



ORIGINAL ARTICLE

# Improving studies of sensitive topics using prior evidence: an informative Bayesian approach for list experiments

Xiao Lu<sup>1</sup>  and Richard Trautmüller<sup>2</sup> 

<sup>1</sup>School of International Studies, Peking University, Beijing, China and <sup>2</sup>School of Social Sciences, University of Mannheim, Mannheim, Germany

**Corresponding author:** Xiao Lu; Email: [xiao.lu.research@gmail.com](mailto:xiao.lu.research@gmail.com)

(Received 27 October 2024; revised 19 May 2025; accepted 18 September 2025)

## Abstract

Estimates of sensitive questions from list experiments are often much less precise than desired. We address this well-known inefficiency problem by presenting an informative Bayesian approach that combines indirect measures with prior information. Specifying informed priors amounts to a principled combination of information that increases the efficiency of model estimates. This framework generalizes a range of different modeling approaches for list experiments, such as the inclusion of direct items, auxiliary information, the double list experiment, and the combination of list experiments with other indirect questioning techniques. As we demonstrate in real-world examples from political science, the informative Bayesian approach not only improves the utility but also changes the substantive implications drawn from list experiments.

**Keywords:** Bayesian analysis; informative priors; list experiments; sensitive questions

## 1. Introduction

Sensitive topics abound in social sciences where subjects are often unwilling to reveal their true preferences or behaviors (Blair *et al.*, 2020). Recent innovations in statistical methodology have made tremendous advances in reducing sensitivity bias with the design and analysis of indirect questioning techniques (see, e.g., Conley and McCabe, 2011; Coutts and Jann, 2011; Rosenfeld *et al.*, 2016; Yamaguchi, 2016). List experiments, in particular, have become a highly popular tool for social science research. While these experimental designs help reduce bias, they are not without cost. It is a well-known problem that estimates from list experiments lack statistical efficiency and are therefore, often less precise than desired or useful in samples of conventional size (Glynn, 2013; Blair *et al.*, 2020).<sup>1</sup>

In this paper, we address the inefficiency problem of list experiments head-on and present an informative Bayesian approach that combines indirect measures with prior evidence. One great benefit of the Bayesian approach is the ability to include additional information in a straightforward manner via prior distributions (Western and Jackman, 1994; Gill and Walker, 2005; Jackman, 2009; Gill, 2015). Specifying informed priors amounts to a principled combination of information, which not

<sup>1</sup> Another issue with list experiments that has received some attention recently is the problem of nonstrategic misreporting (e.g., Ahlquist, 2018; Blair *et al.*, 2019; Kuhn and Vivyan, 2021).

only increases the efficiency<sup>2</sup> of model estimates, but leads to an improved theoretical understanding of sensitive preferences and behaviors of political relevance.

Our contribution is thus fourfold. First, we improve the study of sensitive topics in social sciences by advancing the usability of list experiments. The inefficiency problem has limited a wider use of list experiments because it demands very large sample sizes that come with increased costs (Blair *et al.*, 2020). As a result, many scholars turn back to direct survey questions, trading efficiency gains for bias. We demonstrate that the inefficiency of list experiments can be greatly reduced with readily available prior information.

Second, we connect the literature on list experiments to recent advances in survey methodology that combine probability and non-probability samples (Sakshaug *et al.*, 2019; Wiśniowski *et al.*, 2020; Jerit and Barabas, 2023). This work faces a similar challenge of combining reliable but expensive information (probability samples) and less reliable but cheaper information (non-probability samples) with the ultimate goal to increase the accuracy of inferences. Specifically, we introduce the idea of “conjugate-distance” priors (Wiśniowski *et al.*, 2020) to the analysis of list experiments.

Third, we unify several existing approaches to designing and analyzing list experiments under a single common framework. This is particularly useful, because mastering the many techniques scattered throughout the literature (e.g., Droitcour *et al.*, 1991; Blair *et al.*, 2014; Aronow *et al.*, 2015; Chou *et al.*, 2020) places a high demand on applied researchers to learn multiple disparate approaches that were developed independently along with the accompanying software tools. By proposing a single Bayesian framework for list experiments, we generalize these existing approaches and only require users to do one thing: to specify prior distributions—regardless of the type of information or model class they would like to use. Our approach even allows for any *combination* of previous techniques by sequential Bayesian updating. An accompanying software package, *bayeslist*, which implements the proposed methods of this paper for easy use by applied researchers, is available in the Comprehensive R Archive Network.

Fourth, while social scientists have embraced Bayesian simulation methods as estimation tools for complicated models for quite some time (e.g., Jackman, 2000), they have been much more reluctant to tap into the full potential of the Bayesian approach by leveraging informed prior distributions. Instead, these “Bayesians of convenience” (Gill, 2015) have defaulted to diffuse and often uniform prior forms. Early exceptions are Western and Jackman (1994) who use historical data and Gill and Walker (2005) who elicit expert knowledge to inform regression models. Similar in spirit, Glynn and Ichino (2015), Humphreys and Jacobs (2015), and Pérez Bentancur and Tiscornia (2023) show how qualitative and quantitative evidence can be integrated to improve causal inference. We add to this literature by proposing Bayesian information integration for sensitive questions. Specifically, we introduce new data-derived sources for informative priors, show that they are often readily available and much easier to specify, and demonstrate how to mitigate concerns with biased priors. This latter point is especially valuable to dispel widespread concerns of “prior skeptics.” Thus, our contribution also encourages the wider use of informative priors in social science research.

## 2. Informative priors and list experiments

The utility of list experiments crucially depends on how much information is available for the analysis. Without a sufficiently large sample size, a biased direct question may be as good as or even better than a less biased but also less accurate list experiment (Blair *et al.*, 2020; Li and Van den Noortgate, 2022). The efficiency problem of list experiments has attracted great attention, but existing solutions

<sup>2</sup>Strictly speaking, efficiency is of course a “frequentist” concept referring to the uncertainty of an estimator under repeated sampling, which is at odds with the Bayesian notion of probability. However, this distinction seems of little *practical* relevance. The underlying concern is with the inferential uncertainty attached to an estimate which in Bayesian analysis is expressed in the variance of the posterior probability.

are scattered in the literature and often developed in parallel. For example, different variance reduction strategies have been proposed, including: double list experiments (Miller, 1984; Droitcour *et al.*, 1991), negatively correlated control items (Miller, 1984; Glynn, 2013), combining list and endorsement experiment designs (Blair *et al.*, 2014), combining list and direct question designs (Aronow *et al.*, 2015), and using auxiliary information (Chou *et al.*, 2020).

Thus, while some approaches focus on refinements of list experiments at the design stage, others attempt improvements at the analysis stage when data have already been collected. The former approach is subject to ex-post evaluation and can be very costly when the results are unsatisfying, and experiments have to be re-run. The latter, while arguably much cheaper, requires the mastering of multiple techniques that were developed separately and come with different software packages. All of this can be quite demanding for applied researchers.

Against this background, it would be desirable to have a single framework for the analysis of list experiments that is simple, accurate, and cost-efficient, and that can be done *post hoc* by combining already existing information rather than collecting more data. We propose such a framework based on the idea of Bayesian information updating that is able to incorporate various types of prior information for different models into the analysis of sensitive questions. Thereby, it simplifies the task of improving list experiment analysis by finding available external information and specifying priors.

## 2.1. Notation

### 2.1.1. List experiments

In a list experiment with  $N$  respondents, individual  $i$  is randomly assigned to the control group ( $T_i = 0$ ) and receives  $J$  control items or to the treatment group ( $T_i = 1$ ) and receives  $J + 1$  items by adding the sensitive item. The latent  $Z_i^*$  indicates whether the respondent provides an affirmative answer to the sensitive item, and the latent variable  $Y_i^*$  gives the number of affirmative answers to the control items. Respondents in the treated condition only reveal the aggregated outcome denoted by  $Y_i$ , which is the sum of the two latent variables.

