Measures of Non-Ignorable Selection Bias for Non-Probability Samples

Sergio D. Martinez[1], **Brady T. West**[2], **and Rebecca R. Andridge**[3]

[1]PhD(c) in Survey and Data Science, University of Michigan, sergiomm@umich.edu
[2]Research Professor, Survey Research Center of the Institute for Social Research, University of Michigan, bwest@umich.edu
[3]Professor, Division of Biostatistics in the College of Public Health, The Ohio State University, andridge.1@osu.edu

Abstract

Non-probability samples are increasingly used in applied research, raising concerns about non-ignorable selection bias in estimates based on these samples that traditional diagnostics cannot adequately assess. Conventional diagnostics and inferential approaches for these samples offer limited insight because they ignore the link between selection mechanisms and outcomes. This paper reviews variable-dependent measures for non-ignorable selection bias based on the proxy pattern–mixture model (PPMM), with emphasis on the Standardized Measure of Unadjusted Bias (SMUB) for means and the Measure of Unadjusted Bias for Proportions (MUBP). Both indices are grounded in the PPMM framework, which uses auxiliary variables with reliable population benchmarks to construct a single proxy and summarize departures from ignorability through a single sensitivity parameter. Evidence from simulation studies shows that the SMUB and MUBP can accurately capture the direction and magnitude of bias when auxiliary variables are at least moderately predictive of outcomes, outperforming traditional diagnostics. Empirical applications in health surveys, political polling, pandemic surveillance, and administrative data confirm their practical value while emphasizing the need for strong, harmonized auxiliary covariates. We conclude with guidance on implementation and a brief discussion of ongoing research. Our aim is to motivate broader adoption of these indices as practical and interpretable diagnostics for assessing selection bias in applied survey research, aided by accessible `R` software that facilitates their implementation in practice.

Keywords: selection bias, non-probability samples, proxy pattern–mixture model, sensitivity analysis.

1 Introduction

The cornerstone of survey inference has long rested on a fundamental assumption: that the mechanism by which units are selected into a sample does not depend on the values of the variables being measured. This assumption of ignorable selection, formalized by Rubin (1976), provides the theoretical justification for design-based inference from probability samples. However, the contemporary landscape of survey research presents mounting challenges to this ideal. Declining response rates across all survey modes (Brick and Williams, 2013; Williams and Brick, 2018; de Leeuw, Hox, and Luiten, 2018; Luiten, Hox, and de Leeuw, 2020; Daikeler, Bošnjak, and Lozar Manfreda, 2020; Lundmark and Backström, 2025), escalating costs of maintaining probability samples (Presser and McCulloch, 2011), and the proliferation of readily available non-probability data sources (Baker et al., 2013; Mercer et al., 2017; Cornesse et al., 2020) have created an environment where ignorability can no longer be taken for granted.

Non-probability samples, which lack a formal randomization mechanism, present particular challenges for inference. Unlike probability samples where design-based methods can in principle account for known selection probabilities, non-probability samples require model-based approaches. Elliott and Valliant (2017) outlined two broad approaches for making inferences under this setting: quasi-randomization and superpopulation modeling. Both approaches critically depend on the assumption of ignorable selection, which in practice is unlikely to hold precisely, yet existing adjustment methods provide little guidance on the magnitude of bias that may result from its violation. This gap motivates the development of sensitivity analysis tools that explicitly parameterize departures from ignorability and quantify their impact on estimates of interest.

The methodological response to this challenge has centered on developing model-based approaches that explicitly parameterize departures from ignorability. A particularly influential framework emerged from the work of Andridge and Little, 2011, who introduced the proxy pattern-mixture model (PPMM) as a principled method for sensitivity analysis in the presence of non-ignorable survey non-response. The PPMM compresses auxiliary information observed for both respondents and nonrespondents into a single proxy variable that is predictive of the outcome of interest. By modeling the joint distribution of the proxy and the outcome under different assumptions about the missing data mechanism, the PPMM provides a structured sensitivity analysis for non-response bias.

Building on the PPMM framework, Little et al. (2020) introduced the Standardized Measure of Unadjusted Bias (SMUB), a family of interpretable indices that quantify the degree of departure from ignorable selection in estimated means. Boonstra et al. (2021) later demonstrated that the SMUB correlates more strongly with true bias than traditional diagnostics. Extending this framework to binary outcomes, Andridge et al. (2019) developed the Measure of Unadjusted Bias for Proportions (MUBP), which reformulates the PPMM within a latent variable setting. More recently, West et al. (2021) generalized this framework to regression coefficients in both linear and probit models.

To compute the SMUB and MUBP, three ingredients are required: (1) microdata from a non-probability sample containing the outcome Y (continuous for SMUB or binary for MUBP) and a set of auxiliary variables Z that are predictive of Y; (2) reliable population-level summaries of Z, including means and, when available, variances and covariances, obtained from high-quality data sources such as administrative registers, large probability surveys, or other external benchmarks; and (3) an assumed sensitivity parameter ϕ that governs the degree of non-ignorability. In general terms, the estimation proceeds as follows. A proxy variable X is first constructed for the outcome Y by regressing Y on the auxiliary variables Z using the non-probability sample data, with linear regression used for SMUB and probit regression for MUBP. This proxy represents the best available predictor of the outcome based on the observed covariates, reducing a multidimensional set of auxiliaries to a single composite predictor. Population-level summaries of Z are then used to compute the corresponding population mean and variance of the proxy X. Finally, SMUB(ϕ) or MUBP(ϕ) can be obtained either by fixing specific values for ϕ (commonly 0, 0.5, and 1) or, under a Bayesian framework, by assigning ϕ a noninformative prior distribution that reflects the absence of prior knowledge about the degree of non-ignorability.

The predictive strength of the auxiliary variables plays a central role in this framework. Weakly predictive Z variables may yield highly uncertain bias estimates, making it difficult to assess the direction or magnitude of potential selection bias, reducing the diagnostic value of the indices. Correlations above approximately 0.3–0.4 between the outcome and the proxy are desirable (Andridge et al., 2019; Little et al., 2020). As discussed in Section 3, both simulations and empirical results reinforce the importance of having a strongly predictive set of covariates to ensure reliable inference.

Empirical studies have demonstrated the versatility of these indices across diverse survey domains.

West and Andridge (2023) applied the MUBP to pre-election polling data, showing improved alignment with certified election results; Andridge (2024) used the MUBP to assess bias in COVID-19 vaccine uptake estimates, finding results consistent with non-ignorable selection; Hammon and Zinn (2024) validated the MUBP using population data from the German General Social Survey; and Schroeder and West (2025) applied the MUBP to evaluate potential selection bias in the 2019 Health Survey Mailer (HSM), an off-wave supplement to the longitudinal Health and Retirement Study (HRS). Using harmonized demographic and health covariates shared across waves, they found that MUBP adjustments were small, indicating largely ignorable selection but highlighting the method's value for diagnosing bias in longitudinal survey contexts. Most recently, Gómez-Echeverry et al. (2025) applied the SMUB framework to short-term economic indicators, highlighting the benefits of incorporating historical auxiliary data to improve adjustment accuracy.

