceptive responses) concerns.

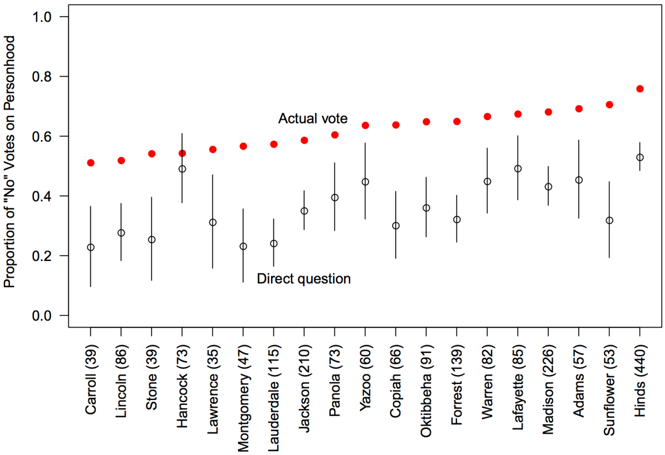
Problems with using direct questions

- **> 50%** refusal rate for Afghanistan Nationwide Quarterly Assessment Research (ANQAR)
  RAND (2011)

- Estimated rate of vote buying in Nicaragua from direct survey item: **2.4%**
  Using indirect survey methods: **24.3%**
  Gonzalez-Ocantos et al. (2011)

- Proportion of civilians who collaborate with militants in Nigeria: **< 10%**
  Using indirect survey methods: **26%**
  Blair, Imai, and Zhou (2015)

Direct Question Example: Underestimating "No" votes on Personhood



Direct question on voting No for "Personhood Amendment" in Mississippi 2011 General Election underestimates actual vote share by > 20 percentage points.

Rosenfeld, Imai, and Shapiro (2015)

How can we elicit truthful answers to sensitive questions?

How can we ask sensitive research questions while protecting individual responses?

Indirect survey experimental methods that obscure the truthful response of individuals

1. **List experiment (Item Count technique)** Aggregation
2. **Endorsement experiment** Evaluation bias
3. **Randomized response** Random noise

Statistical methods efficiently recover underlying responses, multivariate analysis, compare methods, power analysis

Research agenda with Graeme Blair, Kosuke Imai, Yuki Shiraito, Bryn Rosenfeld, Jason Lyall, Will Bullock, Bethany Park, Kenneth Greene, Jacob Shapiro, Winston Chou, Alexander Coppock, Margaret Moor, and many others.

Roadmap

1. Empirical example: 2016 RCT in Kandahar, Afghanistan
2. Overview of List, Endorsement, and Randomized Response techniques
3. Statistical methods to conduct multivariate analysis
4. Comparing methods
5. Additional Resources

Example: Can Economic Assistance Shape Combatant Support in Wartime? Experimental Evidence from Afghanistan


1. Can economic interventions affect support for the Taliban vs. government?
2. RCT in Kandahar, Afghanistan with 2,579 at risk youth
3. Waitlist, factorial design to assess effects of a **skills training program** and/or **unconditional cash transfers**

Joint work with Jason Lyall, Kosuke Imai, and Mercy Corps. Forthcoming in *American Political Science Review.*

List experiment design

I'm going to read you a list with different actions that you could take in your daily life. After I read the entire list, I would like you to tell me how many of these actions you would be willing to do. Please don't tell me which ones you would be willing to do, only tell me how many of these actions you would be willing to do.

Control group

1. Pay additional taxes
2. Report corrupt government officials
3. Enlist in the Afghan National Security Forces

How many, if any, of these actions you would be willing to do?

List experiment design

I'm going to read you a list with different actions that you could take in your daily life. After I read the entire list, I would like you to tell me how many of these actions you would be willing to do. Please don't tell me which ones you would be willing to do, only tell me how many of these actions you would be willing to do.

Control group

1. Pay additional taxes
2. Report corrupt government officials
3. Enlist in the Afghan National Security Forces

Treatment group

1. Pay additional taxes to support the government
2. Report corrupt government officials
3. Enlist in the Afghan National Security Forces
4. Share information about the government with the Taliban

How many, if any, of these actions you would be willing to do?

