Hybrid Methods for Combining Probability and Nonprobability Samples for Estimation

Michael Yang, Jay Breidt, Nada Ganesh, Evan Herring-Nathan, Edward Mulrow, and Vicki Pineau
NORC at the University of Chicago, 4350 East-West Highway, 8th Floor, Bethesda, MD 20814

Abstract

The authors have evaluated a range of estimation methods for combining probability and nonprobability samples using Monte Carlo simulations as well as case studies. Our earlier evaluations show that Statistical Matching, Propensity Weighting, and Small Area Modeling all perform well. In particular, hybrid methods that combine multiple methods perform better than any single method. Based on further Monte Carlo simulations with the CCTC simulation data, now made available by NORC at the University of Chicago, this paper extends our earlier investigations to explore two promising hybrid methods: (1) Matching Propensity (MP) and (2) Matching Propensity with Small Area Modeling (MP-SM). The MP method features a two-way combination of Statistical Matching and Propensity Weighting, and the MP-SM method adds Small Area Modeling to the MP method, resulting in a three-way combination that is doubly robust. The simulations show that both hybrid methods result in substantial bias reduction. The MP-SM method leads to slightly more bias reduction, both overall and for small domains, but at the cost of much greater variance, which is consistent with our earlier observation that calibration to small area model estimates is effective in reducing bias, but the resulting weights are more variable. Overall, the MP method performs better based on our current simulation results. For variance estimation under the MP method, we propose a simple variance estimator that produces approximately unbiased variance estimates in our simulation experiments.

Key words: Nonprobability sample estimation, statistical matching, propensity weighting, small area estimation, variance estimation

1. Introduction

While probability sampling remains the standard basis for inference from a sample to a population, there has been growing interest in methods that combine probability and nonprobability samples to improve the cost efficiency of survey estimation (Elliot, 2009; Wang, et al., 2015; Elliot and Valliant, 2017; and Valliant, 2020). Nonprobability samples provide a lower cost alternative to probability samples; however, estimates based on nonprobability samples may be biased due to unknown population coverage and sample selection mechanisms. A compromise is to combine probability and nonprobability samples to take advantage of the representativeness of the probability sample and the cost effectiveness of the nonprobability sample. At NORC, our typical application is to combine a probability sample from our AmeriSpeak Panel®[1] and an opt-in sample from a sample vendor. The objective of our research is to develop a toolkit of estimation methods to support our practice.

The authors have evaluated a range of estimation methods for combining probability and nonprobability samples using Monte Carlo simulations as well as case studies. Our earlier evaluations show that Statistical Matching, Propensity Weighting, and Small Area Modeling perform well in terms of bias reduction, mean

---

[1] https://amerispeak.norc.org/

square error, and confidence interval coverage. In particular, hybrid methods that combine multiple methods perform better than any single method. Based on further Monte Carlo simulations with the CCTC simulation data, now made available at the Open Science Framework (https://osf.io/ygpzm/), this paper extends our earlier investigations to explore two promising hybrid methods: (1) Matching Propensity (MP) and (2) Matching Propensity with Small Area Modeling (MP-SM). The MP method features the combination of Statistical Matching and Propensity Weighting, while the MP-SM method adds Small Area Modeling to the MP method, resulting in a three-way combination that is also doubly robust. Since the outcome of both methods is a set of analysis weights for the combined probability and nonprobability samples, the two methods are both weighting methods. The simulations show that both hybrid methods provide substantial bias reduction. The MP-SM method leads to slightly more bias reduction, both overall and for small domains, but at the cost of much greater variance, which is consistent with our earlier observation that calibration to small area model estimates is most effective in reducing bias, but the resulting weights are more variable. Overall, the MP method performs better based on our current simulation results. For variance estimation under the MP method, we propose a simple variance estimator that produces approximately unbiased variance estimates in our simulation experiments.

2. Methods Evaluated

In this section, we briefly review our past evaluations, but focus on the two hybrid methods evaluated in the current simulations.