This article provides an overview of the theoretical foundations, empirical performance, and practical implementation of the SMUB and MUBP indices. We begin by outlining the PPMM framework that underlies both measures. We then synthesize evidence from simulation studies that systematically vary proxy strength, selection mechanisms, and outcome distributions, together with validation exercises and empirical applications spanning public health surveillance, demographic surveys, political polling, and administrative data used for economic indicators. Finally, we summarize practical guidance on proxy construction and sensitivity analysis, and discuss methodological extensions already available as well as ongoing research aimed at refining these indices and broadening their applicability. The overarching goal is to show that the SMUB and MUBP constitute accessible, interpretable, and empirically validated tools for diagnosing selection bias, and to motivate their broader adoption in contemporary survey research, where evaluating data quality under potential non-ignorable selection has become increasingly critical.

2 Measures of selection bias

This section summarizes the methodological framework and key formulations introduced by Little et al. (2020) and Andridge et al. (2019). The following exposition outlines the main components, assumptions, and analytical expressions underlying these measures. Readers interested in full derivations and implementation details are referred to the original papers for comprehensive discussions.

2.1 Indices of Non-Ignorable Selection Bias for Means

Little et al. (2020) developed an index-based sensitivity analysis framework that explicitly quantifies potential selection bias under varying assumptions about the degree of non-ignorability. Their approach embeds the PPMM into a tractable parametric framework that enables researchers to bound the range of plausible bias values and assess the robustness of substantive conclusions to departures from ignorable selection.

Suppose the non-probability sample provides data $\mathcal{D} = \{(y_i, z_i) : i = 1, \ldots, n\}$, where y_i denotes the continuous survey outcome for unit i and z_i is a p-dimensional vector of auxiliary variables predictive of y_i and for which summary statistics are available for the population. Let $S_i \in \{0, 1\}$ indicate selection into the non-probability sample, with $S_i = 1$ for selected units and $S_i = 0$ otherwise.

The first step constructs an auxiliary proxy for the unobserved outcome values among non-selected units. Formally, we regress Y on Z using data from selected units ($S = 1$) to obtain the fitted linear predictor $X = Z^\top \hat{\beta}$ where $\hat{\beta}$ denotes the least-squares coefficient vector. This proxy X represents the best linear predictor of Y based on the available auxiliaries and serves as a surrogate for Y in the non-selected population. To facilitate model specification and interpretation, X is rescaled to match the variance of Y within the non-probability sample $X^* = X\sqrt{\frac{s_{YY}^{(1)}}{s_{XX}^{(1)}}}$ where $s_{YY}^{(1)}$ and $s_{XX}^{(1)}$ denote sample variances among selected units. For notational convenience, we denote the rescaled proxy from now on as X, with the rescaling implicit.

The PPMM assumes that the joint distribution of (Y, X) follows a bivariate normal distribution conditional on selection status S:

$$(Y, X) \mid S = j \sim \mathcal{N}_2\left(\begin{pmatrix}\mu_Y^{(j)} \\ \mu_X^{(j)}\end{pmatrix}, \begin{pmatrix}\sigma_{YY}^{(j)} & \sigma_{XY}^{(j)} \\ \sigma_{XY}^{(j)} & \sigma_{XX}^{(j)}\end{pmatrix}\right), \quad j \in \{0, 1\}. \tag{1}$$

Some parameters governing the distribution of (Y, X) among nonselected units ($j = 0$) are not identified from the observed data $(\mu_Y^{(0)}, \sigma_{YY}^{(0)}, \sigma_{XY}^{(0)})$. Identification is achieved by assuming that the probability of selection depends on (X, Y) through a scalar index formed as a convex combination of the two variables:

$$\Pr(S = 1 \mid X, Y) = g\big((1-\phi)X + \phi Y\big),$$

where $g : \mathbb{R} \to (0, 1)$ is an unspecified monotonic function and $\phi \in [0, 1]$ is a scalar sensitivity parameter.

The parameter ϕ quantifies the *degree of non-ignorability* and admits an intuitive interpretation. When $\phi = 0$, selection depends only on the observed proxy X, corresponding to *selection at random* (SAR) conditional on Z, which represents ignorable selection. When $\phi = 1$, selection depends entirely on the outcome Y, representing fully non-ignorable selection where the auxiliary variables provide no direct information about the selection mechanism. For intermediate values $0 < \phi < 1$, selection depends on both X and Y, with larger values indicating stronger dependence on the unobserved outcome.

The maximum likelihood estimator for the population mean of Y as a function of ϕ is given by:

$$\hat{\mu}_Y(\phi) = \bar{y}^{(1)} + \frac{\phi + (1-\phi)\hat{\rho}_{XY}^{(1)}}{\phi\hat{\rho}_{XY}^{(1)} + (1-\phi)}\sqrt{\frac{s_{YY}^{(1)}}{s_{XX}^{(1)}}}\,(\bar{x}^{(1)} - \bar{X}),$$

where $\bar{y}^{(1)}$ and $\bar{x}^{(1)}$ denote sample means among selected units, $\bar{X}$ is the known population mean of X (computed from population-level summaries of Z and the estimated coefficients $\hat{\beta}$), and $\hat{\rho}_{XY}^{(1)}$ is the sample Pearson correlation between Y and X among selected units.

The *Measure of Unadjusted Bias* (MUB) is defined as the difference between the naive sample mean and this model-based adjustment:

$$\text{MUB}(\phi) = \bar{y}^{(1)} - \hat{\mu}_Y(\phi).$$

Because MUB depends on the measurement scale of Y, hindering comparisons across outcomes, Little et al. (2020) recommend standardizing by the sample standard deviation of Y, obtaining the Standardized Measure of Unadjusted Bias (SMUB):

$$\text{SMUB}(\phi) = \frac{\text{MUB}(\phi)}{\sqrt{s_{YY}^{(1)}}} = \frac{\phi + (1-\phi)\hat{\rho}_{XY}^{(1)}}{\phi\hat{\rho}_{XY}^{(1)} + (1-\phi)} \cdot \frac{\bar{x}^{(1)} - \bar{X}}{\sqrt{s_{XX}^{(1)}}}.$$

Critically, ϕ cannot be estimated from the observed data, as there is no information about the distribution of Y among non-selected units. The strategy adopted is therefore to conduct a *sensitivity* *analysis*, computing bias estimates across a range of plausible ϕ values to assess the robustness of conclusions to departures from ignorability. Three particular values of ϕ provide intuitive benchmarks:

$$\text{SMUB}(0) = \hat{\rho}^{(1)}_{XY} \frac{\bar{x}^{(1)} - \bar{X}}{\sqrt{s^{(1)}_{XX}}}, \qquad \text{SMUB}(0.5) = \frac{\bar{x}^{(1)} - \bar{X}}{\sqrt{s^{(1)}_{XX}}}, \qquad \text{SMUB}(1) = \frac{1}{\hat{\rho}^{(1)}_{XY}} \frac{\bar{x}^{(1)} - \bar{X}}{\sqrt{s^{(1)}_{XX}}}.$$

To reflect sensitivity to the choice of ϕ, Little et al. (2020) recommend reporting the *sensitivity interval* $[\text{SMUB}(0), \text{SMUB}(1)]$ to bound the range of plausible bias values under the PPMM assumptions, with SMUB(0.5) serving as a central point estimate when no prior information about ϕ is available. If this interval excludes zero and is substantively meaningful in magnitude, it provides evidence that selection bias may threaten the validity of conclusions drawn from the non-probability sample.