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.

Mean count in Treatment Group = Mean count in Control Group + **Proportion who will share info to Taliban**

Unbiased, standard difference-in-means estimator:

$$\hat{\tau} = \frac{1}{N_1}\sum_{i=1}^{N} T_i Y_i - \frac{1}{N_0}\sum_{i=1}^{N}(1 - T_i) Y_i$$

Design Considerations

- Privacy is NOT protected if respondents' truthful answers are yes (or no) for all sensitive and non-sensitive items
- Consider ceiling and floor effects (Blair and Imai, 2011 to detect this)
- Less efficient than direct questions
- A larger number of non-sensitive items results in a higher variance
- Negative correlation across non-sensitive items is desirable

Joint Distribution allows us to extract more information

Define a “type” for each respondent by $(Y_i(0), Z_{i,J=1})$

- $Y_i(0)$: total number of Yes responses for non-sensitive items $\{0, 1, ..., J\}$
- $Z_{i,J=1}$: truthful answer to the sensitive item $\{0, 1\}$
- e.g. type (2, 1) means *i* would have 2 non-sensitive items and the sensitive item

Total of $(2x(J+1))$ types

Joint Distribution allows us to extract more information

Joint distribution is identified:

$$Pr(type = (y, 1)) = Pr(Y_i \leq y | T_i = 0) - Pr(Y_i \leq y | T_i = 1)$$

$$Pr(type = (y, 0)) = Pr(Y_i \leq y | T_i = 1) - Pr(Y_i < y | T_i = 0)$$

Our example with $J = 3$ non-sensitive items:

Response Y_i	Treatment Group ($T_i = 1$)	Control Group ($T_i = 0$)
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)

Joint Distribution allows us to extract more information

Joint distribution is identified:

$$Pr(type = (y, 1)) = Pr(Y_i \leq y | T_i = 0) - Pr(Y_i \leq y | T_i = 1)$$

$$Pr(type = (y, 0)) = Pr(Y_i \leq y | T_i = 1) - Pr(Y_i < y | T_i = 0)$$

Our example with $J = 3$ non-sensitive items:

Response Y_i	Treatment Group ($T_i = 1$)	Control Group ($T_i = 0$)
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)~~

Joint Distribution allows us to extract more information

Joint distribution is identified:

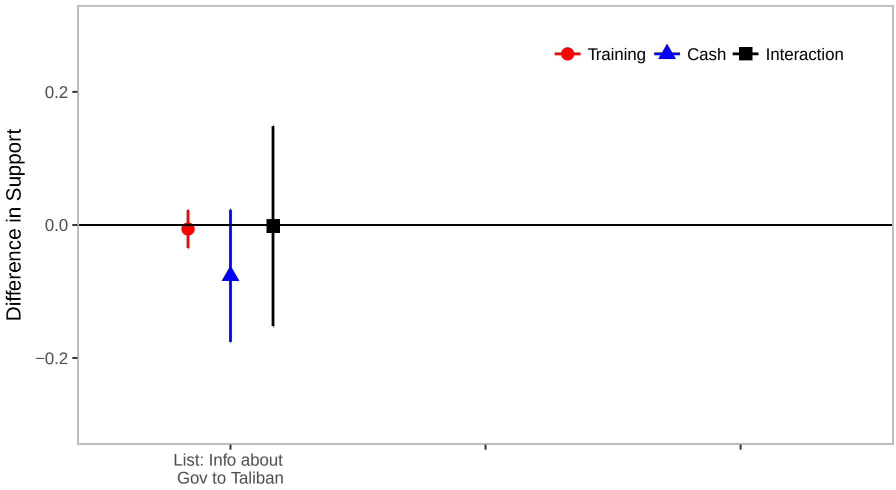
$$Pr(type = (y, 1)) = Pr(Y_i \leq y | T_i = 0) - Pr(Y_i \leq y | T_i = 1)$$

$$Pr(type = (y, 0)) = Pr(Y_i \leq y | T_i = 1) - Pr(Y_i < y | T_i = 0)$$

Our example with $J = 3$ non-sensitive items:

Response Y_i	Treatment Group $(T_i = 1)$	Control Group $(T_i = 0)$
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)~~

Results for Afghanistan RCT Study



Endorsement experiment design

Control group

It has recently been suggested
by the Government of
Afghanistan

that expensive new religious schools be constructed in every district to help provide more opportunities to attend religious schools. How strongly would you support this policy?