2.1 Past Evaluations

In our past evaluations, we reported comparative analysis of the following estimation methods for combining probability and nonprobability samples:

- Calibration: Calibrate the nonprobability sample and the combined sample weights to known benchmarks, typically census demographic benchmarks (Deville and Särndal, 1992).
- Statistical Matching: Statistically match nonprobability and probability samples and derive pseudo weights for nonprobability sample units based on the matching pattern (D'Orazio, and Scanu, 2006; Rivers, 2007; Mulrow, et al., 2018).
- Superpopulation Modeling: Under a linear superpopulation model, express the combined sample weights as a linear combination of a set of covariates with known population totals.
- Propensity Weighting: Use the probability sample as the reference sample to model the propensity of inclusion in a nonprobability sample and then use the estimated propensities to derive pseudo weights for the nonprobability sample units.
- Small Area Modeling: a doubly robust calibration method that rakes an initial set of combined sample weights to both standard demographic benchmarks and additional raking targets for some key response variables and domains predicted from small area models (Kim and Haziza, 2014; Ganesh et al., 2017; Chen, Li, and Wu, 2019).

Most of these methods are also discussed in Elliot, 2009, Elliot and Valliant, 2017, and Valliant, 2020. We evaluated the properties of each method in terms of bias reduction, mean square error, and confidence interval coverage using case studies and Monte Carlo simulations. These evaluations show that both Calibration and Superpopulation Modeling lead to similar results, most likely because they tend to rely on the same set of covariates that have available population benchmarks from official statistics. Under our formulations of these two methods, they are deemed less effective, especially in terms of bias reduction. On the other hand, Statistical Matching, Propensity Weighting, and Small Area Modeling consistently perform better, probably because they can incorporate more covariates in their models that are used to estimate the weights for the nonprobability sample units. Finally, hybrid methods, where two or more methods are combined to estimate the weights, tend to perform better than any single method (Ganesh et al., 2017; Yang, et al, 2018; Yang et al., 2019; Mulrow, et al., 2020; Yang et al., 2020; Yang, et al., 2021).

2.1 Current Evaluations

In this paper, we evaluate two hybrid methods that feature some combination of Statistical Matching, Propensity Weighting, and Small Area Modeling. We describe the two weighting methods in this section.

2.1.1 Matched Propensity (MP)

We start by matching each nonprobability sample unit to a probability sample unit using the R *StatMatch* package (D'Orazio, 2017). Statistical matching is conducted using a nearest neighbor hot deck algorithm based on Gower's dissimilarity measure, which can accommodate both categorical (nominal and ordinal) and continuous variables in the dissimilarity calculation. The matching variables are determined using a gradient boosting tree model (D'Orazio, Di, and Scanu, 2006). The matching process resembles imputation in the sense that a donor from the probability sample is matched to a recipient from the nonprobability sample based on a set of matching variables (Bethlehem, 2015).

Statistical matching divides the probability sample into two disjoint sets: the set of the probability sample units that is matched to the nonprobability sample and the set not matched. Sampling weights for the probability sample are developed following standard weighting procedures. The weights for the nonprobability samples are developed by fitting a propensity model to estimate the inclusion probabilities for the nonprobability sample units, using the matched probability sample as the reference sample. These estimated probabilities are then used to derive the pseudo weights for the nonprobability sample units. The nonprobability sample weights are developed in the following steps:

- Concatenate the matched probability sample and the nonprobability sample
- Create a dichotomous indicator variable, Y, 1 for nonprobability sample units and 0 for matched probability sample units
- Fit a logistic regression model to predict the probability of inclusion for the nonprobability sample units
- Estimate the weights for nonprobability sample units as the reciprocal of the predicted probabilities
- Rake the combined sample weights to population benchmarks

Note that using propensity models to develop nonprobability sample weights has been quite common in practice. The difference here is that our propensity model uses only the matched probability sample instead of the entire probability sample as the reference sample. To the extent that the matching pattern reveals the degree of undercoverage of the nonprobability sample, this procedure should produce better pseudo weights for the nonprobability sample.

2.1.2 Matched Propensity with Small Area Modeling (MP-SM)

The MP-SM method features a three-way combination of Statistical Matching, Propensity Weighting, and Small Area Modeling. In practice, many studies require separate analysis and reporting by domains defined by demographic characteristics such as age, gender, race/ethnicity, education, and the like. The purpose of integrating Small Area Modeling to the MP method is to improve estimation for such domains.