To isolate the component of bias attributable specifically to departures from ignorability (i.e., $\phi > 0$), Little et al. (2020) define the *Standardized Measure of Adjusted Bias* (SMAB) as:

$$\text{SMAB}(\phi) = \text{SMUB}(\phi) - \text{SMUB}(0) = \frac{\phi\{1 - (\hat{\rho}^{(1)}_{XY})^2\}}{\phi\hat{\rho}^{(1)}_{XY} + (1 - \phi)} \cdot \frac{\bar{x}^{(1)} - \bar{X}}{\sqrt{s^{(1)}_{XX}}}.$$

While SMUB quantifies the total bias in the unadjusted sample mean $\bar{y}^{(1)}$, SMAB captures the portion of the overall bias in an unadjusted estimate that exists after adjustment for the known auxiliary variables (given a choice of ϕ), under an assumption that selection is only a function of X (or ignorable).

We note that SMUB(0), SMUB(0.5) and SMUB(1) can be computed without access to microdata for population elements excluded from the non-probability sample. A key advantage of these indices is that they require only the aggregate population mean of the proxy X, which in turn depends on the population means of the auxiliary variables Z. However, these point estimates of bias do not account for sampling variability in constructing the proxy X, that is, in estimating $\hat{\beta}$ from the regression of Y on Z, and may therefore understate total uncertainty. To address this limitation, Little et al. (2020) proposed a fully Bayesian approach that yields posterior draws of SMUB(ϕ), allowing uncertainty to be fully propagated and producing point estimates and credible intervals that can assess whether the estimated bias differs meaningfully from zero or exceeds a substantively important threshold.

Specifically, under a fully Bayesian approach, prior distributions are placed on the regression coefficients β defining the proxy X, the pattern-specific parameters in Equation (1), and the sensitivity parameter ϕ, which can either be fixed or assigned a prior distribution. A common default specification assigns relatively noninformative priors to β and the pattern–mixture parameters and a $\text{Uniform}(0, 1)$ prior to ϕ, reflecting complete ignorance about the degree of non-ignorability. Markov chain Monte Carlo methods then yield posterior draws of $\text{SMUB}(\phi)$ that fully propagate uncertainty, producing credible intervals that can be used to assess whether estimated bias is meaningfully different from zero or exceeds a substantively important threshold. This approach requires the sample mean and variance of X for the non-sampled population, which depend on the sample mean and covariance matrix of Z among non-sampled units. When only the means of Z are available, as is often the case, it can be assumed that the population covariance matrix of Z is the same for sampled and non-sampled units, allowing it to be estimated from the sampled cases.

2.2 Indices of Non-Ignorable Selection Bias for Proportions

The SMUB framework presented in Section 2.1 assumes normally distributed outcomes, limiting its direct applicability to binary variables. To address this limitation, Andridge et al. (2019) extended the proxy pattern–mixture model to binary outcomes by introducing a latent variable formulation, building on earlier developments by Andridge and Little (2020). This extension preserves the intuitive interpretation of the sensitivity parameter ϕ while accommodating the discrete nature of proportions, yielding the Measure of Unadjusted Bias for Proportions (MUBP).

Let Y be a binary variable taking values 0 or 1, representing, for instance, the presence or absence of a particular characteristic in the target population. Following standard probit model conventions, Y is assumed to arise from an underlying continuous latent variable U via the threshold mechanism

$$Y = \begin{cases} 1 & \text{if } U > 0, \\ 0 & \text{if } U \leq 0. \end{cases}$$

The latent variable formulation facilitates the specification of a tractable joint model for Y and auxiliary predictors, enabling the application of normal pattern-mixture modeling techniques analogous to those used for continuous outcomes.

As in the continuous case, let $S \in \{0, 1\}$ denote selection into the non-probability sample, with Y observed only when $S = 1$. The proxy X is constructed by regressing the binary outcome Y on the auxiliaries Z using a probit model fitted to the non-probability sample. In this case, Z must be available for all units in the non-probability sample, and either sufficient statistics (means, variances and covariances) or microdata for Z must be available for the non-selected units. A probit regression model is used for the binary indicator of interest because this model assumes that the observed indicator arises from an underlying, unobserved latent variable that follows a normal distribution.

Following the same pattern-mixture framework used for continuous outcomes, the joint distribution of the latent variable U and proxy X is assumed to follow a bivariate normal distribution conditional on selection status:

$$(U, X \mid S = j) \sim \mathcal{N}_2 \left(\begin{pmatrix} \mu_U^{(j)} \\ \mu_X^{(j)} \end{pmatrix}, \begin{pmatrix} \sigma_{uu}^{(j)} & \rho_{ux}^{(j)} \sqrt{\sigma_{uu}^{(j)} \sigma_{xx}^{(j)}} \\ \rho_{ux}^{(j)} \sqrt{\sigma_{uu}^{(j)} \sigma_{xx}^{(j)}} & \sigma_{xx}^{(j)} \end{pmatrix} \right), \quad j \in \{0, 1\}.$$

Here $\mu_U^{(j)}$ and $\mu_X^{(j)}$ denote the means of the latent variable and proxy in selection pattern j, $\sigma_{uu}^{(j)}$ and $\sigma_{xx}^{(j)}$ are their variances, and $\rho_{ux}^{(j)}$ is their correlation. As in the continuous-outcome case, some parameters governing the distribution of (U, X) among non-selected units ($j = 0$) are not identified without additional assumptions. To achieve identification, the same structural assumption is used in SMUB, namely that selection depends on (U, X) through a scalar index, $\Pr(S = 1 \mid U, X) = g\big((1-\phi)X^* + \phi U\big)$, where $X^* = X\sqrt{\frac{\sigma_{uu}^{(1)}}{\sigma_{xx}^{(1)}}}$, $g(\cdot)$ is an unspecified monotonic function and $\phi \in [0, 1]$ is the sensitivity parameter.