1 I strongly oppose this policy
2 I somewhat oppose this policy
3 I am indifferent to this policy
4 I somewhat support this policy
5 I strongly support this policy

Refused
Don't know

Endorsement experiment design

Control group

It has recently been suggested by the Government of Afghanistan

Treatment group

It has recently been suggested by the Taliban

that expensive new religious schools be constructed in every district to help provide more opportunities to attend religious schools. How strongly would you support this policy?

1 I strongly oppose this policy
2 I somewhat oppose this policy
3 I am indifferent to this policy
4 I somewhat support this policy
5 I strongly support this policy

Refused
Don't know

Endorsement experiment design

Multiple Policies to improve power:

1. constructing religious schools
2. strengthen Independent Election Commission (IEC) to prevent electoral fraud
3. allow Office of Oversight for Anti-Corruption to collect info on corrupt government officials
4. remove former mujahedin from high-ranking government positions

> Need to be on the same policy dimension: domestic public policies on addressing corruption and improving welfare.

Identification Assumptions and Interpretation

1. **No Learning:**
   Endorsements have no influence on respondents' interpretation of policy questions.

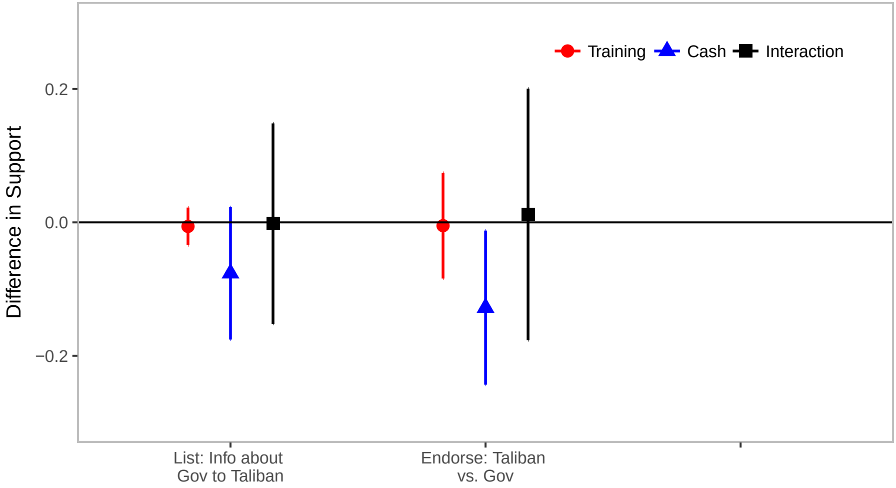
2. All questions occupy a **Single Policy Dimenson** (Shiraito and Imai, 2014: *discrimination parameter* to verify this).

3. **Endorsements are credible**

Response in Treatment Group = Response in Control Group (Policy Preference) + **Endorsement Effect**

To combine responses across policy questions, use IRT model to obtain a single support measure (Bullock, Imai, and Shapiro 2011).

Results for Afghanistan RCT Study



Randomized Response (Forced) design

I want you for each question to spin the spinner twice while my back is turned to you. Remember what you received from the first spin.

If, for the first spin, the arrow lands on the red area, just tell me "no" to the question I ask. If the arrow lands on the green area, just tell me "yes" to the question I ask. But if the arrow lands on either blue area, tell me your true answer to the question.