Consider the MP weights a set of initial weights for the combined sample. We derive the weighted estimates for a set of domains for one or more key response variables, separately for the probability sample and the nonprobability sample. For each response variable y, there are two weighted estimates for each domain d, one from the probability sample, y_d^P, and one from the nonprobability sample, y_d^{NP}.

In essence, the MP-SM method is a hybrid calibration method, where small area models are used to generate additional (i.e., beyond regular demographic benchmarks) raking targets for a set of response variables and domains. The combined sample weights under MP-SM are developed via the following steps:

First, for each selected key variable y, fit a Bivariate Fay-Herriot model (Fay and Herriot, 1979; Rao, 2003; Ganesh et al., 2017) to jointly model the weighted domain-level point estimates:

$$y_d^P = x'_d\beta + \nu_d + \varepsilon_d^P$$

$$y_d^{NP} = b + \alpha_d^{NP} + x'_d\beta + \nu_d + \varepsilon_d^{NP}$$

where

d denotes a domain (e.g., 18-34 years of age, male, Hispanic);
$\boldsymbol{x}_d$ is a vector of covariates associated with domain d;
v_d's are domain level random effects;
b is a fixed effect bias term associated with the nonprobability sample estimate;
α_d^{NP}'s are random effect bias terms associated with the nonprobability sample estimate;
ε_d^P, ε_d^{NP} are the sampling errors associated with y_d^P and y_d^{NP}, respectively

When modeling proportions, the point estimates are typically transformed using an arcsine transformation (see Jiang et al., 2001). The arcsine transformation preserves the bounds of 0 and 1 for a proportion. The transformed point estimate for a proportion z is calculated as two times the arcsine of the square root of the proportion,

$$y = 2\sin^{-1}\sqrt{\mathrm{z}}.$$

Second, based on the fitted small area models, produce predictions for the set of response variables and domains using an Empirical Best Linear Unbiased Predictor (EBLUP). The predicted proportions are converted into estimated population totals for each response category.

Third, rake the initial MP weights such that the distribution of the combined sample weights matches both regular demographic benchmarks and the population totals predicted from the small area models for each response variable and for each domain.

3. Monte Carlo Simulation Setup

The simulation samples are generated using survey completes from a large-scale national study, *Culture & Community in a Time of Crisis (CCTC): A Special Edition of Culture Track*, a national survey of the general public and culturally active Americans (Benoit-Bryan and Mulrow, 2021). The full population (*Frame 1*) for the simulation data consists of 113,549 U.S. adult survey completes. From this frame, 1,000 stratified random samples of size 1,000 are selected. These are considered probability samples and are weighted as such.

To create the nonprobability sample frame (*Frame 2*), a set of population units from *Frame 1* are removed based on some descriptive variables in the file. Additional units are removed based on a highly skewed binomial distribution, resulting in a nonprobability sample frame with 74,202 units. From *Frame 2*, 1,000 samples of size 4,000 are selected. These are considered nonprobability samples whose weights need to be developed and then combined with the probability sample weights.

Both the probability and nonprobability samples contain a large number of demographic variables and substantive survey response variables. As in many studies, CCTC response variables of interest are categorical by nature. Therefore, statistics related to these variables are proportions for the categories of each variable. For each response variable of interest, let $P_{Frame\ 1}$ and $P_{Frame\ 2}$ be the population proportion computed from the probability and nonprobability frame, respectively. The *known absolute bias* associated with the nonprobability frame, B_{pop}, is calculated as the difference of population proportions between the two frames:

$$B_{pop} = |P_{Frame\ 1} - P_{Frame\ 2}|.$$

Twelve behavioral and two attitudinal variables are chosen as the evaluation variables in this investigation, representing a range of absolute bias from close to 0 to over 5 percentage points. The behavioral variables are all binary, so a single proportion is estimated for each variable. Both attitudinal variables are categorical, each with five response categories, hence five estimated proportions. Therefore, across the fourteen evaluation variables, there are a total of 22 estimated proportions. Figure 1 shows the absolute known bias of the behavioral variables, and Figure 2 shows the absolute known bias of the attitudinal variables.