We note that the effectiveness of the auxiliary proxy X in predicting the binary outcome Y is quantified by the *biserial correlation*, which measures the association between a continuous variable (the proxy X) and a binary variable (Y). In the latent variable framework, this is equivalent to the Pearson correlation between U and X among selected units, $\rho_{ux}^{(1)} = \text{Corr}(U, X \mid S = 1) = \text{Biserial Corr}(Y, X \mid S = 1)$.

As is customary with latent variables, $\sigma_{uu}^{(1)} = 1$, since the mean and variance cannot be separately estimated. Under this model specification, the marginal probability that $Y = 1$ in the target population can be expressed by:

$$\mu_Y = \Pr(Y=1) = \pi\Phi(\mu_U^{(1)}) + (1-\pi)\Phi\left(\frac{\mu_U^{(0)}(\phi)}{\sigma_{uu}^{(0)}(\phi)}\right),$$

where $\pi = \Pr(S=1)$ is the proportion of selected cases in the population, $\Phi(\cdot)$ is the standard normal cumulative distribution function, and $(\mu_U^{(0)}(\phi), \sigma_{uu}^{(0)}(\phi))$ are the mean and variance of U among non-selected units, which depend on the assumed value of ϕ. These unidentified parameters for a specific choice of ϕ are given by:

$$\mu_U^{(0)}(\phi) = \mu_U^{(1)} + \frac{\phi + (1-\phi)\rho_{ux}^{(1)}}{\phi\rho_{ux}^{(1)} + (1-\phi)} \cdot \frac{\mu_X^{(0)} - \mu_X^{(1)}}{\sigma_{xx}^{(1)}}, \tag{2}$$

$$\sigma_{uu}^{(0)}(\phi) = 1 + \left[\frac{\phi + (1-\phi)\rho_{ux}^{(1)}}{\phi\rho_{ux}^{(1)} + (1-\phi)}\right]^2 \cdot \frac{\sigma_{xx}^{(0)} - \sigma_{xx}^{(1)}}{\sigma_{xx}^{(1)}}. \tag{3}$$

The difference of the proportion for selected cases from the overall proportion is therefore

$$\mu_y^{(1)} - \mu_y = \mu_y^{(1)} - \left\{\pi\Phi(\mu_u^{(1)}) + (1-\pi)\Phi\left(\frac{\mu_u^{(0)}}{\sqrt{\sigma_{uu}^{(0)}}}\right)\right\}.$$

For a given choice of ϕ, a Measure of Unadjusted Bias for Proportions, MUBP(ϕ), is then defined as the difference between the proportion observed in the non-probability sample and the estimated population proportion:

$$\begin{aligned}\mathrm{MUBP}(\phi) &= \hat{\mu}_y^{(1)} - \hat{\mu}_y \\ &= \hat{\mu}_y^{(1)} - \hat{\pi}\Phi(\hat{\mu}_u^{(1)}) - (1-\hat{\pi}) \\ &\quad \times \Phi\left(\left\{\hat{\mu}_u^{(1)} + \frac{\phi + (1-\phi)\hat{\rho}_{ux}^{(1)}}{\phi\hat{\rho}_{ux}^{(1)} + (1-\phi)} \frac{\hat{\mu}_x^{(0)} - \hat{\mu}_x^{(1)}}{\sqrt{\hat{\sigma}_{xx}^{(1)}}}\right\} \Big/ \right. \\ &\qquad \left. \sqrt{1 + \left\{\frac{\phi + (1-\phi)\hat{\rho}_{ux}^{(1)}}{\phi\hat{\rho}_{ux}^{(1)} + (1-\phi)}\right\}^2 \frac{\hat{\sigma}_{xx}^{(0)} - \hat{\sigma}_{xx}^{(1)}}{\hat{\sigma}_{xx}^{(1)}}}\right).\end{aligned}$$

where $\hat{\mu}_Y^{(1)} = \bar{y}^{(1)}$ is the sample proportion among selected units, and $\hat{\mu}_Y(\phi)$ is computed by replacing the parameters by estimates into Equations (2) and (3). Estimation of MUBP(ϕ) requires computing the sampling fraction π, which may be close to 0 for larger populations, the biserial correlation $\rho_{ux}^{(1)}$ between the latent variable U and the proxy X among selected units, and sufficient statistics for the proxy variable X for both the selected and the non-selected portions of the target population. This last requirement is stronger than that for the SMUB, which only requires the population mean of X, not its variance. Maximum likelihood (ML) estimates of these sufficient statistics for the selected cases can be computed using the observed data from the non-probability sample. Andridge et al. (2019) estimate $\rho_{ux}^{(1)}$ using the two-step approach of (Olsson, Drasgow, and Dorans, 1982), while the remaining parameters are obtained via ML. They refer to the resulting estimates as ‘modified’ ML (MML). To prevent overfitting in the construction of the proxy X and in the estimation of $\rho_{ux}^{(1)}$, they recommend multifold cross-validation.

As with SMUB, extreme and intermediate values of ϕ provide interpretable benchmarks for sensitivity analysis. When $\phi = 0$ (ignorable selection), selection depends only on the observed proxy X. When $\phi = 1$ (fully non-ignorable selection), selection depends entirely on the latent outcome U, and MUBP(1) provides an upper bound on potential bias under the model assumptions. The midpoint MUBP(0.5) represents a compromise assumption of equal dependence on the proxy and latent outcome. Andridge et al. (2019) recommend reporting the sensitivity interval $[\text{MUBP}(0), \text{MUBP}(1)]$ to bound the plausible range of bias values.

Finally, as with the SMUB, the maximum likelihood estimation of the MUBP treats the coefficients in the probit model and therefore the proxy X as fixed, potentially understating total uncertainty in the bias estimates. To address this limitation, Andridge et al. (2019) proposed a fully Bayesian implementation that propagates uncertainty through all levels of estimation. Under this approach, prior distributions are placed on the regression coefficients defining X, on the parameters of the pattern–mixture model, and on the sensitivity parameter ϕ, which can either be fixed or assigned a prior distribution such as $\text{Uniform}(0, 1)$. The Gibbs sampler alternates between imputing the latent variable U from a truncated normal distribution, updating the regression coefficients for the probit model, regenerating the proxy X, and drawing the parameters of the pattern–mixture model. These steps yield posterior draws of $\text{MUBP}(\phi)$ that fully incorporate parameter and model uncertainty.

3 Evidence from Simulations and Empirical Applications of SMUB and MUBP

The development of the SMUB and MUBP has been followed by a series of simulation studies and empirical applications designed to evaluate how well these indices perform in realistic survey conditions. These studies examined their ability to detect and quantify non-ignorable selection bias for continuous and binary outcomes, proxy strengths, and selection mechanisms, providing a clear picture of their practical strengths and limitations.

