Online Appendix

For "A Bias-Corrected Estimator for the Crosswise Model with Inattentive Respondents"

Yuki Atsusaka and Randolph T. Stevenson

Table of Contents

A Additional Discussion on the Bias-Corrected Estimator

A.1 Derivation of the Bias

Here, we derive the bias in the naïve crosswise estimator based on the argument in Section 3. According to the conventional definition of bias in estimators, we define bias as the (signed) difference between the expected value of the naïve crosswise estimator and the true quantity of interest:

$$
\begin{aligned}
B_{CM} &\equiv \mathbb{E}[\widehat{\pi}_{CM}] - \pi && \text{(A.1a)}\\
&= \mathbb{E}\left[\frac{\widehat{\lambda} + p - 1}{2p - 1}\right] - \frac{\lambda - (1-p)\gamma - \kappa(1-\gamma)}{(2p-1)\gamma} && \text{(A.1b)}\\
&= \frac{\gamma(\lambda + p - 1) - (\lambda - \gamma + p\gamma - \kappa + \kappa\gamma)}{(2p-1)\gamma} && \text{(A.1c)}\\
&= \frac{\lambda\gamma + p\gamma - \gamma - \lambda + \gamma - p\gamma + \kappa - \kappa\gamma}{(2p-1)\gamma} && \text{(A.1d)}\\
&= \frac{\lambda\gamma - \kappa\gamma - \lambda + \kappa}{(2p-1)\gamma} && \text{(A.1e)}\\
&= \frac{\lambda - \kappa}{(2p-1)} - \frac{\lambda - \kappa}{(2p-1)\gamma} && \text{(A.1f)}\\
&= \frac{1}{2}\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) - \frac{1}{2\gamma}\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) && \text{(A.1g)}\\
&= \left(\frac{1}{2} - \frac{1}{2\gamma}\right)\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) && \text{(A.1h)}
\end{aligned}
$$

The second term of the second line was obtained by transforming Equation (1c) as:

$$
\begin{aligned}
\lambda &= \Big\{\pi p + (1-\pi)(1-p)\Big\}\gamma + \kappa(1-\gamma) && \text{(A.2a)}\\
\Rightarrow \Big\{\pi p + (1-\pi)(1-p)\Big\}\gamma &= \lambda - \kappa(1-\gamma) && \text{(A.2b)}\\
\Rightarrow \Big\{\pi p + 1 - p - \pi + \pi p\Big\}\gamma &= \lambda - \kappa(1-\gamma) && \text{(A.2c)}\\
\Rightarrow \Big\{2\pi p - \pi + 1 - p\Big\}\gamma &= \lambda - \kappa(1-\gamma) && \text{(A.2d)}\\
\Rightarrow \Big\{\pi(2p-1) + 1 - p\Big\}\gamma &= \lambda - \kappa(1-\gamma) && \text{(A.2e)}\\
\Rightarrow \pi(2p-1)\gamma + (1-p)\gamma &= \lambda - \kappa(1-\gamma) && \text{(A.2f)}\\
\Rightarrow \pi(2p-1)\gamma &= \lambda - (1-p)\gamma - \kappa(1-\gamma) && \text{(A.2g)}\\
\Rightarrow \pi &= \frac{\lambda - (1-p)\gamma - \kappa(1-\gamma)}{(2p-1)\gamma} && \text{(A.2h)}\\
& && \text{(A.2i)}
\end{aligned}
$$

A.2 Behavior of the Bias

By definition, the bias vanishes when the proportion of attentive respondents is 1 ($\gamma = 1$). To see this, simply observe the following limit:

$$\lim_{\gamma \to 1} \left(\frac{1}{2} - \frac{1}{2\gamma}\right)\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) \tag{A.3a}$$
$$= \left(\frac{1}{2} - \frac{1}{2}\right)\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) \tag{A.3b}$$
$$= 0 \tag{A.3c}$$

In contrast, as the proportion of attentive respondents approaches 0 (from the side of 1), the bias term explodes and approaches positive infinity. To see this, observe that the multiplier $\left(\frac{1}{2} - \frac{1}{2\lambda}\right)$ is always negative and the multiplicand $\left(\frac{\lambda-\kappa}{p-\frac{1}{2}}\right)$ is also negative under some regularity conditions. These conditions are that $\pi < 0.5$, and $p < \frac{1}{2}$ (and thus $\lambda > \kappa$). These regularity conditions hold in most surveys that use the crosswise model. Since the absolute value of the multiplier grows as λ approaches 0, the bias term increases as the proportion of attentive responses decreases.

However, the limit itself does not exist as:

$$\lim_{\lambda \to 0} \left(\frac{1}{2} - \frac{1}{2\gamma}\right)\left(\frac{\lambda - \kappa}{p - \frac{1}{2}}\right) \tag{A.4a}$$
$$= \text{Undefined} \tag{A.4b}$$

A.3 Simulation of the Bias

Figure A.1 visualizes the properties of the bias in the conventional crosswise estimator (assuming $\kappa = 0.5$). It shows that the size of the bias increases as (1) the percentage of inattentive respondents increases and (2) the quantity of interest approaches 0, but that it does not change regardless of the value of p.

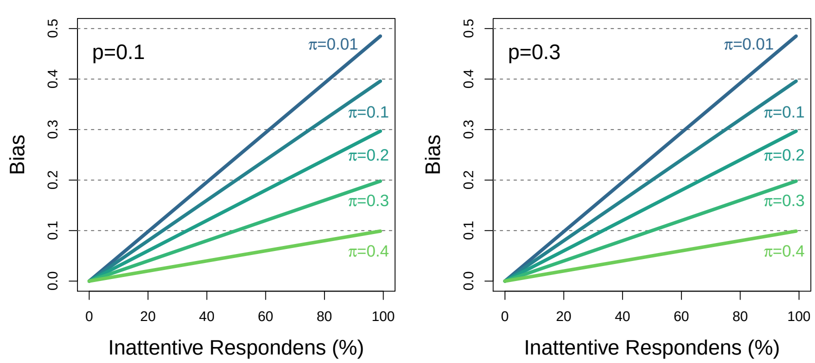

Figure A.1: **Bias in the Naïve Crosswise Estimator.** *Note*: Both panels display the (theoretical) bias in the conventional crosswise estimator with varying levels of inattentive respondents.

A.4 Derivation of the Variance

Here, we derive the population and sample variance of the bias-corrected crosswise estimator discussed in Section 3.

$$\mathbb{V}(\widehat{\pi}_{BC}) = \mathbb{V}\left[\widehat{\pi}_{CM} - \left(\frac{1}{2} - \frac{1}{2\widehat{\gamma}'}\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p - \frac{1}{2}}\right)\right] \tag{A.5a}$$

$$= \mathbb{V}\left[\widehat{\pi}_{CM} - \left(\frac{1}{2} - \frac{1}{2}\left[\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right]\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p - \frac{1}{2}}\right)\right] \tag{A.5b}$$

$$= \mathbb{V}\left[\frac{\widehat{\lambda} + p - 1}{2p - 1} - \left(\frac{1}{2} - \frac{1}{2}\left[\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right]\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p - \frac{1}{2}}\right)\right] \tag{A.5c}$$

$$= \mathbb{V}\left[\frac{\widehat{\lambda} + p - 1}{2p - 1} - \frac{1}{2}\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p - \frac{1}{2}}\right) + \frac{1}{2}\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p - \frac{1}{2}}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5d}$$

$$= \mathbb{V}\left[\frac{\widehat{\lambda} + p - 1}{2p - 1} - \frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1} + \left(\frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5e}$$

$$= \mathbb{V}\left[\frac{p - \frac{1}{2}}{2p - 1} + \left(\frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5f}$$

$$= \mathbb{V}\left[\frac{p - \frac{1}{2}}{2p - 1}\right] + \mathbb{V}\left[\left(\frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5g}$$

$$= 0 + \mathbb{V}\left[\left(\frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5h}$$

$$= \mathbb{V}\left[\left(\frac{\widehat{\lambda} - \frac{1}{2}}{2p - 1}\right)\left(\frac{\frac{1}{2} - p'}{\widehat{\lambda}' - \frac{1}{2}}\right)\right] \tag{A.5i}$$

$$= \mathbb{V}\left[\left(\frac{\widehat{\lambda} - \frac{1}{2}}{\widehat{\lambda}' - \frac{1}{2}}\right)\left(\frac{\frac{1}{2} - p'}{2p - 1}\right)\right] \tag{A.5j}$$

$$= \mathbb{V}\left[\left(\frac{\widehat{\lambda}}{\widehat{\lambda}'}\right)\left(\frac{\frac{1}{2} - p'}{2p - 1}\right)\right] \tag{A.5k}$$

$$\tag{A.5l}$$

Similarly,

$$\widehat{\mathbb{V}}(\widehat{\pi}_{BC}) = \widehat{\mathbb{V}}\left[\left(\frac{\widehat{\lambda}}{\widehat{\lambda}'}\right)\left(\frac{\frac{1}{2} - p'}{2p - 1}\right)\right] \tag{A.5m}$$

$$\tag{A.5n}$$

A.5 Unbiasedness of the Naïve Estimator

To see that the naïve estimator is unbiased when $\gamma = 1$, let Y_i be a binary random variable denoting whether respondent i chooses the crosswise item (i.e., TRUE-TRUE or FALSE-FALSE) and its realization $y_i \in \{0, 1\}$. Let the number of respondents choosing the crosswise item be $k = \sum_{i=1}^{N} y_i$, where $k < n$. Then, the likelihood function for λ given any observed k is $L(\lambda|n, k) = \binom{n}{k}\lambda^k(1 - \lambda)^{n-k}$. Applying the first-order condition yields a maximum likelihood estimate (MLE) of λ, $\widehat{\lambda} = \frac{k}{n}$, where $\mathbb{E}[\widehat{\lambda}] = \lambda$. The unbiasedness follows from the fact that $\mathbb{E}[\widehat{\lambda}] = \mathbb{E}\left[\frac{k}{n}\right] = \frac{1}{n}\mathbb{E}[k] = \frac{1}{n}n\lambda = \lambda$. Following the parameterization invariance property of MLEs, $\mathbb{E}[\widehat{\pi}_{CM}]$ = π. This result, however, does not hold when $\gamma \neq 1$.