### 2.1.2 Simple analysis of list experiments

The prevalence of the sensitive item,  $\pi$ , can be obtained by a simple non-parametric differences-in-means estimator:

$$\hat{\pi} = \frac{1}{N_1} \sum_1^N T_i Y_i - \frac{1}{N_0} \sum_1^N (1 - T_i) Y_i, \tag{1}$$

where  $N_1$  and  $N_0$  are the number of respondents in the treatment and control group, respectively. The variance of the estimated prevalence  $\hat{\pi}$  from this difference-in-means estimator (under a balanced list experiment design) is given as:

$$\sigma_{\text{list}}^2 = \frac{1}{N-1} \left\{ \underbrace{\pi(1-\pi)}_{\text{Variance of sensitive item}} + \underbrace{4\text{Var}(Y_i^*) + 4\text{Cov}(Y_i^*, Z_i^*)}_{\text{Variance of and covariance with control items}} \right\}, \tag{2}$$

which consists of a variance component of the true prevalence  $\pi$ , and variance and covariance components of the control items (Blair *et al.*, 2020). Thus, the variance of the list experiment estimator is almost always larger than that of the direct question estimator, the extent to which depends on the variance of the control items and the correlation between sensitive and control items.

### 2.1.3 Regression models for list experiments

Recent innovations greatly improve the utility of list experiments by modeling answers in the context of regression models (Corstange, 2009; Imai, 2011; Blair and Imai, 2012; Imai *et al.*, 2015; Eady, 2017).

While we obviously cannot observe individual answers to the sensitive item, the list experiment allows us to identify the *joint distribution* of the latent sensitive answer and the control items,  $P(Y_i^*, Z_i^*)$ , if certain assumptions are met (randomization, no design effect, no liars) (Glynn, 2013). This possibility is then exploited in a multivariate modeling strategy that simultaneously models the response to the sensitive item and the control items.

Our focus here is the canonical *sensitive item outcome model*, where the answer to the sensitive item is regressed on and potentially explained by covariates (Imai, 2011; Blair and Imai, 2012). But our Bayesian approach also applies to alternative models, such as the *sensitive item predictor model*, where the answer to the sensitive item is entered as explanatory variable for some outcome of interest (Imai et al., 2015) or the *sensitive item misreporting model* which models misreporting of the sensitive item (Eady, 2017).

#### 2.1.4 Submodel for sensitive item

The core of all regression models for list experiments is the probability of affirming the sensitive item, i.e., its prevalence  $\pi$ :

$$\hat{\pi}_i = P(Z_i^* = 1 | \mathbf{X}_i) = g(\mathbf{X}_i; \boldsymbol{\delta}). \quad (3)$$

In this equation,  $\mathbf{X}_i$  is a vector of covariates for individual respondent  $i$  and  $\boldsymbol{\delta}$  is a vector of coefficients associated with the covariates. The probability is usually modeled using a logistic transformation on the linear combination of  $\mathbf{X}_i$  and  $\boldsymbol{\delta}$ :

$$g(\mathbf{X}_i; \boldsymbol{\delta}) = \text{logit}^{-1}(\mathbf{X}_i' \boldsymbol{\delta}). \quad (4)$$

#### 2.1.5 Submodel for control items

The probability of the number of affirmative answers to the control items is modeled in a similar way except that it uses a beta-binomial with  $J$  trials (i.e., the number of control items) and relates the outcome to a set of coefficients  $\boldsymbol{\psi}$ :

$$P(Y_i^* = y | \mathbf{X}_i, Z_i^*, J) = h(y | \mathbf{X}_i, Z_i^*, J; \boldsymbol{\psi}), \quad (5)$$

where  $h(\cdot)$  is the probability mass function of a beta-binomial distribution (Blair and Imai, 2012). Due to the mixture structure of the likelihood, the coefficients,  $\boldsymbol{\psi}$ , also depend on the values of  $Z_i^*$ , although a common assumption is to simply set  $\boldsymbol{\psi}_0 = \boldsymbol{\psi}_1$  for the treatment and control groups.

#### 2.1.6 Likelihood

Depending on the treatment condition  $T_i$ , observed outcome  $Y_i$ , and latent variables  $Z_i^*$  and  $Y_i^*$ , the observed-data likelihood functions of models for list experiments are combinations of the above mentioned probabilities, yielding a complicated mixture structure. More technical details and likelihood functions for these models can be found in the Supplementary Material A. We also refer readers to Imai (2011), Imai et al. (2015) and Eady (2017). For Bayesian versions of these models of list experiments, we now require prior distributions for their unknown parameters of interest, e.g.,  $P(Z^*, \boldsymbol{\delta}, \boldsymbol{\psi})$ . This is what we discuss next.

### 3. Combining list experiments with prior evidence: an informative Bayesian approach

The statistical analysis of list experiments suffers from what Western and Jackman (1994) have called the “weak data problem,” where a lack of sufficient information leads to high variances. The solution is therefore to bring in more information via prior distributions. Prior distributions are simply formal statements of the existing information a researcher brings to the analysis. That priors may impact statistical results is a long-standing concern with Bayesian analysis, which unfortunately has kept social scientists from exploiting the benefits of systematically including existing information into their models (see Gill and Walker (2005) and Western and Jackman (1994) for exceptions). We argue

instead that researchers should embrace informative priors to strengthen their inferences on sensitive behaviors.

Although extant research has proposed Bayesian methods for list experiments (Blair and Imai, 2012; Rosenfeld *et al.*, 2016; Blair *et al.*, 2019), it only relies on Bayesian simulation algorithms, but does not incorporate informative priors. Studies on sensitive questions that have incorporated prior information, are either not applicable to the analysis of list experiments (e.g., for endorsement experiments, see Groenitz (2015) and Chou *et al.* (2020)), or constrained to very specific types of prior data or prior distributions (Blair *et al.*, 2014; Liu *et al.*, 2019), which limits their overall applicability. Groenitz (2014) proposed to incorporate priors for the analysis of list experiments without a control group. Blair *et al.* (2014) propose a specific Bayesian procedure to combine list and endorsement experiments. Aronow *et al.* (2015) introduce a nonparametric estimator that combines list experiments with direct questions, but their non-Bayesian approach is restricted to prevalence estimates and does not extend to multivariate analysis or other sources of prior information.

Our approach is fully Bayesian and presents a generalized framework for incorporating prior information in regression models of list experiments. Our approach is not only readily adaptable to a) many *different types of prior information* (e.g., direct items (Aronow *et al.*, 2015) or auxiliary information (Chou *et al.*, 2020) from historic sources (Western and Jackman, 1994) or experts (Gill and Walker, 2005), etc.) but also b) to *different model types* like the sensitive item outcome (Blair and Imai, 2012), the sensitive item predictor (Imai *et al.*, 2015) and the misreporting model (Eady, 2017). It also easily handles c) more *specialized list experiment designs* such as the double experiments (Droitcour *et al.*, 1991) and even allows for d) the *combination of list experiments with other types of experiments* (Gerber and Green, 2012; Blair *et al.*, 2014). More generally, our contribution is to demonstrate the benefits of a fully Bayesian approach and to encourage the wider use of informative priors in social science research.

In what follows, we first show why informative priors work and where they come from. The next section then illustrates in great detail how Bayesian analysis of list experiments is done in practice and how it changes our understanding of sensitive topics.

### 3.1. The case for informative priors

Bayesian inference offers a disciplined routine to incorporate prior information into the model estimation process via Bayes' rule. Remember that this rule uses probability statements on how to update *prior* beliefs about unknown quantities and model parameters with the *likelihood* of the data to form *posterior* beliefs (Jackman, 2009; Gill, 2015). The distinction between uninformative or vague priors and informative priors refers to the choice of the parameters and the shape of their respective probability distributions. In the absence of prior information, one could set the mean of the distribution to zero and choose a large variance, resulting in a flat distribution. Model estimates are then almost fully determined by the likelihood alone. But this misses a valuable opportunity to incorporate additional information by specifying the prior mean and variance in more precise terms, which results in a more narrow prior distribution.