Andridge et al. (2019) conducted a simulation study comparing the MUB with the MUBP. The simulation design generated a binary outcome from a latent variable framework, allowing direct comparison of both approaches while varying the correlation between the proxy and the latent outcome and the degree of non-ignorability. Results showed that the MUBP more accurately captured bias in proportions, avoiding implausible estimates outside the [0,1] range that could arise under the linear-normal MUB formulation. Its performance was strong when the proxy was at least moderately predictive, producing well-calibrated sensitivity intervals. The ML-based intervals tend to be wider and to have higher coverage for the normal model than the MML-based intervals for the probit model. Coverage of the Bayesian intervals is higher than that of the MML-based intervals for both models.

The subsequent simulation work of Boonstra et al. (2021) offered a systematic evaluation of SMUB and related diagnostics in settings with continuous outcomes. The authors simulated finite populations where the relationship between outcome, auxiliary variables, and selection could be controlled, manipulating parameters such as the correlation between the outcome and its proxy, the overlap between outcome and selection predictors, and the strength of non-ignorability. Across these conditions, SMUB showed the strongest and most consistent correlation with the realized bias in estimated means, outperforming traditional diagnostics. The SMAB index effectively captured the portion of bias due to non-ignorability, remaining accurate when model assumptions were satisfied. However, as noted by Boonstra et al. (2021), performance declined when the auxiliary variable was only weakly correlated with the outcome, confirming that the usefulness of outcome-based diagnostics depends critically on having a sufficiently informative proxy.

Empirical applications further validated these insights. In the studies introducing these indices, Little et al. (2020) and Andridge et al. (2019) applied them to data from the National Survey of Family Growth (NSFG), treating smartphone owners as a non-probability sample. This design allowed for direct comparison between sample-based estimates and population benchmarks. In Little et al. (2020), the authors demonstrated that SMUB effectively identified survey variables most vulnerable to selection bias, performing well when the proxy–outcome correlation exceeded roughly 0.4. When this relationship was weak, they noted that any diagnostic based solely on auxiliary variables would likely be uninformative. Building on this framework, Andridge et al. (2019) applied the MUBP, showing that it produced narrower and more interpretable sensitivity intervals than its continuous counterpart (MUB) for proportions. The MUBP accurately captured the true bias for most binary outcomes when the proxy was at least moderately predictive of the latent outcome (Pearson correlation above about 0.3) and achieved improved coverage when uncertainty in the probit coefficients was incorporated through Bayesian credible intervals.

Subsequent research has demonstrated the practical value of these indices in diverse real-world settings. West and Andridge (2023) applied the MUBP to evaluate bias in pre-election polling for the 2020 U.S. presidential election. The main case study drew on the ABC/Washington Post polls conducted by Abt Associates in September and October 2020, focusing on likely voters in key states including Wisconsin, Michigan, and Pennsylvania. The estimand of interest was the proportion intending to vote for Donald Trump. Concerns about non-ignorable selection arose from the possibility that Trump supporters were systematically less likely to participate in pre-election polls. Population benchmarks were drawn from three major sources: the November 2020 CPS Voter Supplement, the 2020 ANES pre-election survey, and the AP/NORC VoteCast 2020 data. Each source offered advantages and limitations: CPS lacked direct measures of ideology and party identification, ANES had relatively small state samples, and VoteCast was not entirely probability-based. Covariates harmonized across sources included sex, age, education, race/ethnicity, political ideology, and party identification. Results showed that MUBP-adjusted estimates of Trump support were consistently higher than those produced by standard weighting alone. In many cases, the adjusted estimates narrowed the gap between poll results and the certified election outcomes. At the same time, the authors emphasized the practical challenges of implementing the MUBP, particularly the difficulties of aligning covariates across benchmark datasets.

Applications have also extended beyond political polling. Andridge (2024) investigated estimates of COVID-19 vaccine uptake from the Census Household Pulse Survey (HPS) and the Delphi-Facebook COVID-19 Trends and Impact Survey. Both surveys overestimated uptake relative to CDC benchmarks—by 14 and 17 percentage points, respectively—despite their very large sample sizes. The HPS was treated as a non-probability survey due to its extremely low response rate (6–7%). Auxiliary covariates included sex, age, education, race/ethnicity, and state, harmonized with the American Community Survey. MUBP analysis indicated that the observed overestimation was consistent with non-ignorable selection, especially if unvaccinated individuals were less likely to respond.

Validation studies have reinforced the empirical patterns and limitations observed in earlier applications. Hammon and Zinn (2024) conducted a validation study using the German General Social Survey (GGSS) as a finite population. Ten binary outcomes, including unemployment, union membership, and religious affiliation, were analyzed by comparing the full GGSS population to an artificial non-probability sample defined by internet use and political interest. They concluded that the MUBP performs well in detecting selection bias in estimated proportions when the assumptions of the underlying PPMM are satisfied. They emphasized that a moderate difference in the proxy distributions between sampled and non-sampled cases is crucial for correctly indicating the true bias. In their analysis, this condition was even more relevant than a very high correlation between X and Y, although a strong correlation is an important condition to avoid ineffective and very wide intervals of potential selection bias. The same study applied the MUBP to a large river-sampled online survey in Germany, where the authors demonstrated the practical utility of the MUBP for assessing the robustness of estimated proportions under different assumptions about the selection mechanism.

An evaluation of potential non-ignorable selection bias was conducted using the 2019 Health Survey Mailer (HSM), an off-wave supplement to the Health and Retirement Study (HRS) with an 83% response rate. Despite this high participation, eligibility restrictions raised concerns about systematic exclusion. Using demographic and health covariates common to the HSM and the HRS core, Schroeder and West (2025) estimated MUBP-adjusted proportions for ten binary health outcomes. Weighted and MUBP-adjusted estimates were generally consistent, with overlapping confidence and credible intervals for most outcomes. Larger MUBP shifts were observed only when auxiliary proxies were strong (biserial correlations above 0.5), while weaker proxies yielded wider credible intervals and smaller adjustments. Benchmark analyses treating common covariates as outcomes confirmed that both methods moved estimates toward the population truth. The study also compared results using the National Health Interview Survey (NHIS) as an alternative population source, finding lower biserial correlations and wider credible intervals when fewer common covariates were available. Beyond these empirical findings, the authors emphasized broader implications for survey researchers: MUBP can be especially valuable in panel studies that include informative covariates shared across waves, where population-level information is easier to obtain and proxy correlations tend to be higher. They also highlighted the importance of identifying strong auxiliary predictors to ensure efficiency and interpretability of bias adjustments. Overall, the results suggested that selection bias in the HSM was likely ignorable given the available covariates, and that standard weighting sufficed, while the MUBP provided reassurance and diagnostic insight into potential non-ignorable selection bias.