A.6 Unbiasedness of $\widehat{\lambda}'$

To see that $\widehat{\lambda}'$ is an unbiased estimator of λ', let us define that $\widehat{\lambda}'$ is a binomial random variable (like $\widehat{\lambda}$) with parameters n, λ' and $\widehat{\lambda}' = k'/n$, where k' is the number of people who choose the crosswise item in the anchor question. This is because $k' \sim \text{Binom}(n, \lambda')$ and $\widehat{\lambda}' = k'/n$ suggests $\widehat{\lambda}' \sim \text{Binom}(n, k')$. The probability mass function that $\widehat{\lambda}'$ taking n'/n is given by Pr$(\widehat{\lambda}' = \frac{n'}{n}) = \binom{n}{n'}(k')^{n'}(1-k')^{n-n'}$.

A.7 Unbiasedness of the Proposed Estimator

Here, we offer proof that the proposed estimator is an unbiased estimator of the quantity of interest under two assumptions.

$$
\begin{aligned}
\mathbb{E}[\widehat{\pi}_{BC}] &= \mathbb{E}\left[\widehat{\pi}_{CM} - \left(\frac{1}{2} - \frac{1}{2}\left[\frac{\frac{1}{2}-p'}{\widehat{\lambda}'-\frac{1}{2}}\right]\right)\left(\frac{\widehat{\lambda}-\frac{1}{2}}{p-\frac{1}{2}}\right)\right] && \text{(A.6a)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \mathbb{E}\left[\left(\frac{1}{2} - \frac{1}{2}\left[\frac{\frac{1}{2}-p'}{\widehat{\lambda}'-\frac{1}{2}}\right]\right)\left(\frac{\widehat{\lambda}-\frac{1}{2}}{p-\frac{1}{2}}\right)\right] && \text{(A.6b)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \mathbb{E}\left[\left(\frac{1}{2} - \frac{1}{2\widehat{\gamma}'}\right)\left(\frac{\widehat{\lambda}-\frac{1}{2}}{p-\frac{1}{2}}\right)\right] && \text{(A.6c)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \left(\frac{1}{2} - \frac{1}{2\mathbb{E}[\widehat{\gamma}']}\right)\left(\frac{\mathbb{E}[\widehat{\lambda}]-\frac{1}{2}}{p-\frac{1}{2}}\right) && \text{(A.6d)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \left(\frac{1}{2} - \frac{1}{2\gamma'}\right)\left(\frac{\mathbb{E}[\widehat{\lambda}]-\frac{1}{2}}{p-\frac{1}{2}}\right) && \text{(A.6e)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \left(\frac{1}{2} - \frac{1}{2\gamma}\right)\left(\frac{\mathbb{E}[\widehat{\lambda}]-\frac{1}{2}}{p-\frac{1}{2}}\right) \quad \text{(by Assumption 2)} && \text{(A.6f)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - \left(\frac{1}{2} - \frac{1}{2\gamma}\right)\left(\frac{\lambda-\frac{1}{2}}{p-\frac{1}{2}}\right) \quad \text{(by Assumption 3)} && \text{(A.6g)}\\
&= \mathbb{E}[\widehat{\pi}_{CM}] - B_{CM} && \text{(A.6h)}\\
&= \pi \quad \text{(by the definition of the bias)} && \text{(A.6i)}
\end{aligned}
$$

B Discussions and Evaluations of Alternative Estimators

In this appendix section, we describe two other methods for adjusting crosswise estimates for bias resulting from respondent inattention. In the first subsection we describe these methods a explain why we should only expect them to be useful under ideal conditions unlikely to be met in most settings in which the crosswise model would be used. Next, we present a set of simulation based comparison of the bias, root mean squared error, and coverage of the 95% confidence intervals for these estimators against our own.

B.1 Description of Alternative Estimators Estimator

The first suggestion for correcting bias due to inattention in the crosswise estimator was by Enzmann (2017), who made the suggestion in a book chapter on crosswise methods, but provided no discussion or justification for his proposal. His proposal begins with the formula below and then suggests estimating the unknown proportion of inattentive respondents by asking, in an unprotected direct question, whether the respondent answered the crosswise question randomly. It's unclear from his brief treatment whether this should be done separately for each crosswise question in a survey or only once for the survey as a whole.

$$\frac{\widehat{\pi}_{CM} - 0.5r}{1 - r} \tag{B.1a}$$

where $\widehat{\pi}_{CM}$ is the naive crosswise estimator and r is the population proportion of inattentive respondents.

In a later paper, Schnapp (2019, 311) reproduces Enzmanns formula and then leans into the idea of estimating the percentage of random responders in this equation via direct questioning. Further, he adds an equation for the variance of the estimator. Below, we call this the Enzmann-Schnapp estimator, or just "ES."

Besides describing the Enzmann method, Schnapp (2019) also presents his own proposal for a bias-corrected estimator, which he calls the CMR-I. That estimator also asks respondents if they answered the crosswise question(s) randomly and then uses that answer to adjust individual answers to the crosswise questions accordingly (e.g., changing the answers for respondents that report answering randomly from either the recorded 0 or 1 to 0.5 and then using all the adjusted and unadjusted responses together to calculate $\widehat{\lambda}$ (the proportion of respondents choosing "both or neither is true") in the usual crosswise formula). He then applies the usual variance estimator for crosswise models (not the one he proposed for Enzmann's estimator) to this estimator.

The main difference in our estimator and the ES and CMR-I estimators is that, while we use anchor questions to estimate the share of inattentive respondents, these estimators both do so by directly asking respondents whether they answered the crosswise question(s) randomly. And, of course, the CMR-I estimator also differs more fundamentally in its overall construction. Likewise, Schnapp's variance estimate for this estimator is derived under the assumption that the percentage of random responders is known and so is not appropriate for the full estimator he actually describes (and uses), which includes the use of direct questioning about inattention.

Importantly, in both the ES and CMR-I estimators, the direct question about random responding is asked without the protection of the crosswise format and so it is likely that at least some respondents will hesitate to admit to that socially-undesirable behavior. Thus, in our simulations below, we present results for both The ES and CMR-I estimators with varying levels of veracity for the direct question about random guessing.

B.2 Comparative Evaluation of the Estimators

To perform our comparative evaluation, we replicate the simulations presented in Figure 1. To reiterate, in each simulation, we draw π from continuous uniform distribution (0.1, 0.45), p and p' from continuous uniform distribution (0.088, 0.333) (reflecting the smallest and largest values in existing studies), and γ from continuous uniform distribution (0.5, 1). Finally, we repeat the set of experiments for different sample sizes of 200, 500, 1000, 2000, and 5000 and evaluate the results. To include the ES and CMR-I estimators in the simulations, we simulate direct questioning that asks whether respondents give random answers in the crosswise question with three different percentages of "liars" in the direct question: 0%, 25%, and 50%. We call these conditions "No liars", "25% liars", and "50% liars", respectively. When simulating respondent answers to the direct question about whether they answered randomly, we also make the following assumption:

Assumption B.1 (One-side Lying). *Only inattentive respondents may lie about giving random answers in the crosswise questions.*

In other words, we assume that attentive respondents always report that they did not give random answers in the crosswise model.

Figure B.1 report the results of our comparative assessments. It suggests that the CMR-I estimator does not work at all and the ES estimator only works under the no lair condition, which is hard to satisfy. Additionally, the ES estimator appears to have a wider coverage of the ground truth than necessary (i.e., their 95% confidence intervals appear to cover the ground truth almost 100% of the time). In contrast, the 95% confidence interval of our estimator covers the ground truth 95% of the times as expected.

If one focuses only on the bias and MSE results, it is clear that the our estimator and the ES estimator produce identical results under the No Liars condition. Indeed, in Section B.3 below, we show that if (and only if) that assumption (and one other) holds, the ES estimator is algebraically equivalent to ours. Of course, this key assumption is unlikely to hold in any realistic setting.

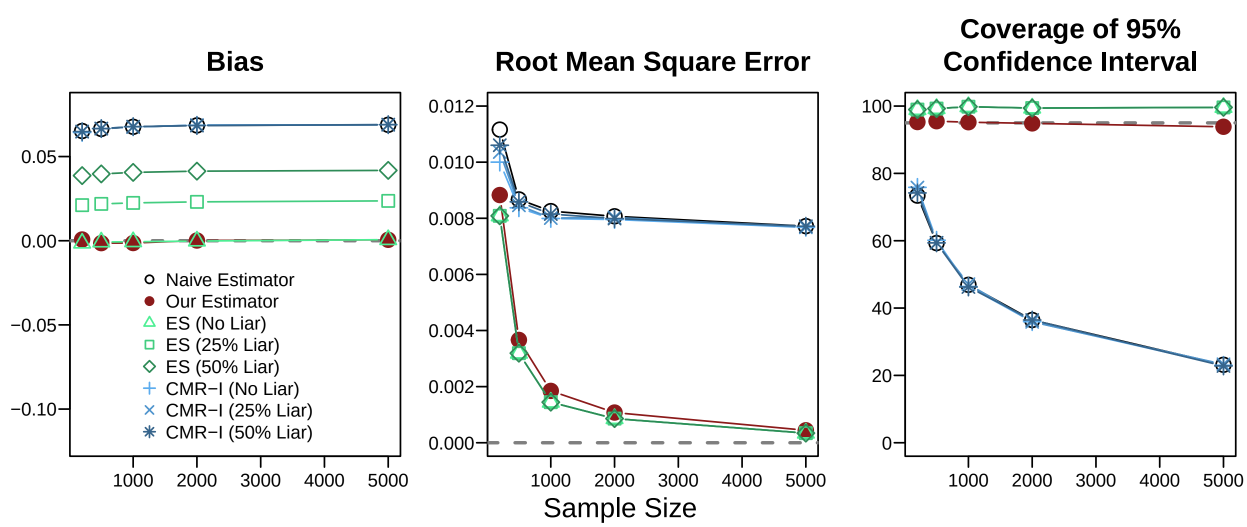

Figure B.1: **Finite Sample Performance of the Naïve, Bias-Corrected, and Enzmann-Schnapp Estimators.** *Note*: This figure displays the bias, root-mean-square error, and the coverage of 95% confidence interval of naïve estimator, our estimator, Enzmann-Schanpp correction (No liars), Enzmann-Schanpp correction (25% liars), Enzmann-Schanpp correction (50% liars), CMR-I (No liars), CMR-I (25% liars), and CMR-I (50% liars).