To provide some intuition for why the incorporation of prior information increases the efficiency of list experiments, we return to the simple difference-in-means estimator (Equation 1). By relying on the central limit theorem, we may treat the estimates as normally distributed with mean equal to the true prevalence  $\pi$  and variance of the sampling distribution  $\sigma_{\text{list}}^2$ , i.e.,  $\hat{\pi} \sim \mathcal{N}(\pi, \sigma_{\text{list}}^2)$  (cf. Gerber and Green, 2012). Additional outside information on the prevalence may be specified as a prior distribution that also takes a normal form, i.e.,  $\pi \sim \mathcal{N}(\pi_0, \sigma_{\text{prior}}^2)$ , where  $\pi_0$  is the prior prevalence suggested by the external information source and  $\sigma_{\text{prior}}^2$  expresses our uncertainty about this prevalence.

Based on a standard result in conjugate Bayesian analysis, the combination of normally distributed data and a normal prior yields a normally distributed posterior with the following *precision*,  $\tau_{\text{posterior}}$ .<sup>3</sup>

$$\tau_{\text{posterior}} = N \times \tau_{\text{list}} + \tau_{\text{prior}}. \quad (6)$$

The above equation summarizes the three main factors influencing estimation efficiency, namely: the sample size, the data precision, and the prior precision. First, everything else equal, the larger the sample size  $N$ , the higher the posterior precision. Second, the higher the precision of the list experiment denoted by  $\tau_{\text{list}}$ , the higher the posterior precision. The key ingredient is the third term, i.e., the prior precision  $\tau_{\text{prior}}$ . The more prior information we have, the higher the prior precision and therefore the overall posterior precision. Indeed, as long as the variance of the prior is not infinite, we will always have more efficiency by incorporating at least *some* prior evidence.<sup>4</sup>

In addition to efficiency gains, another advantage of incorporating prior information is to avoid “inadvertently informative” priors. Remember that a prime interest lies in making inferences about the parameters  $\delta$  in the submodel  $g(\mathbf{X}_i; \delta)$  for the sensitive item and that it often takes the logistic form. Unfortunately, vague normal priors in logistic regression models may have unintended consequences (cf. Seaman III *et al.*, 2012; Northrup and Gerber, 2018). Since the logit transformation is non-linear, large negative and large positive values yield transformed probabilities that are bimodal and bunch up against zero and one. The problem is more serious the larger the prior variance. Therefore, more “uninformative” priors on the logit scale may actually turn out to be highly informative on the probability scale (see Supplementary Material C for an illustration). This may be especially problematic for low-incidence sensitive items and weak data sets with few observations, where the prior is more influential. One solution is to reduce the prior variance, e.g., by specifying uninformative priors on the logit scale as  $\mathcal{N}(0, 1)$ , which translates roughly into a uniform prior on the probability scale (McElreath, 2018). Again, we suggest relying on informative priors with low variance instead. This way, the prior variance is not set in an ad-hoc fashion but guided by existing evidence.

We note that obtaining external information relevant to sensitive questions is often astoundingly easy. This is because it either already exists in direct numerical form or is an already existing element of the research design and therefore available “for free.” For instance, whilst potentially biased, prevalence rates of sensitive behaviors are usually available from governments, NGOs, or field experts, at least for some sub-populations. Many list experimental designs also include direct survey questions on the sensitive topic (Blair *et al.*, 2020). Last, because different researchers usually conduct list experiments with similar sensitive questions for different populations or time points, information from previous list experiments can be used to inform the analysis of the current list experiment (see Gerber and Green (2012) for the integration of experiments).

### 3.2. Bayesian information integration

According to Bayes’ theorem, the joint posterior distribution of the main unknown parameters of interest—i.e., the latent sensitive attitude or behavior  $Z^*$  and the submodel coefficients  $\delta$  and  $\psi$ —is obtained by conditioning on the observed outcome data  $Y$ , treatment status  $T$  and set of control variables  $\mathbf{X}$ , as well as prior distributions for all unknowns via the generic statement:

$$\underbrace{P(Z^*, \delta, \psi | \{\mathbf{X}, Y, T\})}_{\text{Posterior}} \propto \underbrace{P(\{\mathbf{X}, Y, T\} | Z^*, \delta, \psi)}_{\text{Likelihood}} \underbrace{P(Z^*, \delta, \psi)}_{\text{Prior}}. \quad (7)$$

<sup>3</sup>The precision is the inverse of the variance, i.e.,  $\tau = 1/\sigma^2$ . See Supplementary Material B for the derivation of this standard result.

<sup>4</sup>Note that while this statement is true in the normal-normal case, it does not generalize to cases of a beta prior and a binomial likelihood.

While the prevailing maximum likelihood estimation of list experiments focuses on *maximizing* the likelihood  $P(\{X, Y, T\}|Z^*, \delta, \psi)$ ,<sup>5</sup> Bayesian inference instead *multiplies* it with the joint prior distribution for all unknown quantities  $P(Z^*, \delta, \psi)$ . The actual model estimation involves marginalizing the resulting posterior  $P(Z^*, \delta, \psi|\{X, Y, T\})$  using MCMC or related statistical simulation algorithms. Since these are implemented in readily available general purpose software such as STAN or JAGS (Plummer, 2003; Carpenter *et al.*, 2017), we do not address them here.<sup>6</sup> Instead, our focus is specifically on the role of the prior distribution and its effects.

The prime interest when analyzing list experiments usually centers on the prevalence of the sensitive item  $Z^*$  and/or the coefficients in the sensitive outcome equation,  $\delta$ . Given the model structure, we may factor the joint prior distribution into its component marginal and conditional prior distributions:<sup>7</sup>

$$P(Z^*, \delta, \psi) = P(\psi)P(\delta|Z^*)P(Z^*) = P(\psi)P(Z^*|\delta)P(\delta). \quad (8)$$

This factorization shows that we can specify informative priors for  $Z^*$  or  $\delta$  independently of priors for  $\psi$ , which will usually be of the uninformative kind. Because prior information on one parameter propagates to another via the conditional, we are able to formulate a prior for either  $Z^*$  or  $\delta$  while simultaneously informing the other parameter. This is a convenient feature as it allows us to select an appropriate prior depending on what source of information is available. Sometimes, prior information on the latent prevalence  $Z^*$  will be available (see our real-world example in Supplementary Material F). In other instances, as in the real-world examples discussed in Section 4, prior information will directly refer to the coefficients  $\delta$  in the sensitive item outcome model.

### 3.3. Dispelling prior concerns

Many researchers are skeptical of using prior information because of the suspicion that doing so injects “subjectivity” into the analysis instead of “letting the data speak” by itself. This concern has at least three components: a) a general concern with “subjectivity” in data analysis which seems to violate “objective” scientific practice, b) a particular concern with the sources of prior information which may seem arbitrary, and c) a more specific concern with the deliberate use of information that is already *known* to be biased. We dispel these concerns as follows.

First, and as has been noted again and again, *all* decisions in statistical modeling are in a sense “subjective” and require justification. This includes, among many others, distributional assumptions about the data generation process, which variables to include in the model and with which functional form. That “researcher degrees of freedom” are vast and have direct implications for results is recognized in the recent discussion about the credibility crisis in science (Simmons *et al.*, 2011; Breznau *et al.*, 2022). Thus, prior distributions are not qualitatively different from the many other modeling assumptions researchers make.

Second, the source of information used in prior construction is far from subjective but is often itself based on data. In informative priors, there are two major sources: “expert opinion” and “historic data.” In our approach, we focus on the latter and suggest new sources for such already existing information, notably common elements of list experimental designs themselves, such as direct items. At its core,

<sup>5</sup>Again, exact likelihood functions of different sensitive item models are delegated to Supplementary Material A.

<sup>6</sup>We provide an easy-to-use open-source software package, *bayeslist*, in R to implement the Bayesian analysis of list experiments discussed in this paper.

<sup>7</sup>The “no design effect” and “no liars” assumptions identify the joint distribution of the latent sensitive item and number of control items,  $P(Z^*, Y^*)$  (Glynn, 2013). The variables in this joint distribution are then modeled as a mixture of the products of their respective submodels in Equations 4 and 5, which implies that the terms involving  $\delta$  and  $\psi_z$  are independent within each mixture component (see Imai (2011), Blair and Imai (2012), and the Supplementary Material). Furthermore, a common assumption (e.g., Imai, 2011; Blair and Imai, 2012) is to equate  $\psi_0 = \psi_1$  which implies independence of  $Z^*$  and  $\psi$ . Note that the last assumption is not strictly needed because we could still factor the joint prior distribution into  $p(\delta)p(Z^*|\delta)p(\psi|Z^*, \delta)$  without it.