While the HSM study examined a traditional survey application, subsequent research has adapted these indices for use in administrative and short-term estimation contexts. Gómez-Echeverry et al. (2025) applied the MUB to flash estimates constructed from gradually filling non-probability samples, such as administrative data used for short-term economic indicators. The authors proposed three practical implementations that differ in how the sensitivity parameter ϕ is handled: MUB(C.5) fixes $\phi = 0.5$, MUB(C) uses plausible values of ϕ to approximate the range of potential bias values that are consistent with the observed data, and MUB(M) estimates ϕ using lagged and current information. Simulation results showed that MUB(M) achieved the best overall performance, particularly when selection bias was substantial, demonstrating that anchoring ϕ to historical data improves adjustment accuracy. The study also found that the correlation between the target variable and the selection mechanism was more influential than the specific distributional shape of the target variable in determining bias. A case study using turnover data from Statistics Netherlands confirmed these findings, with MUB(M) producing the lowest estimation errors across several economic sectors.

Across simulation studies and empirical applications, several consistent patterns emerge regarding the performance and practical utility of the SMUB and MUBP as diagnostics for non-ignorable selection bias. When auxiliary variables are at least moderately predictive of the outcome, typically with correlations above 0.3 to 0.4, these outcome-aware indices reliably capture both the direction and magnitude of bias, outperforming traditional representativeness diagnostics that ignore outcome distributions. Their performance declines predictably when proxies are weak, signaling insufficient auxiliary information rather than masking uncertainty. The treatment of the sensitivity parameter ϕ plays a key role: fixed values such as $\phi = 0.5$ offer simple summaries but can be less accurate than analyses spanning the full $\phi \in [0, 1]$ range, while approaches that estimate ϕ from historical or lagged data yield the most precise adjustments. Bayesian formulations, which propagate uncertainty from both proxy construction and model estimation, tend to produce better-calibrated intervals than maximum likelihood estimation. Applications across health, political, and administrative domains confirm that the SMUB and MUBP can uncover non-ignorable selection risks overlooked by conventional diagnostics, provided that proxies are strong and covariates are harmonized across data sources.

4 Discussion

The goal of this paper was to review and synthesize recent developments in diagnostic measures of selection bias for non-probability samples, focusing on the SMUB and the MUBP. Both indices provide accessible and interpretable tools for quantifying the sensitivity of survey estimates to non-ignorable selection. Their main strengths lie in their parsimony and practical feasibility. The use of a single sensitivity parameter ϕ captures the continuum between ignorable and non-ignorable selection, while the construction of a proxy variable summarizes the influence of multiple auxiliary covariates into a single dimension, simplifying implementation. Moreover, these indices can be computed even in the absence of population microdata, provided that sufficient population-level summary statistics for the auxiliary variables are available. Compared to traditional approaches, the SMUB and MUBP have been shown to detect non-ignorable bias in both simulations and empirical applications more effectively.

Building on this foundation, selection of the auxiliary variables of Z should be guided by both predictive power for the outcome of interest and availability of reliable population benchmarks. In the PPMM framework, Z is used to construct the proxy X and to supply population summaries that anchor identification, so variables that are strongly related to Y and measured consistently across surveys are preferred. These covariates must be predictive of Y but also collected in comparable form across data sources to ensure valid application of the SMUB and MUBP. Weak correlations inflate the sensitivity of results to ϕ and may produce wide, uninformative sensitivity intervals. Researchers should report the estimated correlation to communicate the strength of the proxy.

Reliable population statistics are usually obtained from large probability surveys. However, due to the increasing challenges faced by these methods, some government agencies are turning to administrative data sources to produce official statistics (Berzofsky et al., 2025), making them a promising source of auxiliary information to implement PPMM-based indices. It is worth noting that administrative data also face issues related to quality and coverage, and because of their nonprobabilistic nature, a non-ignorable missing data mechanism can cause systematic biases that standard adjustment methods may not fully correct. Nonetheless, the PPMM-based measures discussed here could also be applied to evaluate bias in administrative datasets and support their use in the production of official statistics.

In some applications, the auxiliary variables Z needed to construct the proxy may not be directly available in the non-probability sample. In such cases, these variables can be obtained by linking the sample to external sources. Little et al. (2020) suggest that, when suitable auxiliary variables are unavailable, data fusion techniques can be used to integrate variables with the required properties from another independent dataset. Linking to administrative register data can be particularly advantageous, as these sources often provide rich and reliable information. However, when such linking procedures are employed, uncertainty arising from potential mismatch errors should be properly accounted for. Recent work by Slawski et al. (2025) has developed a general framework for valid post-linkage inference in the presence of mismatch error. Incorporating these ideas into the estimation of the indices discussed in this paper represents a promising direction for future research.

Once auxiliary variables are selected and population benchmarks are identified, estimation can proceed directly. Closed-form expressions for SMUB and SMAB permit straightforward maximum likelihood (ML) estimation using sample statistics and external population summaries. Little et al. (2020) and Andridge et al. (2019) provide accompanying `R` functions `github.com/bradytwest/IndicesOfNISB`) that implement both ML and Bayesian estimation for these indices. A preliminary `R` package is also available at `github.com/randridge/ppmm`. Together, these open-source tools facilitate replication of published results, illustrate practical implementation of the indices, and, importantly, are designed to lower the barrier to their application across a wide range of research domains.

The next step involves assessing how sensitive the conclusions are to different assumptions about the selection process. Because the true selection mechanism is rarely known, sensitivity analysis provides a transparent way to evaluate robustness. A practical approach is to report sensitivity intervals, which bound the plausible range of bias under varying assumptions. The midpoint $\phi = 0.5$ offers a convenient single-number summary corresponding to equal dependence on X and Y, and has simple closed-form expressions in the continuous case. A Bayesian formulation further extends this approach by propagating uncertainty across all model components.

The PPMM framework relies on the assumption that (Y, X) follows a bivariate normal distribution. Gómez-Echeverry et al. (2025) reported that deviations of Y from normality exert a weaker influence on the performance of the MUB than changes in the strength of the non-ignorable selection mechanism or the predictive power of the proxy. Although the distributional shape plays a secondary role, it can marginally affect estimator accuracy when the selection mechanism is strongly non-ignorable or when the auxiliary variables are only moderately informative. Scenarios that combine current auxiliary variables with lagged information on the target variable appear to offer some protection against departures from normality. Overall, results obtained under normality tend to perform better than those with larger deviations. Further research examining the robustness of the SMUB and MUBP to their distributional assumptions is needed; the gamma-based extension proposed by Andridge and Thompson (2015) could provide additional insights.

Recent work has extended this framework beyond its original focus on means and proportions. West et al. (2021) extended the approach to regression coefficients in both linear and probit models, while ongoing research by Andridge and colleagues is adapting the method to ordinal and nominal outcomes through versions of the MUBP based on ordinal and multinomial probit models. Complementary methodological developments have re-expressed the PPMM as a selection model (Yiadom and Andridge, 2024) and extended the framework to subgroup estimation. Together, these efforts reinforce the conceptual foundation of the indices and expand their applicability across a broader range of survey estimation problems.