▸ Other RR designs

Randomized Response (Forced) design

Sensitive Questions

```
Would you be willing...
1.  to share information with the government about the Taliban?
2.  to enlist in the Afghan National Security Forces?
3.  to give money to the Taliban?
4.  to shelter the Taliban in your house?...
```

Identification Assumptions

1. **Randomization Distribution** is known to researcher.
2. **Compliance**
   Actually use the randomization device and comply with the directions (Blair, Imai, and Zhou 2015: design based ways to address noncompliance)

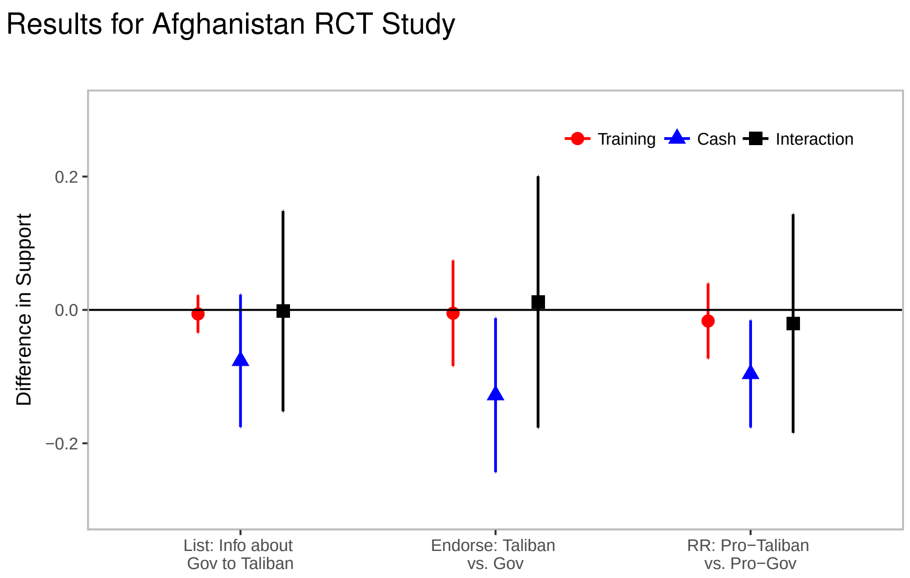
Probability of a 'yes' response is,

$$\Pr(Y_i = 1) = p_1 + (1 - p_1 - p_0) \Pr(Z_i = 1)$$

Probability of truthful 'yes' is,

$$\Pr(Z_i = 1) = \frac{\Pr(Y_i = 1) - p_1}{1 - p_1 - p_0}$$

where Y_i is the observed response, Z_i is the latent response to the sensitive item, R_i is the latent variable for randomization outcome, p_1 and p_0 are $\Pr(R_i = 1)$ and $\Pr(R_i = -1)$, 1/6 and 1/6 in our case.



Results for Afghanistan RCT Study
Training
Cash
Interaction
Difference in Support
0.2
0.0
−0.2
List: Info about Gov to Taliban
Endorse: Taliban vs. Gov
RR: Pro−Taliban vs. Pro−Gov

Additional statistical methods to get more out of these techniques

1. What types of respondents are more likely to have sensitive trait?
2. Can we calculate predicted responses to the sensitive item for each individual?
3. Can the sensitive trait predict other behaviors and attitudes?
4. Can we provide tools for research design?
    - Power analysis
    - Guidance on choosing between designs
    - Detecting violations

Statistical Modeling for Multivariate Analysis

- List: Imai (2011); Blair and Imai (2012) treat $Z_{i,j+1}$ as missing data, model the joint distribution and propose a ML estimator. ▸ Details
  - Regression command in **R**:

```
ictreg(y.variable ~x.variable,
treat = "treatment.variable", data = my.data)
```

- Endorse: Bullock, Imai, Shapiro (2011) use IRT to average over multiple policies and model ideal points and support levels. ▸ Details
- Randomized Response: Blair, Imai, Zhou (2015) treat Z_i as missing data, create a likelihood function generalizeable across all RR designs. ▸ Details