To further investigate the relative efficiency of our estimator vs. the ES in its most favorable condition (i.e., no lair), Figure B.2 shows the relative length of the 95% confidence interval of our estimator (left panel) and the ES (right panel) to that of the conventional estimator. The figure illustrates that regardless of the sample size our estimator has a narrower confidence interval (about 1.3 times larger than the conventional crosswise confidence interval) than does the ES (about 4 times larger than the conventional crosswise confidence interval).

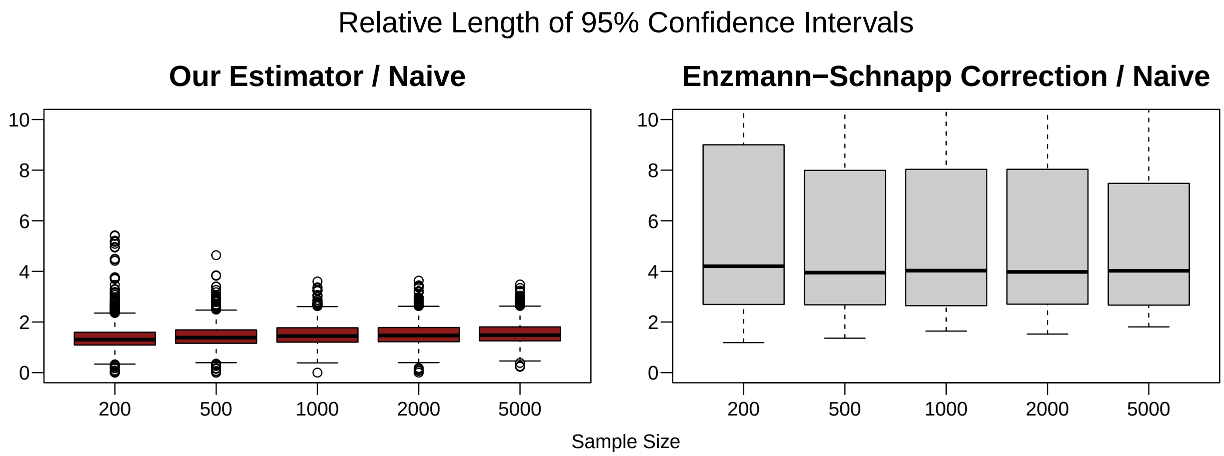

Figure B.2: **Relative Lengths of 95% Confidence Interval with respect to the Naïve estimator.** *Note*: This figure displays the length of the 95% confidence intervals of each estimator relative to that of the conventional estimator on the y-axis. For example, for 200 observations, the confidence interval of the Enzmann-Schnapp correction is about 4 times larger than that of the naïve estimator.

B.3 Proof that Equation (3d) and Equation (B.1.a) are Equivalent under Several Conditions

In this section we show that our bias-corrected estimator is equivalent to the Enzmann-Schnapp correction if and only if we assume that r is an *estimated* proportion of inattentive respondents (i.e., $r = 1 - \widehat{\gamma}$), Assumption 1 holds (i.e., $\kappa = 0.5$), and the proportion of inattentive respondents is estimated in the same way using the anchor question. To prove the equivalence, we show that the difference between the two estimators is zero under these conditions.

$$
\begin{aligned}
\widehat{\pi}_{\text{Schnapp-Enznmann}} - \widehat{\pi}_{\text{BC}} &= \frac{\widehat{\pi}_{CM} - 0.5r}{1-r} - \left[\widehat{\pi}_{CM} - \widehat{B}_{CM}\right] && \text{(B.2a)}\\
&= \frac{\widehat{\pi}_{CM} - \frac{1}{2}(1-\widehat{\gamma})}{\widehat{\gamma}} - \widehat{\pi}_{CM} + \widehat{B}_{CM} && \text{(B.2b)}\\
&= \frac{\widehat{\pi}_{CM}}{\widehat{\gamma}} - \frac{1}{2\widehat{\gamma}}(1-\widehat{\gamma}) - \widehat{\pi}_{CM} + \widehat{B}_{CM} && \text{(B.2c)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\widehat{\pi}_{CM} - \left(\frac{1}{\widehat{\gamma}} - 1\right)\frac{1}{2} + \widehat{B}_{CM} && \text{(B.2d)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\widehat{\pi}_{CM} - \frac{1}{2}\right) + \widehat{B}_{CM} && \text{(B.2e)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\frac{\widehat{\lambda} + p - 1}{2p-1} - \frac{1}{2}\right) + \widehat{B}_{CM} && \text{(B.2f)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\frac{\widehat{\lambda} + p - 1}{2(p-\frac{1}{2})} - \frac{1}{2}\right) + \widehat{B}_{CM} && \text{(B.2g)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\frac{\widehat{\lambda} + p - 1 - (p-\frac{1}{2})}{2(p-\frac{1}{2})}\right) + \widehat{B}_{CM} && \text{(B.2h)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\frac{\widehat{\lambda} + p - 1 - p + \frac{1}{2}}{2(p-\frac{1}{2})}\right) + \widehat{B}_{CM} && \text{(B.2i)}\\
&= \left(\frac{1}{\widehat{\gamma}} - 1\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{2(p-\frac{1}{2})}\right) + \widehat{B}_{CM} && \text{(B.2j)}\\
&= \left(\frac{1}{2\widehat{\gamma}} - \frac{1}{2}\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p-\frac{1}{2}}\right) + \widehat{B}_{CM} && \text{(B.2k)}\\
&= -\left(\frac{1}{2} - \frac{1}{2\widehat{\gamma}}\right)\left(\frac{\widehat{\lambda} - \frac{1}{2}}{p-\frac{1}{2}}\right) + \widehat{B}_{CM} && \text{(B.2l)}\\
&= -\widehat{B}_{CM} + \widehat{B}_{CM} && \text{(B.2m)}\\
&= 0 && \text{(B.2n)}
\end{aligned}
$$

C Additional Discussion and Simulations for Extensions

In this section, we present additional information for the proposed extensions of the bias-corrected estimator.

C.1 Sensitivity Analysis

Here, we present the full results of our sensitivity analysis. We applied our sensitivity analysis to all existing studies (49 estimates reported by 21 unique articles) for which we were able to collect essential information. For consistency, if the original estimate ($\widehat{\pi}$) is between 0.5 and 1 (where we expect over-reporting), we flip the direction of the estimate so that it will be between 0 and 0.5 (where we expect under-reporting). We did not apply this transformation for Figure 5. The featured studies are from Coutts et al. (2011); Jann, Jerke and Krumpal (2012); Kundt, Misch and Nerré (2013); Korndörfer, Krumpal and Schmukle (2014); Shamsipour et al. (2014); Kundt (2014); Hoffmann et al. (2015); Gingerich et al. (2016); Waubert de Puiseau, Hoffmann and Musch (2017); Höglinger and Jann (2018); Nasirian et al. (2018); Kuhn and Vivyan (2018); Banayejeddi et al. (2019); Hoffmann and Musch (2019); Hopp and Speil (2019); Klimas et al. (2019); Meisters, Hoffmann and Musch (2020*a*); Hoffmann, Meisters and Musch (2020); ÖZGÜL (2020); Mieth et al. (2021); Canan et al. (2021).