Bayes rule provides a principled mechanism to *combine existing information*. So instead of injecting “subjectivity,” we actually make better use of the available evidence and rely *even more* on data than we would without informative priors.

Third, regarding the concerns with the biased nature of the information used, we respond in two ways. First, it is a well-known result that the accuracy of a statistical estimator (expressed as mean squared error) can be decomposed into bias and variance, i.e.,  $MSE = Bias^2 + Variance$  (e.g., James *et al.*, 2013). This re-expresses the well-known variance-bias tradeoff and suggests that the costs of accepting a little bias may be completely offset by the benefit of a decrease in variance. The overall net effect is improved accuracy. Second, we provide an explicit mechanism to deal with biased prior information via the so-called conjugate-distance prior (Wiśniowski *et al.*, 2020), and show that it yields unbiased estimates on average in the next section.

**3.4. Bias correction in prior specification**

While trading bias for increased efficiency may be a good bargain, we anticipate that many researchers will be hesitant to incorporate potentially biased information in their models of list experiments. This worry is clearly warranted if both the bias’s direction and magnitude are unknown. Fortunately, this is not a concern in the context of list experiments, where we often have information on the direction and size of the bias *by design*. In the following, we will demonstrate how to exploit this additional knowledge for bias correction in prior specification.

Our solution is inspired by recent advances in survey methodology that seek to combine inferences from probability-based and non-probability samples (Sakshaug *et al.*, 2019; Wiśniowski *et al.*, 2020; for a concise review, see Jerit and Barabas 2023). This research faces the related challenge of combining reliable but sparse data from expensive probability-based samples with less reliable, probably biased but abundant, cheap data from non-probability samples (e.g., commercial access panels). Using a Bayesian framework, this amounts to treating the probability-based sample as the “likelihood” and the non-probability sample as “prior information.” Their combination results in posterior survey estimates that have less variance (Sakshaug *et al.*, 2019). Just like our approach, this involves the deliberate incorporation of potentially biased information in the form of a prior to yield more efficient posterior estimates. Specifically, we consider the so-called “conjugate-distance” prior introduced in this literature (Wiśniowski *et al.*, 2020).

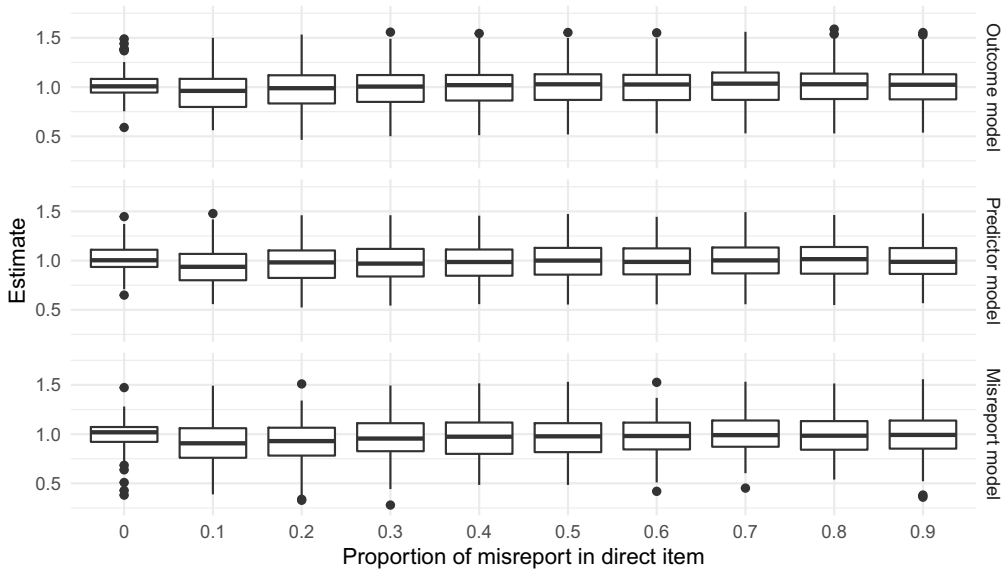
The conjugate-distance prior allows us to gain efficiency, while bounding the influence of biased information on the posterior. Let us assume a normal prior for the coefficient of covariate  $k$ , with prior mean equal to the point estimate  $\hat{\delta}_{D,k}$  from a direct item model. While *biased*, this still represents the best knowledge we have before running the sensitive item model. We then run the sensitive item model with an uninformative prior to get the *unbiased* estimate  $\hat{\delta}_{S,k}$ . The key ingredients of the conjugate-distance prior are the distance between  $\hat{\delta}_{D,k}$  and  $\hat{\delta}_{S,k}$ , and the standard errors  $\hat{\sigma}_{D,k}$  from the estimation of the direct items.

The intuition behind this prior specification is as follows. Whenever the difference between the potentially biased  $\hat{\delta}_{D,k}$  and the unbiased  $\hat{\delta}_{S,k}$  is smaller than the standard error  $\hat{\sigma}_{D,k}$  from the direct item, we can just use this standard error to shrink the prior distribution. Whenever there is greater disagreement between the biased and unbiased coefficient estimates, we instead let the prior variances be proportional to the squared differences between the two sets of estimates. The prior variance is thus calculated as

$$\hat{\sigma}_{P,k}^2 = \max \left[ (\hat{\delta}_{D,k} - \hat{\delta}_{S,k})^2, \hat{\sigma}_{D,k}^2 \right], \tag{9}$$

and for the complete prior specification, we use

$$\delta_k \sim \mathcal{N}(\hat{\delta}_{D,k}, \hat{\sigma}_{P,k}^2). \tag{10}$$



**Figure 1.** Estimates using conjugate-distance priors as misreport proportion in direct items increases (true coefficient value equals one).

Thereby, we allow the amount of sensitivity bias to rescale the prior variance and prevent the danger that the posterior distribution will be dominated by biased estimates.<sup>8</sup>

To demonstrate that the conjugate-distance prior indeed works as intended, we ran a detailed simulation study (see Supplementary Material D for details on the exact set-up of this simulation exercise). **Figure 1** shows the simulation results of using informative priors based on biased direct items scaled by the conjugate-distance method. The key takeaway is that regardless of how likely the respondents are to misreport their answers to a direct question, the resulting estimates of our suggested approach remain unbiased on average.

A possible remaining concern with the conjugate-distance prior is an alleged “double use” of the same data. But apart from the fact that no such concerns are raised in the related survey literature (e.g., Sakshaug *et al.*, 2019; Wiśniowski *et al.*, 2020; Jerit and Barabas, 2023), we note that almost all common prior modeling techniques are implicitly motivated by a reference likelihood (Gelman *et al.*, 2017). In fact, our approach has a close connection to “empirical Bayes” (EB) methods. EB methods have in common that they empirically estimate some part of the Bayesian procedure (see Efron (2024) for a recent overview). In this sense, our use of the “conjugate-distance prior” can also be labeled EB because it uses empirical data to adjust the variance of the prior distribution. We provide a more detailed discussion of the “double-use” problem and why it is easy to dispel in Supplementary Material E.

#### 4. Illustration of Bayesian analysis of list experiments using informative priors

We now illustrate our Bayesian approach to list experiments in a real-world application. Specifically, we identify commonly available prior evidence for list experiments, detail the steps needed to quantify prior evidence as prior probability distributions, and demonstrate how combining list experiments with prior evidence improves our substantive understanding of sensitive topics in the social sciences.

<sup>8</sup>Note that  $\hat{\sigma}_{p,k}^2$  may also be multiplied by a scalar, e.g.,  $k_0 = 1/N_D$  to modify the prior variance based on the degree of freedom  $N_D$  of the direct item data (Wiśniowski *et al.*, 2020). That is,  $\delta_k \sim \mathcal{N}(\hat{\delta}_{D,k}, k_0 \hat{\sigma}_{p,k}^2)$ .