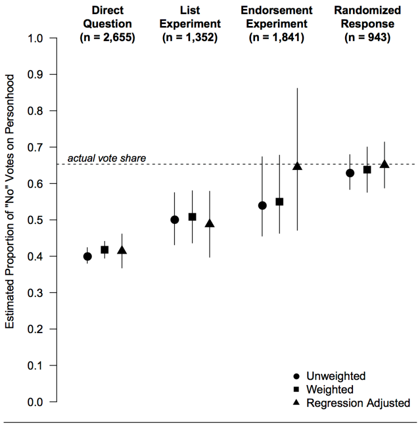
Example: Randomized Response Multivariate Analysis Findings in Nigeria Study

Estimated Logistic Regression Coefficients for Social Connection with Militants

	est.	s.e.
Asset Index	0.079	0.041
Married	−0.267	0.255
Age	−3.528	2.642
Age, Squared	4.099	2.603
Education level	−0.007	0.046
Female	−0.554	0.162
(Intercept)	−0.340	0.509

Respondents who have more household assets and men are substantially more likely to be socially connected to militants.

Validation using Mississippi Study



All techniques improve upon direct question, Randomized Response performs the best.

Rosenfeld, Imai, and Shapiro (2015)

Considerations for Applied Researchers

Tradeoff between direct and indirect methods: Is the topic truly sensitive to justify loss of power (bias-variance tradeoff) and more complex design? (see Blair, Coppock, Moor 2019)

1. List
   - **Pros**: Easy to implement and understand, widely applicable
   - **Cons**: Some individual responses are not protected
   - **Advice**: Need to carefully choose non-sensitive items
2. Endorsement
   - **Pros**: Most indirect questioning, easy to implement and understand
   - **Cons**: Limited applicability, greatest loss of efficiency, difficult to interpret effect magnitudes
   - **Advice**: Need to carefully choose policies
3. Randomized Response
   - **Pros**: Cannot identify individual responses, level of protection is chosen by researcher, many available designs
   - **Cons**: Instructions can be confusing for respondents and enumerators
   - **Advice**: Include a practice question with a non-sensitive question

General Overview and Meta-analysis

- Check out http://sensitivequestions.org/ and http://imai.princeton.edu/projects/sensitive.html
- Blair, Graeme. 2015. "Survey Methods For Sensitive Topics." APSA Comparative Politics Newsletter.
- Blair, Graeme, Alexander Coppock, Margaret Moor. "When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." Working paper.

Open-Source Software available on CRAN and GitHub

- Blair, Graeme, and Kosuke Imai. "`list`: Statistical Methods for the Item Count Technique and List Experiment."
- Shiraito, Yuki, and Kosuke Imai. "`endorse`: R Package for Analyzing Endorsement Experiments."
- Blair, Graeme, Yang-Yang Zhou, and Kosuke Imai. "`rr`: Statistical Methods for the Randomized Response Technique."

Papers that develop methods

- Bullock, Will, Kosuke Imai, and Jacob Shapiro. (2011). "Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan." Political Analysis, Vol. 19, No. 4 (Autumn), pp. 363-384.
- Imai, Kosuke. (2011). "Multivariate Regression Analysis for the Item Count Technique." Journal of the American Statistical Association, Vol. 106, No. 494 (June), pp. 407-416.
- Blair, Graeme and Kosuke Imai. (2012). "Statistical Analysis of List Experiments." Political Analysis, Vol. 20, No. 1 (Winter), pp. 47-77.
- Blair, Graeme, Kosuke Imai, and Jason Lyall. (2014). "Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan." American Journal of Political Science, Vol. 58, No. 4 (October), pp. 1043-1063.

Papers that develop methods

- Imai, Kosuke, Bethany Park, and Kenneth Greene. (2015). "Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models." Political Analysis, Vol. 23, No. 2 (Spring), pp. 180-196.
- Blair, Graeme, Kosuke Imai, and Yang-Yang Zhou. (2015). "Design and Analysis of the Randomized Response Technique." Journal of the American Statistical Association, Vol. 110, No. 511 (September), pp. 1304-1319.
- Blair, Graeme, Winston Chou, and Kosuke Imai. (2019). "List Experiments with Measurement Error." Political Analysis, Vol. 27, No. 4 (October), pp. 455-480.
- Chou, Winston, Kosuke Imai, and Bryn Rosenfeld. "Sensitive Survey Questions with Auxiliary Information." Sociological Methods & Research, Forthcoming.