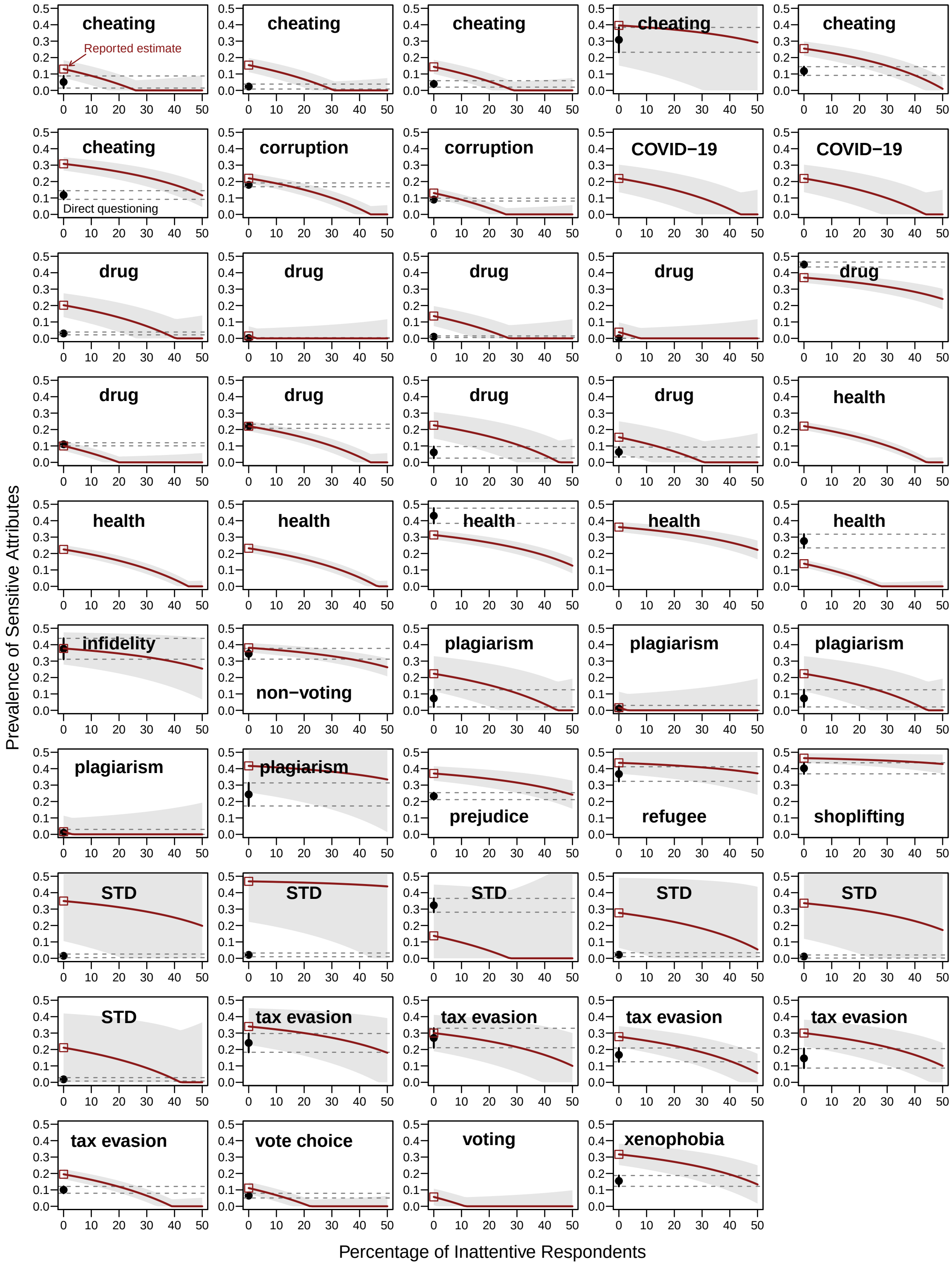

Figure C.1: **Sensitivity Analysis of Previous Crosswise Estimates.** *Note*: For each estimate, the bias correction is applied with varying percentages of inattentive respondents.

C.2 Weighting

Here, we present our strategy to incorporate sample weights into our bias-corrected estimator. Recall that the only sample statistics we observe in our framework are $\hat{\lambda}$ and $\hat{\lambda}'$, which are observed proportions of respondents choosing the crosswise item in the crosswise and anchor questions, respectively. The key idea here is that we can apply a Horvitz-Thompson-type estimator of the mean (and thus the inverse probability weighting more generally) to the crosswise proportions, where weights are the inverse of the probabilities that respondents in different strata will be in the sample. Namely, we can apply a weight $w_i = \frac{1}{\text{Pr}(S_i=1|\mathbf{X_i})}$, where $S_i = \{0, 1\}$ is a binary variable denoting if respondent i is in the sample and $\mathbf{X_i}$ is a vector of the respondent's background characteristics.[1]

Let $Y_i \in \{0, 1\}$ be a binary variable denoting if respondent i chooses the crosswise item in the crosswise question and $A_i \in \{0, 1\}$ be a binary variable denoting if the same respondent chooses the crosswise item in the anchor question. We then estimate the weighted crosswise proportion in the crosswise question and the weighted proportion of attentive respondents in the following way:

$$\widehat{\lambda}_w = \frac{\sum_{i=1}^n w_i Y_i}{\sum_{i=1}^n w_i} \quad \text{and} \quad \widehat{\gamma}_w = \frac{\frac{\sum_{i=1}^n w_i A_i}{\sum_{i=1}^n w_i} - \frac{1}{2}}{\frac{1}{2} - p'}, \tag{C.1a}$$

The proof is straightforward. Assuming that $Y_i \perp\!\!\!\perp S_i|X$ (choosing the crosswise item and being in the sample are statistically independent conditional upon a covaraite), weighting can recover the population crosswise proportion λ from the sample crosswise response $Y_i S_i$:

$$\begin{aligned}
&\mathbb{E}\left[\frac{Y_i S_i}{\text{Pr}(S_i = 1|X)}\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[\frac{Y_i S_i}{\text{Pr}(S_i = 1|X)}\middle| X\right]\right] \quad \text{(Iterative Expectation)} \\
=&\mathbb{E}\left[\frac{\mathbb{E}[Y_i|X]\mathbb{E}[S_i|X]}{\text{Pr}(S_i = 1|X)}\right] \quad \text{(Conditional Independence)} \\
=&\mathbb{E}\left[\frac{\mathbb{E}[Y_i|X]\text{Pr}(S_i = 1|X)}{\text{Pr}(S_i = 1|X)}\right] \quad \text{(Definition of Expectation)} \\
=&\mathbb{E}[\mathbb{E}[Y_i|X]] \\
=&\mathbb{E}[Y_i] \quad \text{(Iterative Expectation)} \\
=&\lambda
\end{aligned}$$

[1] In this article, we only consider the base weight, but one can naturally include other weights such as non-response weights to construct the final survey weights.

Similarly, we can show that

$$\begin{aligned}
&\mathbb{E}\left[\frac{A_i S_i}{\Pr(S_i = 1|X)}\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[\frac{A_i S_i}{\Pr(S_i = 1|X)}\middle| X\right]\right] \\
=&\mathbb{E}[A_i] \\
=&\lambda'
\end{aligned}$$

In practice, researchers can calculate weights using their favorite weighting techniques such as raking (or iterative proportional fitting), matching, propensity score weighting, or sequential applications of these. Recent research shows that "when it comes to accuracy, choosing the right variables for weighting is more important than choosing the right statistical method" (Mercer, Lau and Kennedy, 2018, 4). Thus, we recommend that researchers think carefully about the association between the sensitive attribute of interest and basic demographic and other context-dependent factors when using weighting. To choose the "right" variables, our proposed regression models can also be useful exploratory aids. When generalizing the results on sensitive attributes to a larger population, however, it is strongly advised to elaborate on how weights are constructed and what potential bias may exist (Franco et al., 2017).

Another possible approach to deal with highly selected samples is to employ multilevel regression and post-stratification (MRP) (Downes et al., 2018). While we do not consider MRP with crosswise estimates in this article, future research should explore the optimal strategy to use MRP in sensitive questions.

To illustrate our weighting strategy, we simulate crosswise data with two covariates: $X_1 \sim \text{Binomial}(0.5)$ and $X_2 \sim \text{Poisson}(30)$ for, for example, 100,000 voters. Specifically, we simulate the true prevalence rates in the crosswise and anchor questions according to the following generative models:

$$\begin{aligned}
\pi &= \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)} \\
\gamma &= \frac{\exp(\theta_0 + \theta_1 X_1 + \theta_2 X_2)}{1 + \exp(\theta_0 + \theta_1 X_1 + \theta_2 X_2)},
\end{aligned}$$

where we set $\beta_0 = -1.5$, $\beta_1 = 0.5$, $\beta_2 = 0.02$ and $\theta_0 = 2$, $\theta_1 = -0.1$, $\theta_2 = -0.01$.. For example, we can consider X_1 as a binary indicator for being female (as opposed to non-female) and female voters are more likely to have sensitive traits than non-female voters (i.e., $\beta_1 = 0.5$).

Under this generative model, the population-level proportion of individuals with sensitive traits is **0.35** (and the population proportion is **0.40** for female and **0.29** for non-female voters). Now, from the population of 100,000 voters, we sample 1000, 2000, 3000, 4000, and 5000 individuals. In this process, we intentionally oversample female voters with probability 0.7. Consequently, we obtain sample weights 1.43 for female voters and 3.33 for non-female voters. We then generate bias-corrected crosswise estimates with and without incorporating the sample weights.

Figure C.2 compares bias-corrected crosswise estimates based on simulated unrepresentative samples with and without sample weights. It demonstrates that while the unweighted estimator always overestimates the ground truth, the weighted estimator captures the population-level quantity of interest.

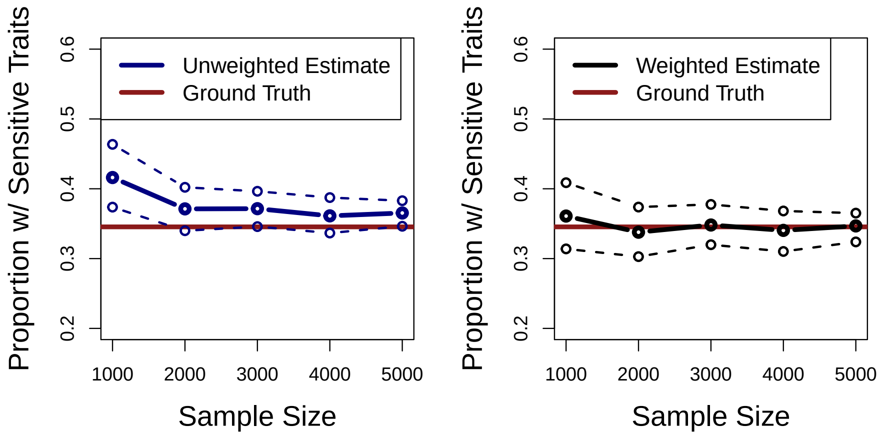

Figure C.2: **Weighting in the Crosswise Estimator.** *Note*: This figure shows bias-corrected crosswise estimates without weighting (left panel) and estimates with weighting (right panel) based on simulated unrepresentative samples.