A very useful and straightforward source of prior information is direct survey items about the sensitive issue (Aronow *et al.*, 2015). Indeed, around one hundred of the list experiment studies summarized in Blair *et al.* (2020) also asked respondents direct sensitive questions. These direct questions are usually asked for the purpose of comparison and are often biased due to social desirability. What is less emphasized is the fact that these direct questions are often biased in a systematic way and thus, still provide valuable information that can be exploited (Aronow *et al.*, 2015).

As an illustrative example, we consider the study by Gonzalez-Ocantos *et al.* (2012) on vote buying in Nicaragua. Political clientelism, where electoral support is traded for particularistic favors, is not only a problem in Latin American politics but plagues many democracies around the world (Mares and Young, 2016). Because of the sensitive nature of vote buying, Gonzalez-Ocantos *et al.* (2012) use a list experiment that elicits whether respondents were ever offered a gift for their vote. They also include two direct survey questions asking about “individual gifts” and “neighborhood gifts.” The study shows that while less than 3% report individual vote buying behavior when asked directly, some 18% report vote buying behavior in their neighborhood. The list experiment yields an even higher prevalence rate of 24%. But next to the question of the true extent of this form of political clientelism, there is also considerable debate about the characteristics of voters who are most likely to experience and respond to such economic inducements (Stokes, 2005; Nichter, 2008; Mares and Young, 2016). As we will demonstrate, including direct items as prior information also has the potential to change the substantive conclusions regarding the determinants of vote buying.

We would like to stress that our Bayesian framework for the analysis of list experiments also applies to a wider range of data situations and is not restricted to the use of direct questions as prior information. In the Supplementary Material, we provide two more real-world applications. The first demonstrates the use of auxiliary information on the prevalence of the sensitive item, including in the form of a “conjugate-distance” prior (see Supplementary Material F). The second application illustrates how our Bayesian approach can be used to analyze a specialized design called the “double-list experiment,” providing a framework for combining different experiments via the use of informative priors (see Supplementary Material G).

#### 4.1. Prior specification and sequential Bayesian updating

To incorporate direct items as prior information into the estimation of list experiments, we suggest the following simple procedure to derive prior distributions for the coefficients in the sensitive item outcome model:

##### Prior Specification Using Direct Items

- **Step 1.** Run the model without priors and using direct item  $D$  as the outcome to obtain coefficient estimates  $\hat{\delta}_{D,k}$  and standard errors  $\hat{\sigma}_{D,k}$  for each covariate  $k \in \{0, \dots, K\}$ , where  $K$  is the number of covariates and  $k = 0$  for the intercept.
- **Step 2.** Use these estimates to inform the location and spread of the priors for the sensitive item outcome model of the list experiment, e.g.,  $\delta_k \sim \mathcal{N}(\hat{\delta}_{D,k}, \hat{\sigma}_{D,k}^2)$  for each covariate  $k$ .

Since a Bayesian model with the correctly bounded uniform priors is equivalent to a likelihood model without priors (Diaconis and Freedman, 1986, 1990), the first step is to simply run the desired model without priors but using the direct item,  $D$ , as the outcome. Because this is usually a binary variable, a popular model choice is to use logistic regression:

$$P(D = 1) = \text{logit}^{-1}(\mathbf{X}'_i \hat{\boldsymbol{\delta}}_D). \quad (11)$$

This is convenient because we can now directly work with these model estimates  $\hat{\boldsymbol{\delta}}_D$  and their associated standard errors  $\hat{\boldsymbol{\sigma}}_D$  to form our prior distributions on the logit scale without having to

worry about transformations. More generally, in the presence of  $k$  covariates, logistic regression based on direct items will yield a prior for the coefficient of the  $k$ -th covariate  $\delta_k$  with mean  $\hat{\delta}_{D,k}$  and the standard error  $\hat{\sigma}_{D,k}$  (note the use of subscripts, e.g.,  $D$ , to distinguish priors and model parameters). Relying on the central limit theorem, we are able to assume that coefficients in the sensitive item outcome model will be normally distributed and thus pick a prior distribution for  $\delta_k$  which also follows a normal distribution.

In the second step, we can impose this normal distribution as an informative prior on  $\delta_k$  in the sensitive item outcome model, simply plugging in the coefficient estimates and (squared) standard errors as the required means and variances:<sup>9</sup>

$$\delta_k \sim \mathcal{N}(\hat{\delta}_{D,k}, \hat{\sigma}_{D,k}^2), \tag{12}$$

where the actual derivation of the posterior distribution and its marginalization are then left to the statistical software. Note that in contrast to the method described by Aronow *et al.* (2015), the procedure outlined above allows for multivariate analysis of list experiments and is also applicable when direct questions and list experiments are answered by distinct sets of respondents or part of completely independent studies.

To make things more concrete, we turn to the vote buying example. Which voter characteristics explain whether they received a gift or favor in exchange for their vote? A logistic regression model for the direct report of individual gift receiving behavior yields the following coefficient estimates (with standard errors in parentheses):

$$\begin{aligned} P(\text{Individual.Gift} = 1) = & \text{logit}^{-1}(-3.58(1.07) + 1.31(0.56) \times \text{Support.FSLN} \\ & -0.16(0.87) \times \text{Support.PLC} + 1.20(0.63) \times \text{Very.Poor} + 0.62(0.60) \times \text{Poor} \\ & -0.33(0.29) \times \text{Education} - 0.46(0.31) \times \text{Age} - 0.40(0.43) \times \text{Male}). \end{aligned}$$

Thus, based on this *direct item model*, we would conclude that voters’ socio-demographic characteristics have little to do with their susceptibility to vote buying. For instance, poor and very poor voters do not differ significantly from more well-off voters in admitting individual vote buying (logit coefficient  $\hat{\delta}_{D,\text{Poor}} = 0.62$  with standard error  $\hat{\sigma}_{D,\text{Poor}} = 0.60$  and  $\hat{\delta}_{D,\text{Very.Poor}} = 1.20$  with  $\hat{\sigma}_{D,\text{Very.Poor}} = 0.63$ ). But interestingly, party supporters of the leftist party *Frente Sandinista de Liberación Nacional* (FSLN) are markedly more likely to be targeted by politicians and their brokers ( $\hat{\delta}_{D,\text{Support.FSLN}} = 1.31$  with  $\hat{\sigma}_{D,\text{Support.FSLN}} = 0.56$ ).

This clearly is valuable information and suggests the following prior distributions for the coefficients of the *sensitive item outcome model*:<sup>10</sup>

$$\begin{aligned} \delta_{\text{Intercept}} & \sim \mathcal{N}(-3.58, 1.07^2); & \delta_{\text{Support.FSLN}} & \sim \mathcal{N}(1.31, 0.56^2); & \delta_{\text{Support.PLC}} & \sim \mathcal{N}(-0.16, 0.87^2); \\ \delta_{\text{Very.Poor}} & \sim \mathcal{N}(1.20, 0.63^2); & \delta_{\text{Poor}} & \sim \mathcal{N}(0.62, 0.60^2); & \delta_{\text{Education}} & \sim \mathcal{N}(-0.33, 0.29^2); \\ \delta_{\text{Age}} & \sim \mathcal{N}(-0.46, 0.31^2); & \delta_{\text{Male}} & \sim \mathcal{N}(-0.40, 0.43^2). \end{aligned}$$

The resulting posterior distributions for all coefficients of the sensitive item model are presented in Figure 2 (Panel a). As is common practice, we summarize them in terms of prior means and 95% credible intervals. We also contrast them with the posterior estimates derived from a model with uninformative priors (i.e.,  $\delta_k \sim \mathcal{N}(0, 10^2)$  for all  $k$ ), which therefore only reflects the information provided by the list experiment alone. It is immediately apparent that informed priors improve the efficiency. By using the direct item “individual gifts” as the prior information, the variance of the

<sup>9</sup>While we use univariate normal distributions throughout the paper, we can also impose a multivariate normal distribution given a variance-covariance matrix of prior uncertainties.