Ongoing research aims to refine the SMUB and MUBP by accounting for sampling uncertainty from finite probability survey benchmarks. This refinement is especially relevant when benchmarks are drawn from moderately sized reference samples rather than large-scale data sources, where sampling error can materially affect the accuracy of the indices. Work in progress is also focused on improving proxy construction using machine learning methods such as Bayesian additive regression trees (BART), which can capture complex non-linear relationships between auxiliary variables and outcomes, potentially yielding stronger proxies and tighter sensitivity intervals.

The original articulation of proxy pattern–mixture models emphasized that auxiliary covariates are indispensable for evaluating bias, a principle that equally applies to traditional methods for inference from non-probability samples. Valid estimation ultimately depends on the availability of high-quality auxiliary information. While the SMUB and MUBP can be computed using summary-level rather than microdata, their effectiveness still hinges on the accuracy and relevance of the covariates used.

In general, best practices for drawing valid inferences from non-probability samples (or from probability samples with low response rates) call for the identification of a large reference probability survey targeting the same population. Such a reference data source supplements the non-probability sample by providing auxiliary information on population characteristics that are essential for bias adjustment. Both data sources must include a set of common, harmonized covariates measured in the same way for individuals from the same target population. These shared covariates should be strong predictors of the key variables observed only in the non-probability sample. In practice, this requires identifying a large, representative probability survey such as the ACS, CPS, NHIS, or ANES that includes comparable measures, allocating sufficient time for harmonization when variables differ across surveys, and verifying after data collection that the chosen covariates are indeed predictive of the outcomes of interest. These steps are by no means trivial but remain indispensable, as methods for non-probability inference, whether based on microdata modeling or on population-level sufficient statistics, all rely on the consistency and predictive strength of harmonized covariates.

In this regard, we strongly support the argument made by Elliot (2022), who noted that the growing reliance on non-probability samples creates an urgent need for well-supported probability surveys to provide reliable benchmark information. Sustained investment in government-funded probability surveys is critical not only to preserve their role as independent data sources but also to strengthen their capacity to serve as analytical partners for non-probability survey inference, ensuring coverage of key covariates across the many domains where inference from non-probability samples is needed.

Acknowledgments

This work was supported in part by the National Cancer Institute of the National Institutes of Health through grant R03CA280007 (PI: Andridge).

References

Andridge, R. R. (Aug. 2024). “Using proxy pattern-mixture models to explain bias in estimates of COVID-19 vaccine uptake from two large surveys”. In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 187.3, pp. 831–843. ISSN: 0964-1998. DOI: `10.1093/jrsssa/qnae005`. URL: `https://doi.org/10.1093/jrsssa/qnae005` (visited on 09/25/2025).

Andridge, R. R. and R. J. A. Little (2011). “Proxy Pattern-Mixture Analysis for Survey Nonresponse”. In: *Journal of Official Statistics* 27, pp. 153–180.

Andridge, R. R. and R. J. A. Little (July 2020). “Proxy Pattern-Mixture Analysis for a Binary Variable Subject to Nonresponse”. en. In: *Journal of Official Statistics* 36.3, pp. 703–728. DOI: `10.2478/jos-2020-0035`. URL: `https://sciendo.com/article/10.2478/jos-2020-0035` (visited on 09/29/2025).

Andridge, R. R. and K. J. Thompson (Dec. 2015). “Assessing nonresponse bias in a business survey: Proxy pattern-mixture analysis for skewed data”. In: *The Annals of Applied Statistics* 9.4, pp. 2237–2265. ISSN: 1932-6157, 1941-7330. DOI: `10.1214/15-AOAS878`. URL: `https://projecteuclid.org/journals/annals-of-applied-statistics/volume-9/issue-4/Assessing-nonresponse-bias-in-a-business-survey--Proxy-pattern/10.1214/15-AOAS878.full` (visited on 09/25/2025).

Andridge, R. R., B. T. West, R. J. A. Little, P. S. Boonstra, and F. Alvarado-Leiton (Nov. 2019). "Indices of Non-Ignorable Selection Bias for Proportions Estimated from Non-Probability Samples". In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 68.5, pp. 1465–1483. ISSN: 0035-9254. DOI: 10.1111/rssc.12371. URL: https://doi.org/10.1111/rssc.12371 (visited on 09/25/2025).

Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau (Nov. 2013). "Summary Report of the AAPOR Task Force on Non-probability Sampling". In: *Journal of Survey Statistics and Methodology* 1.2, pp. 90–143. ISSN: 2325-0984. DOI: 10.1093/jssam/smt008. URL: https://doi.org/10.1093/jssam/smt008 (visited on 10/21/2025).

Berzofsky, M. E., D. Liao, C. Barnett-Ryan, and E. L. Smith (June 2025). "A total data quality paradigm for official statistics based on administrative data". EN. In: *Statistical Journal of the IAOS* 41.2, pp. 205–215. ISSN: 1874-7655. DOI: 10.1177/18747655251337671. URL: https://doi.org/10.1177/18747655251337671 (visited on 10/30/2025).

Boonstra, P. S., R. J. Little, B. T. West, R. R. Andridge, and F. Alvarado-Leiton (Sept. 2021). "A Simulation Study of Diagnostics for Selection Bias". EN. In: *Journal of Official Statistics* 37.3, pp. 751–769. ISSN: 0282-423X. DOI: 10.2478/jos-2021-0033. URL: https://journals.sagepub.com/action/showAbstract (visited on 09/25/2025).

Brick, J. M. and D. Williams (Jan. 2013). "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys". EN. In: *The ANNALS of the American Academy of Political and Social Science* 645.1, pp. 36–59. ISSN: 0002-7162. DOI: 10.1177/0002716212456834. URL: https://doi.org/10.1177/0002716212456834 (visited on 10/13/2025).

Cornesse, C., A. G. Blom, D. Dutwin, J. A. Krosnick, E. D. De Leeuw, S. Legleye, J. Pasek, D. Pennay, B. Phillips, J. W. Sakshaug, B. Struminskaya, and A. Wenz (Feb. 2020). "A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research". In: *Journal of Survey Statistics and Methodology* 8.1, pp. 4–36. ISSN: 2325-0984. DOI: 10.1093/jssam/smz041. URL: https://doi.org/10.1093/jssam/smz041 (visited on 10/21/2025).

Daikeler, J., M. Bošnjak, and K. Lozar Manfreda (June 2020). "Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates". In: *Journal of Survey Statistics and Methodology* 8.3, pp. 513–539. ISSN: 2325-0984. DOI: 10.1093/jssam/smz008. URL: https://doi.org/10.1093/jssam/smz008 (visited on 10/21/2025).