C.3 Multivariate Regression: Sensitive Attribute as an Outcome

To validate this regression framework, we simulate crosswise data with two covariates: $X_1 \sim \text{Binomial}(0.5)$ and $X_2 \sim \text{Poisson}(30)$. Specifically, we simulate the true prevalence rates in the crosswise and anchor questions according to the following generative models:

$$\pi = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}$$
$$\gamma = \frac{\exp(\theta_0 + \theta_1 X_1 + \theta_2 X_2)}{1 + \exp(\theta_0 + \theta_1 X_1 + \theta_2 X_2)},$$

where we set $\beta_0 = -1.5$, $\beta_1 = 0.5$, $\beta_2 = 0.02$ and $\theta_0 = 2$, $\theta_1 = -0.1$, $\theta_2 = -0.01$..

Finally, we estimate the crosswise regression with the latent sensitive trait as the outcome variable. For estimation, we take a natural log of this likelihood function and maximize it by an iterative maximization method. For inference, we use the negative inverse of the Hessian matrix of the log-likelihood evaluated at the maximum likelihood estimates to compute standard errors of the estimated coefficients. The same applies to the next regression model. Figure C.3 displays the estimated parameters and confidence intervals with different sample sizes. The results suggest that the proposed model and estimation strategy can recover the true parameters (asymptotically). It is also straightforward to compute predicted probabilities of having a sensitive attribute with 95% confidence intervals using the parametric bootstrap. Figure C.4 displays the predicted probabilities of having sensitive attributes in this particular simulation with different values for X_1.

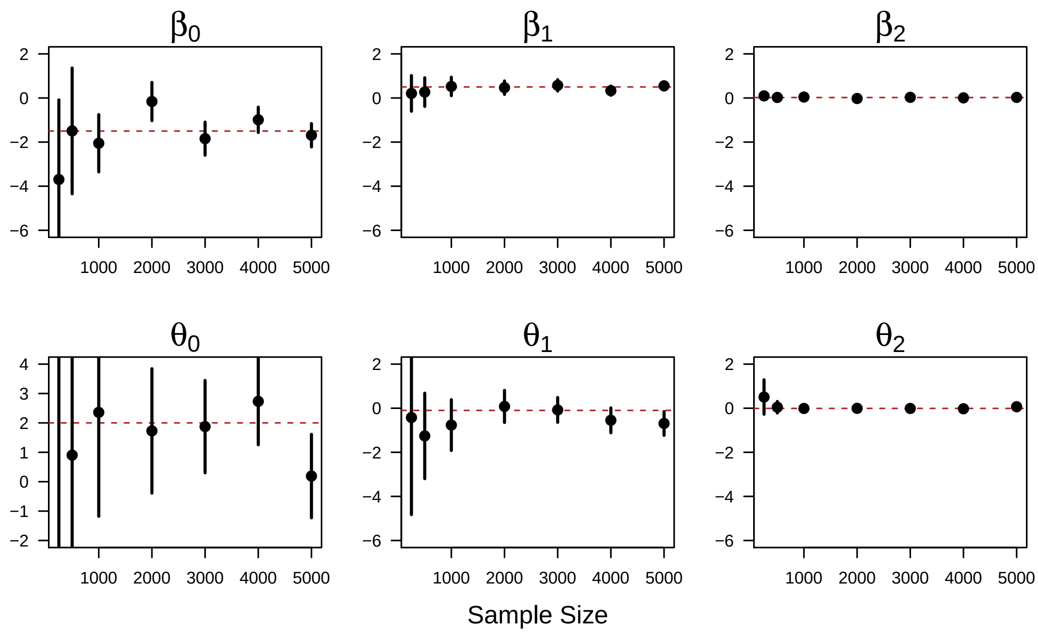

Figure C.3: **Finite Sample Performance of Regression Estimator (Sensitive Trait as a Predictor).** *Note*: Regression estimates of six parameters in simulated data. The dashed lines indicate the true values for the parameters.

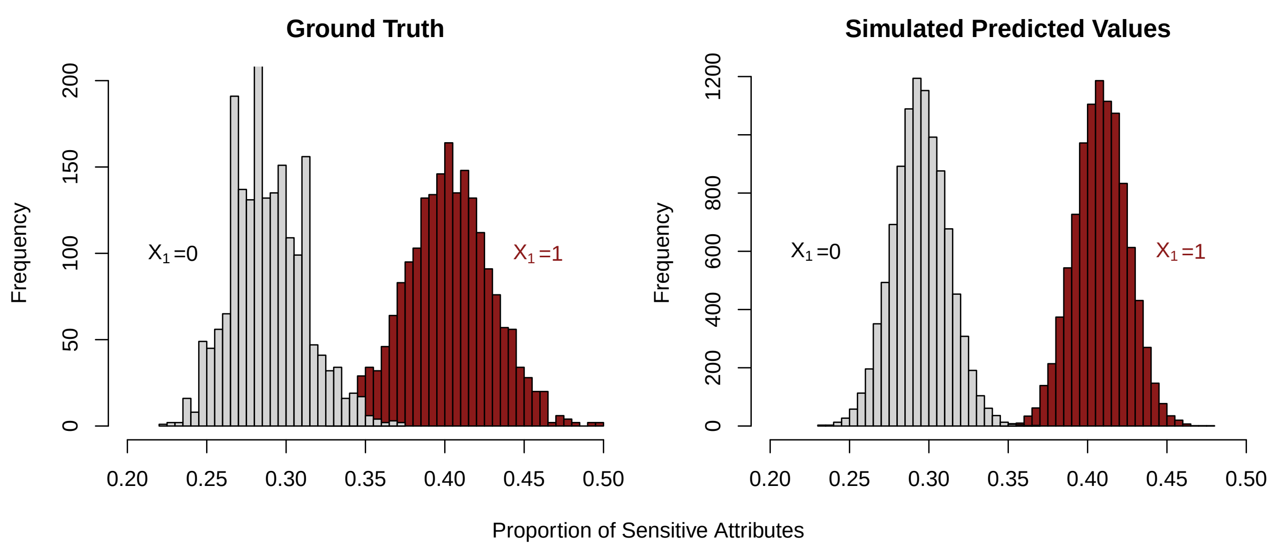

Figure C.4: **True and Predicted Proportions of Sensitive Attributes** *Note*: This graph visualizes the empirical distribution of the probability of having sensitive attributes for $X_1 = 0$ and $X_1 = 1$ based on the ground truth with $N = 4000$ (left panel) and simulated predicted values (right panel).

C.4 Multivariate Regression: Sensitive Attribute as a Predictor

Next, we propose regressions models in which the latent sensitive attribute is used as a predictor or explanatory variable. Let V_i be a continuous or discrete outcome variable for respondent i (other types of outcome variables are also possible). We define the regression model (conditional expectation) of interest as:

$$g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, Z_i), \tag{C.2a}$$

where $\mathbf{\Theta}$ is a vector of parameters that associate a set of predictors plus the indicator for having a sensitive trait ($\mathbf{X_i}$, Z_i) and the response variable (V_i). For example, for a normally distributed outcome variable, we can consider $g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, Z_i) = \mathcal{N}(\alpha + \boldsymbol{\gamma}^\top \mathbf{X_i} + \delta Z_i, \sigma^2)$ with $\mathbf{\Theta} = (\alpha, \boldsymbol{\gamma}, \delta, \sigma^2)$. For a binary response variable, we can consider $g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, Z_i) = \text{Bernoulli}(\phi)$, where $\frac{\phi}{1-\phi} = \alpha + \boldsymbol{\gamma}^\top \mathbf{X_i} + \delta Z_i$ and $\mathbf{\Theta} = (\alpha, \boldsymbol{\gamma}, \delta)$. Our goal is to make inferences about the association between the latent sensitive attribute (Z_i) and the response variable (V_i) after controlling for other covariates (δ is our primary quantity of interest).

Using all the available information from data, the observed data likelihood function then becomes:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{\Theta}|\{V_i, \mathbf{X_i}, Y_i, A_i\}_{i=1}^n, p, p') &= \prod_{i=1}^n g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, Z_i, T_i)\mathbb{P}(Y_i = 1, Z_i, T_i|\mathbf{X_i})\mathbb{P}(A_i = 1, Z_i, T_i|\mathbf{X_i}) \\
&= \prod_{i=1}^n \Bigg\{ g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 1, 1)p^{Y_i}(1-p)^{1-Y_i}\pi_{\boldsymbol{\beta}}(\mathbf{X_i})(1-p')^{A_i}p'^{1-A_i}\gamma_{\boldsymbol{\theta}}(\mathbf{X_i}) \\
&\quad + g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 0, 1)(1-p)^{Y_i}p^{1-Y_i}(1-\pi_{\boldsymbol{\beta}}(\mathbf{X_i}))(1-p')^{A_i}p'^{1-A_i}\gamma_{\boldsymbol{\theta}}(\mathbf{X_i}) \\
&\quad + g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 1, 0)\frac{1}{2}\pi_{\boldsymbol{\beta}}(\mathbf{X_i})\frac{1}{2}(1-\gamma_{\boldsymbol{\theta}}(\mathbf{X_i})) \\
&\quad + g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 0, 0)\frac{1}{2}(1-\pi_{\boldsymbol{\beta}}(\mathbf{X_i}))\frac{1}{2}(1-\gamma_{\boldsymbol{\theta}}(\mathbf{X_i})) \Bigg\},
\end{aligned} \tag{C.2b}$$

where each part inside the bracket is $g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, z, t)\mathbb{P}(Y_i = 1|Z_i = z, T_i = t)\mathbb{P}(Z_i = z|\mathbf{X_i})\mathbb{P}(A_i = 1|Z_i = z, T_i = t)\mathbb{P}(T_i = 1|\mathbf{X_i})$, where $z = \{0, 1\}$ and $t = \{0, 1\}$.