<sup>10</sup>Note that in the original analysis, Gonzalez-Ocantos *et al.* (2012) used the ordinary least squares estimator with interactions between the treatment and covariates. While efficient, this estimator suffers from linear extrapolation where predictions can fall outside the range of item counts of list experiments.

estimated prevalence was reduced by 45% compared to the estimation using an uninformative normal prior (Panel a). The second direct item “neighborhood gifts” produces an even higher variance reduction rate of 75% (Panel b). But before we turn to the substantive interpretation of these results, it is worth considering whether even more prior information can be brought to the analysis.

When there is more than one direct item available, one could include this additional information by sequentially updating the prior information with the following steps:

#### Sequential Bayesian Updating for Several Direct Items

- **Step 1.** Run a model without priors and using the first direct item  $D_1$  as the outcome to obtain coefficient estimates  $\hat{\delta}_{D_1,k}$  and standard errors  $\hat{\sigma}_{D_1,k}$  for each covariate  $k \in \{0, \dots, K\}$  in the first direct item model.
- **Step 2.** Repeat the following procedure to incorporate all available information up to the  $q$ th direct item.  
For  $i$  from 1 to  $q - 1$ :
  - (i) Use the estimates from the  $i$ th direct item to produce informed priors for the next direct item model, e.g.,  $\delta_{D_{i+1},k} \sim \mathcal{N}(\hat{\delta}_{D_i,k}, \hat{\sigma}_{D_i,k}^2)$ .
  - (ii) Run a model with the above informed priors using the  $(i + 1)$ th direct item  $D_{i+1}$  as outcome to obtain coefficient estimates  $\hat{\delta}_{D_{i+1},k}$  and standard errors  $\hat{\sigma}_{D_{i+1},k}$  for each covariate  $k$ .
- **Step 3.** Use the final estimates to inform the priors for the sensitive item outcome model of the list experiment, e.g.,  $\delta_k \sim \mathcal{N}(\hat{\delta}_{D_q,k}, \hat{\sigma}_{D_q,k}^2)$  for each covariate  $k$ .

Applying the above procedure to the vote buying example, the estimates from the direct item on individual gift receiving behavior  $D_1$  can now be used to inform the model of the direct item on neighborhood gift receiving behavior  $D_2$  (i.e., instead of  $\delta_k$  we put the priors on  $\delta_{D_2,k}$ ). Following the logic of sequential updating, the resulting coefficient estimates and standard errors can then, in turn, be used to inform the prior distributions for the coefficients of the sensitive item outcome model.

The informative priors that contain the information of two direct items ( $D_1$  and  $D_2$ ) will have considerably smaller variances. For instance, the prior for the coefficient of FSLN support is now:

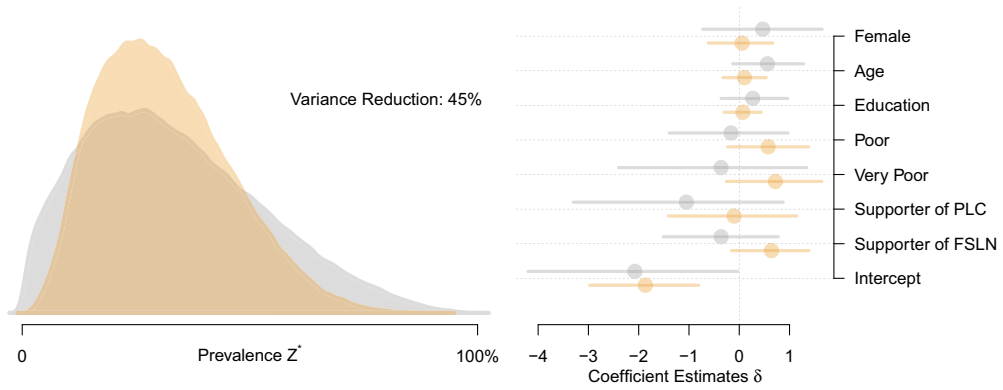
$$\delta_{\text{Support.FSLN}} \sim \mathcal{N}(0.45, 0.19^2),$$

with variance  $\sigma_{\text{Support.FSLN}}^2 = 0.19^2$ , which is much smaller than the prior variance suggested by the first direct item model alone,  $\sigma_{\text{Support.FSLN}}^2 = 0.56^2$ . Note that sequential updating yields exactly the same posterior as combining all information *at the same time*. Put differently, the order in which we update does not matter. Panel c of [Figure 2](#) shows that incorporating both direct items significantly improves the efficiency of the list experiment analysis. Compared to Panels a and b, where only one direct item is incorporated, the variance of the posterior prevalence in Panel c is reduced by 82% with the inclusion of both direct items.

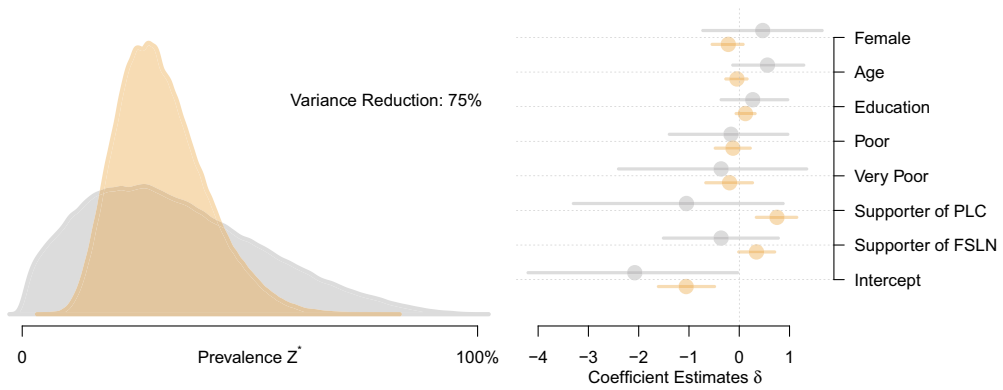
In addition to variance reduction for the estimated prevalence, the coefficient estimates also show a significant efficiency improvement when direct items are exploited as prior information. This in turn leads to substantively different results on the mechanisms of vote buying. A central question and long-standing debate in the literature concerns who is more likely to be targeted for vote buying: swing voters and those weakly opposed (Stokes, 2005; Stokes *et al.*, 2013) or core party supporters (Nichter, 2008; Calvo and Murillo, 2013; Albertus, 2015). Relying on the results of the list experiment *alone*, we find that there are no statistically reliable differences between these voter groups. The coefficient for supporters of the FSLN is -0.36 [-1.51, .078] and for supporters of the PLC -1.05 [-3.30, 0.87] (see [Figure 2](#)). Accordingly, Gonzalez-Ocantos *et al.* (2012), 210 conclude that there is “little support for Nichter’s (2008) argument that machines mainly target core supporters in an effort to buy turnout.”

But this conclusion may simply be due to the high variance and indeed changes as soon as we consider the direct item about neighborhood vote buying as prior evidence. By incorporating this additional piece of information into the model, we find that supporters of the center-right party

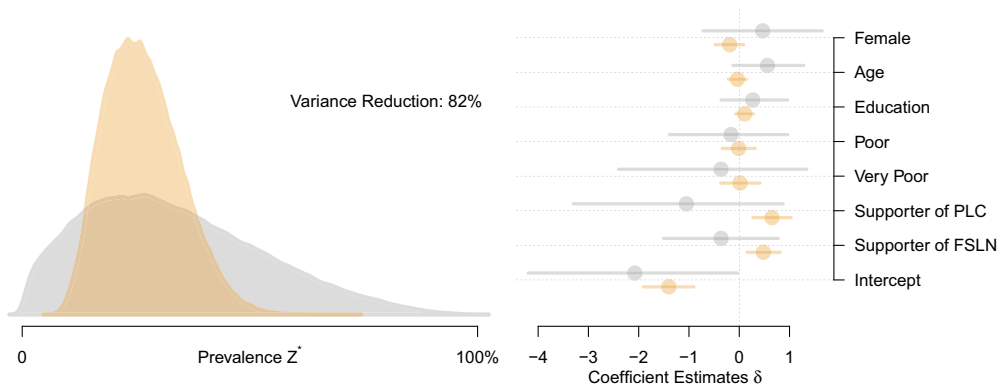
(a) Direct Item: Individual Gifts



(b) Direct Item: Neighbor Gifts



(c) Two Direct Items Combined



**Figure 2.** Comparison between estimates using uninformative (highlighted in gray) and informative priors (highlighted in orange) from direct items on vote buying (posterior means along with 95% credible intervals).