Papers that describe applications/provide validation

- Lyall, Jason, Graeme Blair, and Kosuke Imai. (2013). "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." American Political Science Review, Vol. 107, No. 4 (November), pp. 679-705.
- Lyall, Jason, Kosuke Imai, and Yuki Shiraito. (2015). "Coethnic Bias and Wartime Informing." Journal of Politics, Vol. 77, No. 3 (July), p. 833-848.
- Rosenfeld, Bryn, Kosuke Imai, and Jacob Shapiro. (2016). "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions." American Journal of Political Science, American Journal of Political Science, Vol. 60, No. 3 (July), pp. 783-802.
- Hirose, Kentaro, Kosuke Imai, and Jason Lyall. (2017). "Can Civilian Attitudes Predict Insurgent Violence?: Ideology and Insurgent Tactical Choice in Civil War." Journal of Peace Research, Vol. 51, No. 1 (January), pp. 47-63.
- Jason Lyall, Yang-Yang Zhou, Kosuke Imai (2019). "Can Economic Assistance Shape Combatant Support in Wartime? Experimental Evidence from Afghanistan." American Political Science Review, Forthcoming.

Thank you. Please send questions/comments to yangyang.zhou@ubc.ca.

Statistical Modeling for List

Setup:

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

- Sub-model for sensitive item: e.g., probit regression

$$\Pr(Z_i^* = 1 \mid X_i) \ = \ \Phi(X_i^\top \delta)$$

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

$$\Pr(Y_i(0) = y \mid X_i, Z_i^* = z) \ = \ J \times \Phi(X_i^\top \psi_z)$$

Maximum likelihood with the EM algorithm or Bayes with MCMC.

Back

Statistical Modeling for Endorsement

Setup:

- T_i: treatment
- Y_i: observed (ordinal) response
- Y_i^*: latent (continuous) response
- V_i^*: latent ideological position
- E_i^*: latent endorsement effect
- X_i: observed covariates

Latent measurement model:

$$Y_i^* \overset{\text{indep.}}{\sim} \mathcal{N}(\beta_j(V_i^* + T_i E_i^*) - \alpha_j, 1)$$

and

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

$$E_i^* \overset{\text{indep.}}{\sim} \mathcal{N}(\gamma^T X_i, \omega^2)$$

Probability of being a “supporter”: $\Pr(Z_i^* > 0 \mid X_i)$

Back

Statistical Modeling for Randomized Response

- Setup:
  - Y_i: observed response
  - Z_i: latent response to the sensitive item
  - R_i: latent variable for randomization outcome
  - X_i: covariates

- The model is,

$$\Pr(Z_i = 1 | X_i) = logit^{-1}(\alpha + \beta^T X_i)$$

- The likelihood function is,

$$\mathcal{L}(\beta | \{X_i, Y_i\}_{i=1}^{n}) = \prod_{i=1}^{N} \{p f_\beta(X_i) + p_1\}^{Y_i} \{1 - (p f_\beta(X_i) + p_1)\}^{1 - Y_i}$$

Maximum likelihood with the EM algorithm.

Back

RR Comparison of Standard Designs

Design	Randomization determines	Pros	Cons
Mirrored Question	Whether answers sensitive item ("I have the sensitive trait") or its inverse ("I do not have the sensitive trait")	Simple implementation	Low respondent confidence in the answer being hidden
Forced Response	Whether answers sensitive item or with forced 'yes' or 'no'	Simple implementation	Respondents with forced 'yes' may fail to say 'yes' due to concern that their response might be interpreted as an affirmative admission to the sensitive item
Disguised Response	Order of red and black cards in two decks of cards. Respondent states the color chosen from the right deck for 'yes' to the sensitive item and the color chosen from the left deck for 'no'	Best for items where even saying 'yes' out loud is sensitive	Complicated randomization device requires in-person implementation
Unrelated Question	Whether answers sensitive item or unrelated, non-sensitive item	High respondent confidence in the answer being hidden	The response to the unrelated question must be either independent of respondent characteristics or modeled

▸ Back