We assume that Assumptions 1-3 hold as well as $V_i \perp\!\!\!\perp Y_i|\mathbf{X_i}$, $V_i \perp\!\!\!\perp A_i|\mathbf{X_i}$, and $Y_i \perp\!\!\!\perp A_i|\mathbf{X_i}$. The key insight is that, under these assumptions, we can rewrite the entire likelihood of the observed crosswise data as a product of three conditional probabilities. We can then marginalize the product over the two latent variables Z_i and T_i by summing up the conditional probabilities that we could in principle obtain for all possible combinations of the latent variables.[2] Our simulation studies confirm that this regression model can recover both the main ($\mathbf{\Theta}$) and auxiliary ($\boldsymbol{\beta}, \boldsymbol{\theta}$) parameters (Section C.4). We plan to demonstrate the effectiveness of the proposed regression frameworks with empirical examples in future research.

To validate the proposed framework, we simulate crosswise data with two covariates as in Online Appendix C.3. We then simulate the response variable according to the following generative model:

$$V_i = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \delta Z_i + \epsilon_i,$$

[2]For example, the third component inside the bracket represents $g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 1, 0)\mathbb{P}(Y_i = 1|Z_i = 1, T_i = 0)\mathbb{P}(Z_i = 1|\mathbf{X_i})\mathbb{P}(A_i = 1|Z_i = 1, T_i = 0)\mathbb{P}(T_i = 1|\mathbf{X_i})$. Here, $\mathbb{P}(Y_i = 1|Z_i = 1, T_i = 0)$ is the conditional probability that respondents choose the crosswise item when they actually have a sensitive trail *and* do not provide attentive responses. Because they do not follow the instruction, Assumption 1 states that this probability is $\frac{1}{2}$ (regardless of Z_i). Next, $\mathbb{P}(Z_i = 1|\mathbf{X_i})$ is the conditional probability that respondents have a sensitive trait, and we defined this quantity as $\pi_{\boldsymbol{\beta}}(\mathbf{X_i})$. Now, $\mathbb{P}(A_i = 1|Z_i = 1, T_i = 0)$ is the conditional probability that respondents choose the crosswise item in the anchor question when they actually have a sensitive trait *and* do not provide attentive responses. Because they do not follow the instruction, Assumption 1 states that this probability is $\frac{1}{2}$ (regardless of Z_i). Finally, $\mathbb{P}(T_i = 1|\mathbf{X_i})$ is the conditional probability that respondents *do not* provide attentive responses, and we defined this quantity as $1 - \gamma_{\boldsymbol{\theta}}(\mathbf{X_i})$. Hence, the joint probability for this component is $g_{\mathbf{\Theta}}(V_i|\mathbf{X_i}, 1, 0)\frac{1}{2}\pi_{\boldsymbol{\beta}}(\mathbf{X_i})\frac{1}{2}(1-\gamma_{\boldsymbol{\theta}}(\mathbf{X_i}))$.

where we set $\gamma_0 = 0$, $\gamma_1 = 0.3$, $\gamma_2 = 0.01$, $\delta = 1$, and $\epsilon_i \sim N(0, 1)$. Recall that Z_i is a latent variable for having a sensitive trait and we cannot observe its value directly (and thus crosswise data do not contain Z_i).

We then estimate the above crosswise regression model with the simulated observed outcome and crosswise data. Figure C.5 shows the estimates for our quantity of interest with different sample size. It demonstrates that the proposed regression model and estimation strategy can recover the latent magnitude of the association between the latent sensitive trait and the response variable ($\delta = 1$). It also shows that other ten parameters can be properly estimated by the proposed regression model. It is also straightforward to compute predicted values of the outcome variable with 95% confidence intervals using the parametric bootstrap. Figure C.6 displays the predicted values of the outcome in this particular simulation with different values for Z.

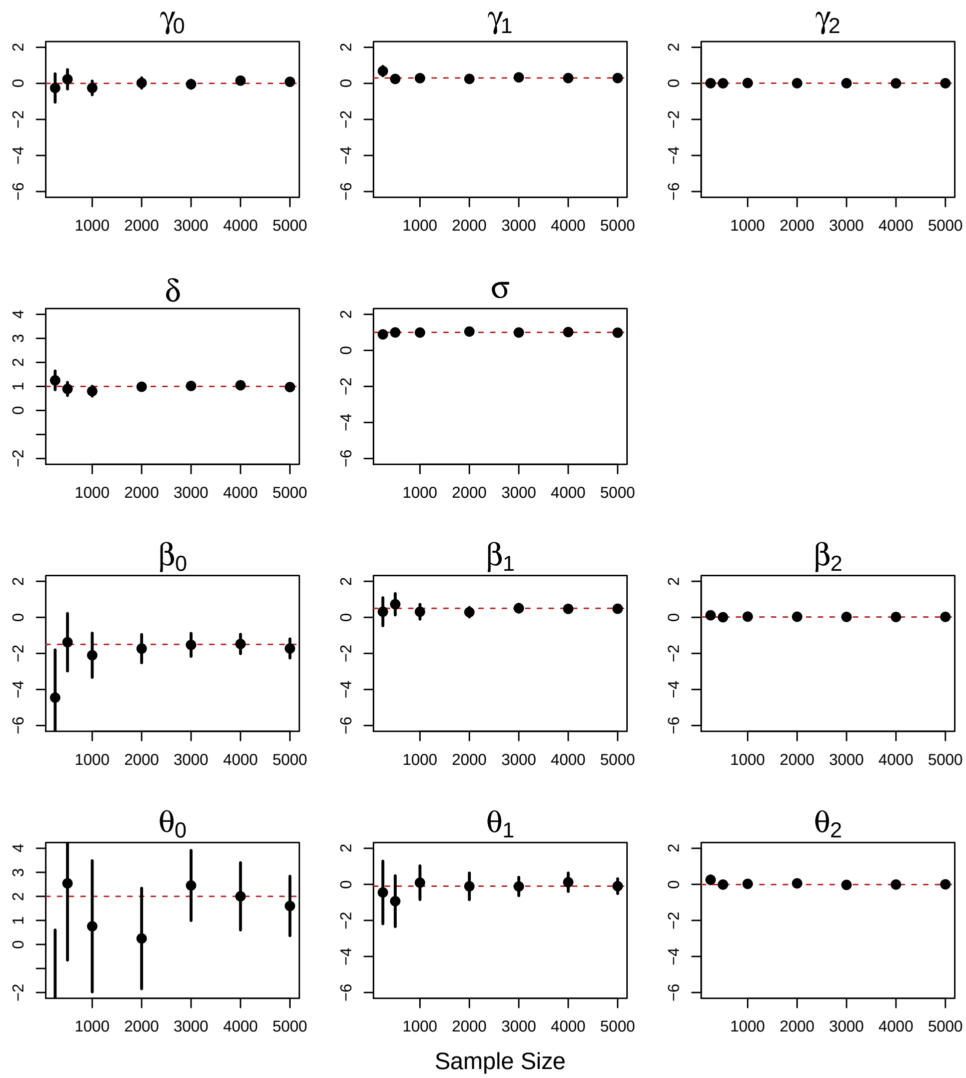

Figure C.5: **Finite Sample Performance of Regression Estimator (Sensitive Trait as a Predictor)**. *Note*: The dashed lines indicate the true values for the parameters.

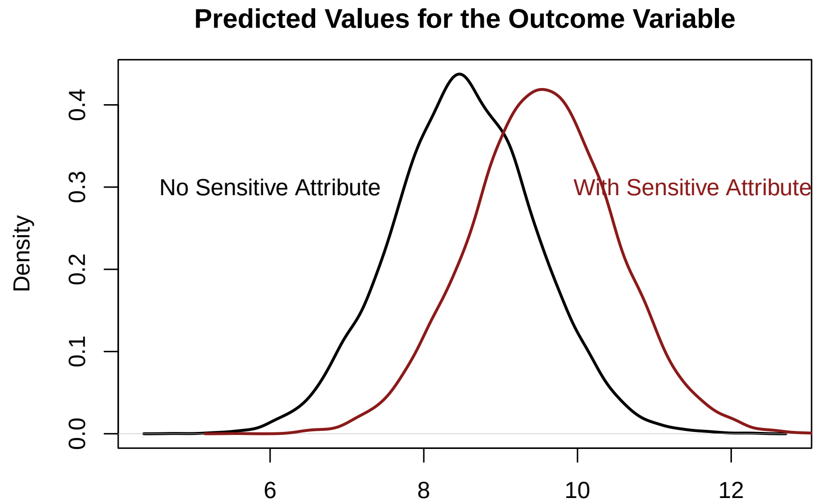

Figure C.6: **Simulated Outcome Values with and without the Sensitive Attribute.** *Note*: This graph visualizes the density of simulated (predicted) values for the outcome variable in the absence (left density) and presence (right density) of the sensitive attribute.

C.5 Sample Size Determination and Parameter Selection

When using the crosswise model with our procedure, researchers may wish to choose the sample size and specify other design parameters (i.e., π', p, and p') so that they can obtain (1) high statistical power for hypothesis testing and/or (2) narrow confidence intervals for precise estimation. To fulfill these needs, we develop power analysis and data simulation tools appropriate for our bias-corrected estimator.