*Partido Liberal Constitucionalista* (PLC) are reliably more susceptible to vote buying than independent voters (logit coefficient of 0.75 with 95% credible interval of [0.34, 1.14]). Sequentially updating

the priors and therefore exploiting the information contained in both direct items leads to further substantive insights. Now, not only supporters of the PLC but also supporters of the FSLN are more likely to report being targeted for vote buying. Combining the list experiment with prior information lends support to the notion that it is core party supporters rather than swing voters who are approached by parties. This changes the substantive conclusion of the original study by Gonzalez-Ocantos *et al.* (2012) and suggests that the findings for Nicaragua in fact correspond to those in other Latin American countries such as Argentina (Nichter, 2008), Chile (Calvo and Murillo, 2013), and Venezuela (Albertus, 2015).

#### 4.2. Bias correction in prior specification

Despite the improved efficiency of our Bayesian approach, researchers may still be concerned about the potential bias of the direct item, which derives from voters' tendency to under-report this particular behavior. To address this potential bias and provide a meaningful interpretation, we utilize the conjugate-distance prior.

To illustrate the conjugate-distance prior, we turn to a *sensitive item misreporting model* (Eady, 2017), which adds a separate equation for misreporting to the sensitive outcome and control times equations (see Equation A3 in the Supplementary Material A for more details). For the coefficients in this misreporting equation,  $\gamma$ , we simply choose priors from a direct item model of "neighborhood gifts," e.g.

$$\begin{aligned}\gamma_{\text{Support.PLC}} &\sim \mathcal{N}(0.35, 0.20^2), \\ \gamma_{\text{Support.FSLN}} &\sim \mathcal{N}(0.79, 0.22^2).\end{aligned}$$

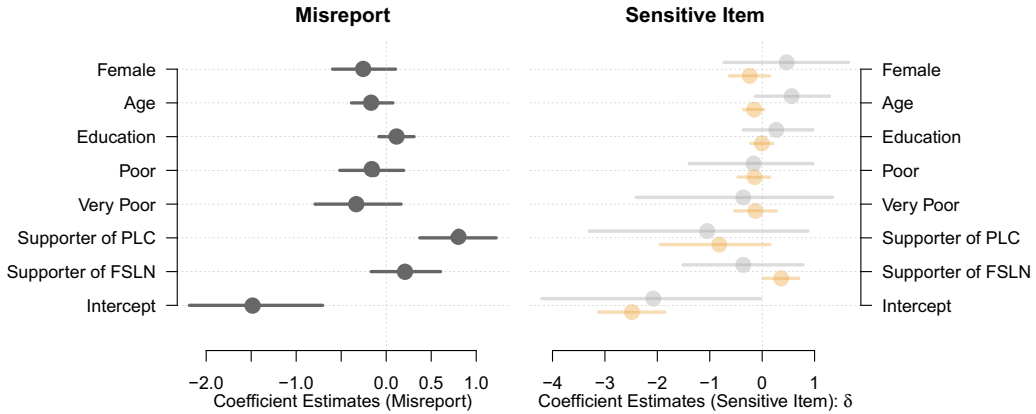
Note that these priors are highly informative for misreporting because they explicitly capture respondents' biased answering behavior. In contrast, for the coefficients in the sensitive outcome equation,  $\delta$ , we would instead want to correct for this known bias by specifying the following conjugate-distance priors

$$\begin{aligned}\delta_{\text{Support.PLC}} &\sim \mathcal{N}(0.35, \hat{\sigma}_{P,\text{Support.PLC}}^2), \\ \delta_{\text{Support.FSLN}} &\sim \mathcal{N}(0.79, \hat{\sigma}_{P,\text{Support.FSLN}}^2),\end{aligned}$$

where

$$\begin{aligned}\hat{\sigma}_{P,\text{Support.PLC}}^2 &= \max [(0.35 - (-0.36))^2, 0.20^2] = 0.50, \\ \hat{\sigma}_{P,\text{Support.FSLN}}^2 &= \max [(0.79 - (-1.05))^2, 0.22^2] = 3.39.\end{aligned}$$

The results in Figure 3 suggest that when it comes to misreporting "individual gifts," supporters of the PLC are reliably more likely to misreport than independents, while no such difference exists for supporters of the FSLN. This is an interesting finding, one that has not been captured in the original study by Gonzalez-Ocantos *et al.* (2012). Looking at the sensitive item outcome of individual gift receiving, we first note that for supporters of both parties, it is better to rely on wider prior variances which reflect the biased responses to the direct "neighborhood gifts" item. But even when correcting for the obvious bias in direct items using a conjugate-distance prior, the posterior coefficient estimates are still considerably more efficient. And again, this prior information has important substantive implications. While supporters of the center-right party PLC are now no longer different from independents (-0.82 [-1.94, 0.15]), we find reliable evidence that supporters of the leftist FSLN are more likely to be targeted for vote buying than independents (0.36 [0.02, 0.70]). Again, this qualifies the original findings in Gonzalez-Ocantos *et al.* (2012) and has important substantive implications for the debate on the partisan targets of vote buying in Latin America and beyond.



**Figure 3.** Comparison between estimates using uninformative (highlighted in gray) and conjugate-distance priors (highlighted in orange) from direct items on vote buying (estimates of misreporting model on the left; posterior means along with 95% credible intervals).

### 5. Conclusion

We have introduced the Bayesian analysis of list experiments using informative priors. This way, we generalize a whole range of different modeling approaches for list experiments under a common and principled framework. Importantly, we can think of our Bayesian approach as a *modular framework* where different approaches can be mixed and matched as required and desired by applied researchers. For instance, one could use *auxiliary information* on the prevalence of some sensitive behavior to inform a regression model for a *direct survey item*. These combined estimates could then be used to derive updated priors for a model of a *first list experiment*. Its results could then inform a model for a *second list experiment*, and so on. All of this is achieved by following a simple Bayesian logic: specifying the respective informative prior distributions and updating sequentially.

We give the following practical suggestions to applied researchers who wish to improve the efficiency of their list experiments. In terms of *sources* of prior information, we encourage applied researchers to consider a wide array of information from both external sources (e.g., archival reports, expert opinions, and qualitative accounts) and additional design elements (e.g., direct items, repeated measures, and double list experiments as well as other experimental techniques). The latter are a particularly underappreciated source of prior information. In terms of prior *forms*, researchers primarily interested in the prevalence of sensitive items can express their prior intuitively using a beta distribution. If priors for multivariate models of list experiments are desired, informative normal priors on coefficients are a straightforward solution.

While our focus is on the *analysis* of list experiments, our approach also provides a straightforward *design* recommendation for list experiments. Whenever possible, researchers should include one or several direct questions of the sensitive item in the survey. Although estimates based on the direct items are likely to be biased, they have low variance and can be combined with list experiments to considerably increase the overall efficiency. Concerns over bias are mitigated by relying on the conjugate-distance prior formulation. Adding a simple direct item is much easier to implement than classical design suggestions, such as negatively correlated control items (Glynn, 2013) or designing a double list experiment (Droitcour *et al.*, 1991). This being said, and to reiterate, these design elements are not at all exclusive but may be combined via Bayesian updating to further maximize efficiency and change our substantive understanding of sensitive topics in the social sciences.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/psrm.2025.10084>. To obtain replication material for this article, please visit <https://doi.org/10.7910/DVN/QE1XNM>.

**Acknowledgements.** The previous versions of the paper have been presented at the 2021 APSA Annual Meeting, the 8th Asian Political Methodology Meeting, and the Data Science in Action Lecture Series of the Mannheim Center for Data Science. We thank Daina Chiba, Kentaro Fukumoto, Jeff Gill, Erin Hartman, Thomas König, Patrick Kuhn, Yoshikuni Ono, and Marc Ratkovic for their helpful comments and suggestions.