De Leeuw, E. d., J. Hox, and A. Luiten (Dec. 2018). "International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data". en-US. In: *Survey Methods: Insights from the Field (SMIF)*. ISSN: 2296-4754. DOI: 10.13094/SMIF-2018-00008. URL: https://surveyinsights.org/?p=10452 (visited on 10/21/2025).

Elliot, M. R. (2022). "Comments on "Statistical inference with non-probability survey samples"". en. In: *Survey Methodology, Statistics Canada* 48.2. URL: https://www150.statcan.gc.ca/n1/pub/12-001-x/2022002/article/00004-eng.htm (visited on 09/30/2025).

Elliott, M. R. and R. Valliant (May 2017). "Inference for Nonprobability Samples". In: *Statistical Science* 32.2, pp. 249–264. ISSN: 0883-4237, 2168-8745. DOI: 10.1214/16-STS598. URL: https://projecteuclid.org/journals/statistical-science/volume-32/issue-2/Inference-for-Nonprobability-Samples/10.1214/16-STS598.full (visited on 10/13/2025).

Gómez-Echeverry, S., A. Van Delden, T. De Waal, D. Pavlopoulos, and R. Stoel (Sept. 2025). "Quantifying Selection Bias in Flash Estimates Derived from Cumulative Non-Probability Samples". In: *Journal of Survey Statistics and Methodology* 00, pp. 1–26. ISSN: 2325-0992. DOI: 10.1093/jssam/smaf016. URL: https://doi.org/10.1093/jssam/smaf016 (visited on 09/25/2025).

Hammon, A. and S. Zinn (2024). "Validating an Index of Selection Bias for Proportions in Non-Probability Samples". en. In: *International Statistical Review*. ISSN: 1751-5823. DOI: 10.1111/

insr.12590. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12590 (visited on 09/25/2025).

Little, R. J. A., B. T. West, P. S. Boonstra, and J. Hu (Nov. 2020). “Measures of the Degree of Departure from Ignorable Sample Selection”. In: *Journal of Survey Statistics and Methodology* 8.5, pp. 932–964. ISSN: 2325-0984. DOI: 10.1093/jssam/smz023. URL: https://doi.org/10.1093/jssam/smz023 (visited on 09/25/2025).

Luiten, A., J. Hox, and E. de Leeuw (Sept. 2020). “Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys”. EN. In: *Journal of Official Statistics* 36.3, pp. 469–487. ISSN: 0282-423X. DOI: 10.2478/jos-2020-0025. URL: https://doi.org/10.2478/jos-2020-0025 (visited on 10/21/2025).

Lundmark, S. and K. Backström (Oct. 2025). “Predicting Survey Nonresponse with Registry Data in Sweden between 1993 and 2023: Cohort Replacement or a Deteriorating Survey Climate?” en. In: *Survey Research Methods* 19.3, pp. 247–265. ISSN: 1864-3361. DOI: 10.18148/srm/2025.v19i3.8278. URL: https://ojs.ub.uni-konstanz.de/srm/article/view/8278 (visited on 10/21/2025).

Mercer, A. W., F. Kreuter, S. Keeter, and E. A. Stuart (Apr. 2017). “Theory and Practice in Non-probability Surveys: Parallels between Causal Inference and Survey Inference”. In: *Public Opinion Quarterly* 81.S1, pp. 250–271. ISSN: 0033-362X. DOI: 10.1093/poq/nfw060. URL: https://doi.org/10.1093/poq/nfw060 (visited on 10/21/2025).

Olsson, U, F Drasgow, and N. Dorans (1982). “The polyserial correlation coefficient”. In: *Psychometrika* 47.3, pp. 337–347.

Presser, S. and S. McCulloch (July 2011). “The growth of survey research in the United States: Government-sponsored surveys, 1984–2004”. In: *Social Science Research* 40.4, pp. 1019–1024. ISSN: 0049-089X. DOI: 10.1016/j.ssresearch.2011.04.004. URL: https://www.sciencedirect.com/science/article/pii/S0049089X11000718 (visited on 10/13/2025).

Rubin, D. B. (1976). “Inference and Missing Data”. In: *Biometrika* 63.3, pp. 581–592. ISSN: 0006-3444. DOI: 10.2307/2335739. URL: https://www.jstor.org/stable/2335739 (visited on 10/10/2025).

Schroeder, H. M. and B. T. West (Feb. 2025). “Analyzing Potential Non-Ignorable Selection Bias in an Off-Wave Mail Survey Implemented in a Long-Standing Panel Study”. In: *Journal of Survey Statistics and Methodology* 13.1, pp. 100–127. ISSN: 2325-0992. DOI: 10.1093/jssam/smae039. URL: https://doi.org/10.1093/jssam/smae039 (visited on 09/25/2025).

Slawski, M., B. T. West, P. Bukke, Z. Wang, G. Diao, and E. Ben-David (July 2025). “A general framework for regression with mismatched data based on mixture modelling”. In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 188.3, pp. 896–919. ISSN: 0964-1998. DOI: 10.1093/jrsssa/qnae083. URL: https://doi.org/10.1093/jrsssa/qnae083 (visited on 11/02/2025).

West, B. T. and R. R. Andridge (Aug. 2023). “Evaluating Pre-election Polling Estimates Using a New Measure of Non-ignorable Selection Bias”. In: *Public Opinion Quarterly* 87.S1, pp. 575–601. ISSN: 0033-362X. DOI: 10.1093/poq/nfad018. URL: https://doi.org/10.1093/poq/nfad018 (visited on 09/25/2025).

West, B. T., R. J. Little, R. R. Andridge, P. S. Boonstra, E. B. Ware, A. Pandit, and F. Alvarado-Leiton (Sept. 2021). “Assessing selection bias in regression coefficients estimated from nonprobability samples with applications to genetics and demographic surveys”. In: *The Annals of Applied Statistics* 15.3, pp. 1556–1581. ISSN: 1932-6157, 1941-7330. DOI: 10.1214/21-AOAS1453. URL: https://projecteuclid.org/journals/annals-of-applied-statistics/volume-

15/issue-3/Assessing-selection-bias-in-regression-coefficients-estimated-from-nonprobability-samples/10.1214/21-AOAS1453.full (visited on 09/25/2025).

Williams, D. and J. M. Brick (June 2018). "Trends in U.S. Face-To-Face Household Survey Nonresponse and Level of Effort". In: *Journal of Survey Statistics and Methodology* 6.2, pp. 186–211. ISSN: 2325-0984. DOI: 10.1093/jssam/smx019. URL: https://doi.org/10.1093/jssam/smx019 (visited on 10/13/2025).

Yiadom, S. A. and R. Andridge (Sept. 2024). *Formulating the Proxy Pattern-Mixture Model as a Selection Model to Assist with Sensitivity Analysis*. arXiv:2409.18117. DOI: 10.48550/arXiv.2409.18117. URL: http://arxiv.org/abs/2409.18117 (visited on 09/25/2025).