First, our power analysis uses a one-sided hypothesis test based on the Wald test (Ulrich et al., 2012). We consider the null hypothesis $H_0 : \pi \leq \pi_0$, where π_0 (prevalence rate under the null) may be zero or a particular value obtained from direct questioning. Assuming that the crosswise estimate is larger than the direct questioning estimate — and larger than zero (i.e., more-is-better assumption), we then consider the alternative hypothesis $H_1 : \pi > \pi_0$ when the true value of π is π_1. Based on the normal approximation, the power function becomes:

$$\underbrace{\mathbb{P}(\{\text{Reject } H_0 | H_1 \text{ is true}\})}_{\text{Power}} = \beta = 1 - \underbrace{\Phi\Big(\frac{\pi_0 - \pi_1 - c\tilde{\sigma}_0}{\tilde{\sigma}_1}\Big)}_{\mathbb{P}(\text{Type II Error})}, \tag{C.3}$$

where Φ is the cumulative distribution function of the standard normal distribution, $c = \Phi(1-\alpha)$ is the critical value given size $\alpha = \mathbb{P}(\text{Type I Error})$, and $\tilde{\sigma}_0$ and $\tilde{\sigma}_1$ are *simulated* standard errors of the bias-corrected estimates under H_0 and H_1, respectively.[3]

Panel A of Figure C.7 plots power curves against the sample sizes. For this illustration, we assume that $\pi_0 = 0$, $\pi_1 = 0.1$, $\pi' = 0$, $p' = 0.1$, $\gamma = 0.8$, and $\alpha = 0.05$. Substantively, this means that we hope to distinguish the estimated prevalence of 0.1 from zero with a 95% (100% $\times$ $1-\alpha$) level of confidence when we expect that only 80% of respondents are attentive. Panel A displays multiple power curves for different

[3]We simulate the standard errors in repeated Monte Carlo experiments at $n = \{1, 500, 1000, 1500, 2000, 2500\}$. Our software also allows one to draw more fine-grained power curves.

values of p. It suggests that if researchers want to reject H_0 with power $\beta = 0.8$ they need to have about 400 (when $p = 0.1$), 600 (when $p = 0.2$), and 1400 (when $p = 0.3$) respondents.

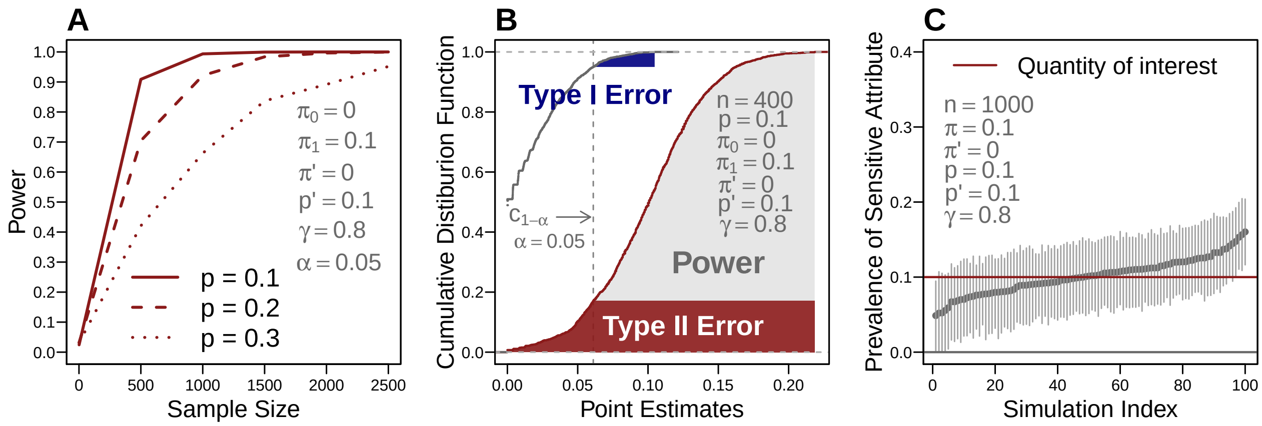

Figure C.7: **Tools for Power Analysis and Parameter Selection.** *Note*: Panel A shows power curves for three different values of p. Panel B offers visualizes type I error, type II error, and power as areas under the empirical cumulative distribution of point estimates under H_0 (left) and H_1 (right). Panel C display 100 point estimates and 95% confidence intervals based on 100 simulated data sets.

After choosing the sample size n, researchers may wish to verify if the selected n and design parameters achieve the desired level of power and re-adjust their design parameters if necessary. To further assist applied researchers, Panel B visualizes type I error, type II error, and statistical power as areas under the (empirical) cumulative distribution function of the sampling distribution of the bias-corrected estimates based on H_0 (left) and H_1 (right). We assume the same parameter values as above and $n = 400$ and $p = 0.1$. Panel B shows a small type I error and high power (close to 0.8) or equivalently small type II error as we expected.

Next, our data simulation tool allows researchers to see what they would obtain from using the crosswise model with some fixed sample size and parameter values. The key idea is to give analysts a sense of how large their 95% confidence interval would be and let them change the sample size and design parameters to achieve their desired precision (i.e., desired length of the confidence interval). To illustrate, Panel C shows the point estimates and 95% confidence intervals based on 100 simulated data sets with $n = 1000$ and $\pi = 0.1$, $\pi' = 0$, $p = p' = 0.1$, $\gamma = 0.8$. It suggests that the resulting interval estimate would be approximately the point estimate ± 0.05. If a narrower confidence interval is needed, researchers should increase the sample size and/or choose lower values for p and p'.

C.6 Secret Number Approach to Non-Sensitive Statements

This section introduces another extension of our method, which was not fully discussed in the main manuscript. One of the obstacles for some researchers to use the crosswise model is that they need to find an appropriate non-sensitive statement with a known prevalence in the crosswise question. Further, to maintain respondent privacy it is essential that a *different* non-sensitive item be used with each crosswise question in the survey. This can be quite difficult to do in practice when one asks many crosswise questions. To remedy this problem, we propose a *secret number approach* to the non-sensitive statement. The essence of the secret number approach is to use a different virtual die roll for the non-sensitive crosswise item in each question, but to do so in a way that makes it clear to respondents the critical information remains private even if they believe surveyors are recording the result of the virtual die rolls.

If respondents believe that a virtual die roll is private, there are many advantages of using it in online surveys for the non-sensitive piece of private information needed in the crosswise question (e.g., after they roll the die the statement used is something like the value of my die roll was 4 rather than a statement like my mother was born in January.) These advantages are the same ones as using a physical die in a paper and pencil survey and include:

• The necessary probabilities are known completely and do not change with the population being surveyed.

• The die roll can be used in the same format in many different questions, alleviating the need to develop a large library of items with known prevalence for surveys with many crosswise items.

• There is no chance that respondents will not know the value of the die roll or be uncertain about it (unlike, for example, their mother's birthday).

• The die roll is independent of sensitive item by construction.

The problem with the virtual die roll is that respondents will often not believe that it is not being recorded and so their privacy protection will be undermined.

However, a simple addition to the instructions for rolling a virtual die alleviates this problem completely. To reassure respondents that such virtual dice rolls cannot be recorded and used by researchers to undermine privacy, one can simply ask the respondent to first pick (and perhaps write down) a "secret number", which must be an integer between 1 and 6, and then roll the virtual die. Then, the statement they evaluate is simply "The value of the dice roll for this question was equal to my secret number."

Further, in surveys with multiple crosswise items, respondents can be told that the can change their secret number on each question or not as they like. Since researchers will not know whether they have or not, their privacy is assured.

D Examples and Discussion of Constructing Anchor Questions

As discussed in the main text, anchor questions should be topically connected to the other questions that are being asked in the crosswise format and should be sufficiently sensitive that they do not stand out relative to the other crosswise items.

Our purpose here is to first provide some examples of anchor questions that could be used for typical kinds of political science surveys that might use crosswise questions and then some discussion and practical advice about how to choose/construct good anchor questions.

D.1 Practice Guide

We first offer some practical advice for researchers when designing a bias-corrected crosswise model. First, to satisfy the random pick assumption (Assumption 1), researchers should randomize the order of the two crosswise answer-choices both in the crosswise and anchor questions. This can be done at the question level or (if there are many crosswise questions presented sequentially) at the respondent level (to avoid respondent confusion).

Second, attention consistency (Assumption 2) is satisfied when the crosswise and anchor questions have the same *proportion* of attentive respondents. If respondents, on average, perceive these two types of questions to be somehow different, attention consistency may be violated. Thus, we recommend that researchers design the anchor questions to fit in topically with the other crosswise questions in the survey, have a similar level of sensitivity, and look quite similar to them (e.g., the anchor question should use the same format for the non-sensitive item as the other crosswise questions (e.g., a mother's birthday) and the sensitive item should be about the same length in both the anchor and other crosswise questions).

We discuss more of the subtleties of writing good anchor questions (and provide examples) below. In addition, researchers may also compare the duration of time spent on the crosswise and anchor questions as a way of diagnosing any differences in how respondents have treated them.

Finally, randomizing the position of the anchor question in the survey relative to the crosswise questions may help guarantee that there is no carryover effect from one type of question to another (Assumption 3).

D.2 Examples

To help researchers get started on writing useful anchor questions, we provide a set of examples for anchor questions that might fit into crosswise survey modules on sensitive political topics (see Table D.1). Of course, these examples are meant only as a starting point and may or may not make sense in any given setting. As always researchers should think carefully (and creatively) about the specific survey and sample in which they might be used.