**Competing interests.** The authors declare none.

## References

- Ahlquist JS (2018) List experiment design, non-strategic respondent error, and item count technique estimators. *Political Analysis* 26, 34–53.
- Albertus M (2015) The role of subnational politicians in distributive politics: political bias in Venezuela's land reform under Chávez. *Comparative Political Studies* 48, 1667–1710.
- Aronow PM, Coppock A, Crawford FW and Green DP (2015) Combining list experiment and direct question estimates of sensitive behavior prevalence. *Journal of Survey Statistics and Methodology* 3, 43–66.
- Bentancur P, V and Tiscornia L (2023) Iteration in mixed-methods research designs combining experiments and fieldwork. *Sociological Methods & Research* 53, 729–759.
- Blair G, Chou W and Imai K (2019) List experiments with measurement error. *Political Analysis* 27, 455–480.
- Blair G, Coppock A and Moor M (2020) When to worry about sensitivity bias: a social reference theory and evidence from 30 years of list experiments. *American Political Science Review* 114, 1297–1315.
- Blair G and Imai K (2012) Statistical analysis of list experiments. *Political Analysis* 20, 47–77.
- Blair G, Imai K and Lyall J (2014) Comparing and combining list and endorsement experiments: evidence from Afghanistan. *American Journal of Political Science* 58, 1043–1063.
- Breznau N, Rinke EM, Wuttke A, Nguyen HHV, Adem M, Adriaans J, Alvarez-Benjumea A, Andersen HK, Auer D and Azevedo F Bahnsen O (2022) Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proceedings of the National Academy of Sciences* 119, e2203150119.
- Calvo E and Murillo MV (2013) When parties meet voters: assessing political linkages through partisan networks and distributive expectations in Argentina and Chile. *Comparative Political Studies* 46, 851–882.
- Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Peter L and Riddell A (2017) Stan: a probabilistic programming language. *Journal of Statistical Software* 76, 1–32.
- Chou W, Imai K and Rosenfeld B (2020) Sensitive survey questions with auxiliary information. *Sociological Methods & Research* 49, 418–454.
- Conley D and McCabe BJ (2011) Body mass index and physical attractiveness: evidence from a combination image-alteration/list experiment. *Sociological Methods & Research* 40, 6–31.
- Corstange D (2009) Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis* 17, 45–63.
- Coutts E and Jann B (2011) Sensitive questions in online surveys: experimental results for the randomized response technique (RRT) and the unmatched count technique (UCT). *Sociological Methods & Research* 40, 169–193.
- Diaconis P and Freedman D (1986) On the consistency of Bayes estimates. *The Annals of Statistics* 14, 1–26.
- Diaconis P and Freedman D (1990) On the uniform consistency of Bayes estimates for multinomial probabilities. *The Annals of Statistics* 18, 1317–1327.
- Caspar RA, Hubbard ML, Parsley TL, Visscher W and Ezzati TM (1991) The item-count technique as a method of indirect questioning: A review of its development and a case study application. In Biemer, PP, RM Groves, LE Lyberg, NA Mathiowetz and S Sudman (eds), *Measurement Errors in Surveys*. New York: John Wiley & Sons, pp.185–210.
- Eady G (2017) The statistical analysis of misreporting on sensitive survey questions. *Political Analysis* 25, 241–259.
- Efron B (2024) Empirical Bayes: Concepts and Methods. In Berger, J, X-L Meng, N Reid and M Xie (eds.) *Handbook of Bayesian, Fiducial, and Frequentist Inference*. New York: Chapman and Hall/CRC, pp. 8–34.
- Gelman A, Simpson D and Betancourt M (2017) The prior can often only be understood in the context of the likelihood. *Entropy* 19, 555.
- Gerber AS and Green DP (2012) *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton.
- Gill J (2015) *Bayesian Methods: A Social and Behavioral Sciences Approach*, 20. New York: CRC Press.
- Gill J and Walker LD (2005) Elicited priors for Bayesian model specifications in political science research. *The Journal of Politics* 67, 841–872.
- Glynn AN (2013) What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly* 77, 159–172.
- Glynn AN and Ichino N (2015) Using qualitative information to improve causal inference. *American Journal of Political Science* 59, 1055–1071.
- Gonzalez-Ocantos E, De Jonge CK, Meléndez C, Osorio J and Nickerson DW (2012) Vote buying and social desirability bias: experimental evidence from Nicaragua. *American Journal of Political Science* 56, 202–217.

- Groenitz H** (2014) Improvements and extensions of the item count technique. *Electronic Journal of Statistics* **8**, 2321–2351.
- Groenitz H** (2015) Using prior information in privacy-protecting survey designs for categorical sensitive variables. *Statistical Papers* **56**, 167–189.
- Humphreys M and Jacobs AM** (2015) Mixing methods: a Bayesian approach. *American Political Science Review* **109**, 653–673.
- Imai K** (2011) Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association* **106**, 407–416.
- Imai K, Park B and Greene KF** (2015) Using the predicted responses from list experiments as explanatory variables in regression models. *Political Analysis* **23**, 180–196.
- Jackman S** (2000) Estimation and inference via Bayesian simulation: an introduction to Markov chain Monte Carlo. *American Journal of Political Science* **44**, 375–404.
- Jackman S** (2009) *Bayesian Analysis for the Social Sciences*. Hoboken: John Wiley & Sons.
- James G, Witten D, Hastie T and Tibshirani R** (2013) *An Introduction to Statistical Learning With Applications in R*. London: Springer.
- Jerit J and Barabas J** (2023) Are nonprobability surveys fit for purpose? *Public Opinion Quarterly* **87**, 816–840.
- Jiayuan L and Van den Noortgate W** (2022) A meta-analysis of the relative effectiveness of the item count technique compared to direct questioning. *Sociological Methods & Research* **51**, 760–799.
- Kuhn PM and Vivyan N** (2021) The misreporting trade-off between list experiments and direct questions in practice: partition validation evidence from two countries. *Political Analysis* **30**, 381–402.
- Liu Y, Tian G-L, Qin W and Tang M-L** (2019) Poisson–Poisson item count techniques for surveys with sensitive discrete quantitative data. *Statistical Papers* **60**, 1763–1791.
- Mares I and Young L** (2016) Buying, expropriating, and stealing votes. *Annual Review of Political Science* **19**, 267–288.
- McElreath R** (2018) *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*. New York: Chapman and Hall/CRC.
- Miller JD** (1984) *A New Survey Technique for Studying Deviant Behavior*, Ph.D. thesis. Washington, DC: The George Washington University.
- Nichter S** (2008) Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review* **102**, 19–31.
- Northrup JM and Gerber BD** (2018) A comment on priors for Bayesian occupancy models. *PLOS One* **13**, 1–13.
- Plummer M** (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. **124**, 1–8.
- Rosenfeld B, Imai K and Shapiro JN** (2016) An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science* **60**, 783–802.
- Sakshaug JW, Wiśniowski A, Ruiz DAP and Blom AG** (2019) Supplementing small probability samples with nonprobability samples: a Bayesian approach. *Journal of Official Statistics* **35**, 653–681.
- Seaman III, John W, John W Seaman J and James DS** (2012) Hidden dangers of specifying noninformative priors. *The American Statistician* **66**, 77–84.
- Simmons JP, Nelson LD and Simonsohn U** (2011) False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* **22**, 1359–1366.
- Stokes SC, Dunning T, Nazareno M and Brusco V** (2013) *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge: Cambridge University Press.
- Stokes SC** (2005) Perverse accountability: a formal model of machine politics with evidence from Argentina. *American Political Science Review* **99**, 315–325.
- Western B and Jackman S** (1994) Bayesian inference for comparative research. *American Political Science Review* **88**, 412–423.
- Wiśniowski A, Sakshaug JW, Ruiz DAP and Blom AG** (2020) Integrating probability and nonprobability samples for survey inference. *Journal of Survey Statistics and Methodology* **8**, 120–147.
- Yamaguchi K** (2016) Cross-sectional and panel data analyses of an incompletely observed variable derived from the nonrandomized method for surveying sensitive questions. *Sociological Methods & Research* **45**, 41–68.