Anchor statement	May be useful for surveys with crosswise items about	True prevalence	Note
I have paid a bribe to be at the top of a waiting list for organ transplants	corruption	0	
In 2019, I became a naturalized citizen of the United States	political participation, immigration attitudes	.0025 in a US sample	
In the last election, I voted in a state other than the one I in which I registered	elections	0	
In 2019 or 2020, I donated at least $200 to NARAL, a Political Action Committee that advocates for abortion rights	elections, participation	598 people (US sample)	
At some time in my life, I have worked for a political organization that advocates the violent overthrow of the government	radicalism, trust in government or political institutions	$\approx$ 0 for western samples	*
Last year, the gross annual income of my household was more than 600,000 dollars	inequality, policy opinion	$\approx$ .005 in US sample	**
At sometime in my life, I have physically threatened and member of Congress	public opinion surveys	≈ 0	***
I am a registered lobbyist for the Iron and Steel Association of America	policy opinion surveys	0 (this industry group does not exist)	****
I am going to vote for the [name of a fictional party]	election, voting for extremist parties, "shy voters"	0	

Table D.1: **Examples of Anchor Statements.** *Note*: * = Assuming this is not the object of inquiry, this would be a good match to other cross-wise questions that are very intrusive/sensitive, ** = Given the difficult of sampling high earners, this is likely 0 in any real sample. Also, much evidence suggests people think of income questions as sensitive, *** = Keep in mind that the public opinion survey we have in mind would be probing sensitive topics about the relationship of citizens and leaders so these kinds of question may seem jarring outside of that context but within that context they, **** = The industry can be chosen to fit the topic of the survey.

D.3 Discussion

The examples above illustrate some of the different kinds of sensitive statements that may have known prevalence in a given population. As the diversity of the examples illustrate, however, statements differ in the extent to which they are sensitive and so they should be matched in terms of the level of sensitivity to the other crosswise questions in the survey.

In addition, some of the examples suggest that researchers should try to avoid statements that, while they have might have a true prevalence of zero, might nevertheless attract false positives for reasons other than inattention. One way this can happen is if an item attracts support for purely symbolic or affecting reasons. For example, in the example that uses a fictional party, care must be taken in choosing a neutral name of the fictional party. Prior et al. (2015), for example, have shown that partisans sometimes report factual opinions on surveys that they know to be wrong as a kind of partisan cheerleading. Applying this to an election survey aimed at detecting "shy" Trump voters, for example, one might use an anchor question like "In the 2020 election I voted for the Anarchy Party of Texas." Since there is no such party, the prevalence rate should be zero. However, we might expect this statement to attract some support from anarchists who engaged in a kind of partisan cheerleading.

A better alternative would be a party with a more neutral name. This would be less sensitive, but in the context of a survey asking crosswise questions which other (extremist and non-extremist) parties one would vote for, it would likely fit in well. This illustrates again that whether an anchor question matches the set of crosswise questions well depends entirely on the context of the other crosswise questions.

Finally, its important to point out that a given anchor question can be made more typical of the whole set of crosswise questions not only by altering the anchor question itself but also by altering the composition of the whole set of crosswise questions in ways that make a given anchor question more representative of the whole set.

Specifically, those who are finding it difficult to construct a topically relevant anchor question that is sufficiently sensitive to match their crosswise questions of interest can simply add several topically relevant crosswise questions that are less sensitive than their crosswise questions of interest. In this way, the distribution of all the crosswise questions that the respondent sees (before encountering the anchor question) can be made to be more diverse in terms of sensitivity. This may make the anchor question less jarring when it is encountered. Likewise, if the trouble is finding a topically relevant anchor question, then several off-topic crosswise questions can be added to alter the mix of crosswise questions in ways that make the anchor question less of a stand-out.

References for Online Appendix

Banayejeddi, Mortaza, Sima Masudi, Sakineh Nouri Saeidlou, Fatemeh Rezaigoyjeloo, Fariba Babaie, Zahra Abdollahi and Fatemeh Safaralizadeh. 2019. "Implementation evaluation of an iron supplementation programme in high-school students: the crosswise model." *Public health nutrition* 22(14):2635–2642.

Canan, Chelsea E, Geetanjali Chander, Richard Moore, G Caleb Alexander and Bryan Lau. 2021. "Estimating the prevalence of and characteristics associated with prescription opioid diversion among a clinic population living with HIV: Indirect and direct questioning techniques." *Drug and alcohol dependence* 219:108398.

Coutts, Elisabethen, Ben Jann, Ivar Krumpal and Anatol-Fiete Näher. 2011. "Plagiarism in student papers: prevalence estimates using special techniques for sensitive questions." *Jahrbücher für Nationalökonomie und Statistik* 231(5-6):749–760.

Downes, Marnie, Lyle C Gurrin, Dallas R English, Jane Pirkis, Dianne Currier, Matthew J Spittal and John B Carlin. 2018. "Multilevel Regression and Poststratification: A Modeling Approach to Estimating Population Quantities From Highly Selected Survey Samples." *American journal of epidemiology* 187(8):1780–1790.

Enzmann, Dirk. 2017. Die Anwendbarkeit des Crosswise-Modells zur Prfung kultureller Unter schiede sozial erwnschten Antwortverhaltens. In *Methodische Probleme von Mixed-Mode-Anstzen in der Umfrageforschung*, ed. Stefanie Eifler and Frank Faulbaum. Springer VS chapter 10, pp. 239–277.

Franco, Annie, Neil Malhotra, Gabor Simonovits and LJ Zigerell. 2017. "Developing standards for post-hoc weighting in population-based survey experiments." *Journal of Experimental Political Science* 4(2):161–172.

Gingerich, Daniel W, Virginia Oliveros, Ana Corbacho and Mauricio Ruiz-Vega. 2016. "When to protect? Using the crosswise model to integrate protected and direct responses in surveys of sensitive behavior." *Political Analysis* pp. 132–156.

Hoffmann, Adrian, Birk Diedenhofen, Bruno Verschuere and Jochen Musch. 2015. "A strong validation of the crosswise model using experimentally-induced cheating behavior." *Experimental Psychology* .

Hoffmann, Adrian and Jochen Musch. 2019. "Prejudice against women leaders: insights from an indirect questioning approach." *Sex Roles* 80(11):681–692.

Hoffmann, Adrian, Julia Meisters and Jochen Musch. 2020. "On the validity of non-randomized response techniques: an experimental comparison of the crosswise model and the triangular model." *Behavior research methods* pp. 1–15.

Höglinger, Marc and Ben Jann. 2018. "More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model." *PloS one* 13(8):e0201770.

Hopp, Christian and Alexander Speil. 2019. "Estimating the extent of deceitful behaviour using crosswise elicitation models." *Applied Economics Letters* 26(5):396–400.

Jann, Ben, Julia Jerke and Ivar Krumpal. 2012. "Asking sensitive questions using the crosswise model: an experimental survey measuring plagiarism." *Public opinion quarterly* 76(1):32–49.

Klimas, C, Ulrike Ehlert, TJ Lacker, Patricia Waldvogel and Andreas Walther. 2019. "Higher testosterone levels are associated with unfaithful behavior in men." *Biological psychology* 146:107730.

Korndörfer, Martin, Ivar Krumpal and Stefan C Schmukle. 2014. "Measuring and explaining tax evasion: Improving self-reports using the crosswise model." *Journal of Economic Psychology* 45:18–32.

Kuhn, Patrick M and Nick Vivyan. 2018. "Reducing turnout misreporting in online surveys." *Public Opinion Quarterly* 82(2):300–321.

Kundt, Thorben. 2014. "Applying'Benford's Law'to the crosswise model: Findings from an online survey on tax evasion." *Available at SSRN 2487069* .

Kundt, Thorben, Florian Misch and Birger Nerré. 2013. "Re-assessing the merits of measuring tax evasions through surveys: Evidence from Serbian firms." *ZEW-Centre for European Economic Research Discussion Paper* (13-047).

Meisters, Julia, Adrian Hoffmann and Jochen Musch. 2020. "Can detailed instructions and comprehension checks increase the validity of crosswise model estimates?" *PloS one* 15(6):e0235403.

Mercer, Andrew, Arnold Lau and Courtney Kennedy. 2018. "For Weighting Online Opt-In Samples, What Matters Most." *Pew Research Center* .

Mieth, Laura, Maike M Mayer, Adrian Hoffmann, Axel Buchner and Raoul Bell. 2021. "Do they really wash their hands? Prevalence estimates for personal hygiene behaviour during the COVID-19 pandemic based on indirect questions." *BMC public health* 21(1):1–8.

Nasirian, Maryam, Samira Hosseini Hooshyar, Arezoo Saeidifar, Leila Taravatmanesh, Aboubakr Jafarnezhad, Sina Kianersi and Ali Akbar Haghdoost. 2018. "Does crosswise method cause overestimation? An example to estimate the frequency of symptoms associated with sexually transmitted infections in general population: A cross sectional study." *Health Scope* 7(3).

ÖZGÜL, Nilgün. 2020. "A Survey on Illicit Drug Use among University Students by Binary Randomized Response Technique: Crosswise Design." *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 24(2):377–388.

Prior, Markus, Gaurav Sood, Kabir Khanna et al. 2015. "You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions." *Quarterly Journal of Political Science* 10(4):489–518.

Schnapp, Patrick. 2019. "Sensitive Question Techniques and Careless Responding: Adjusting the Crosswise Model for Random Answers." *methods, data, analyses* 13(2):13.

Shamsipour, Mansour, Masoud Yunesian, Akbar Fotouhi, Ben Jann, Afarin Rahimi-Movaghar, Fariba Asghari and Ali Asghar Akhlaghi. 2014. "Estimating the prevalence of illicit drug use among students using the crosswise model." *Substance Use & Misuse* 49(10):1303–1310.

Ulrich, Rolf, Hannes Schröter, Heiko Striegel and Perikles Simon. 2012. "Asking sensitive questions: a statistical power analysis of randomized response models." *Psychological methods* 17(4):623.

Waubert de Puiseau, Berenike, Adrian Hoffmann and Jochen Musch. 2017. "How indirect questioning techniques may promote democracy: A preelection polling experiment." *Basic And Applied Social Psychology* 39(4):209–217.