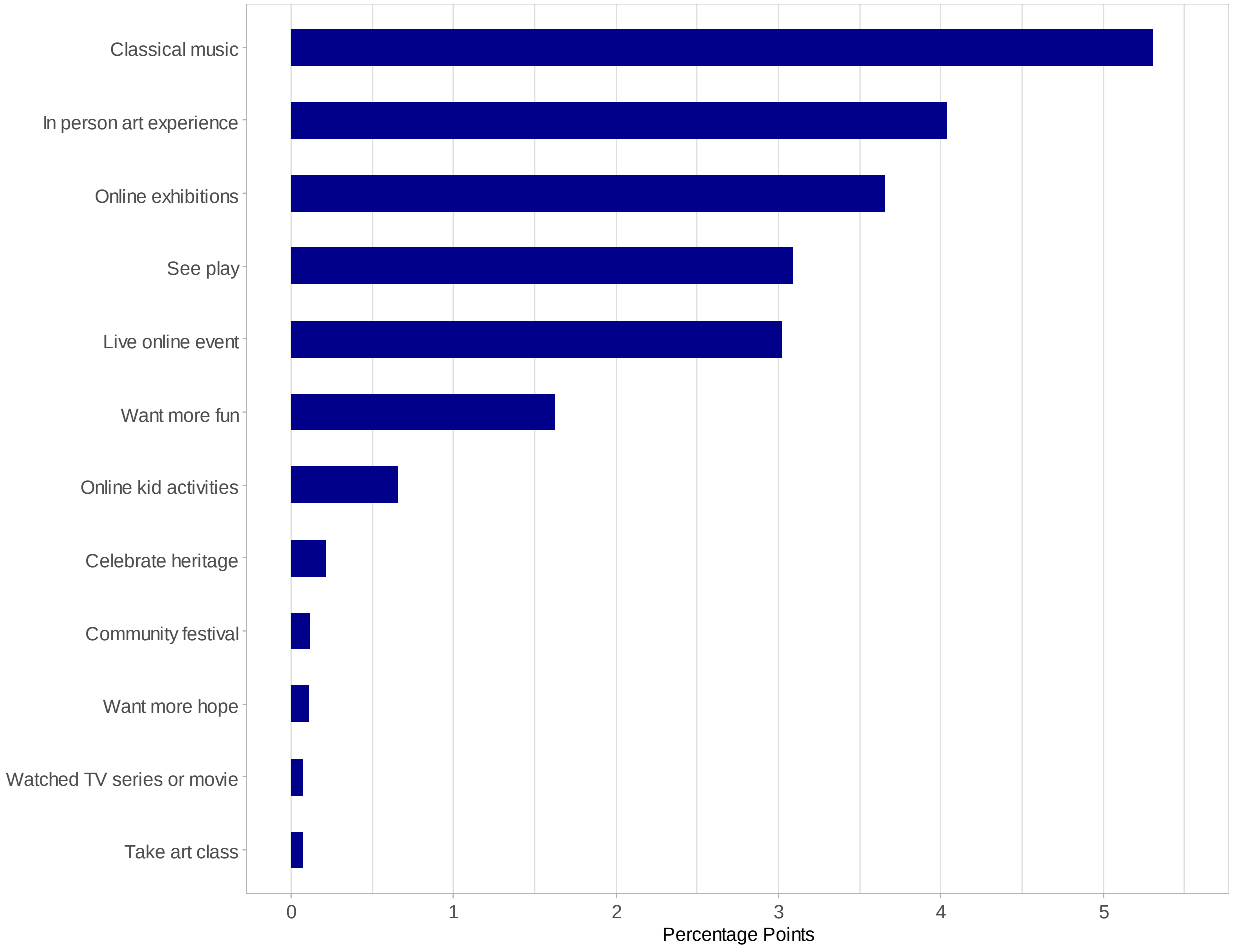

Figure 1: Absolute Known Bias of Behavioral Variables

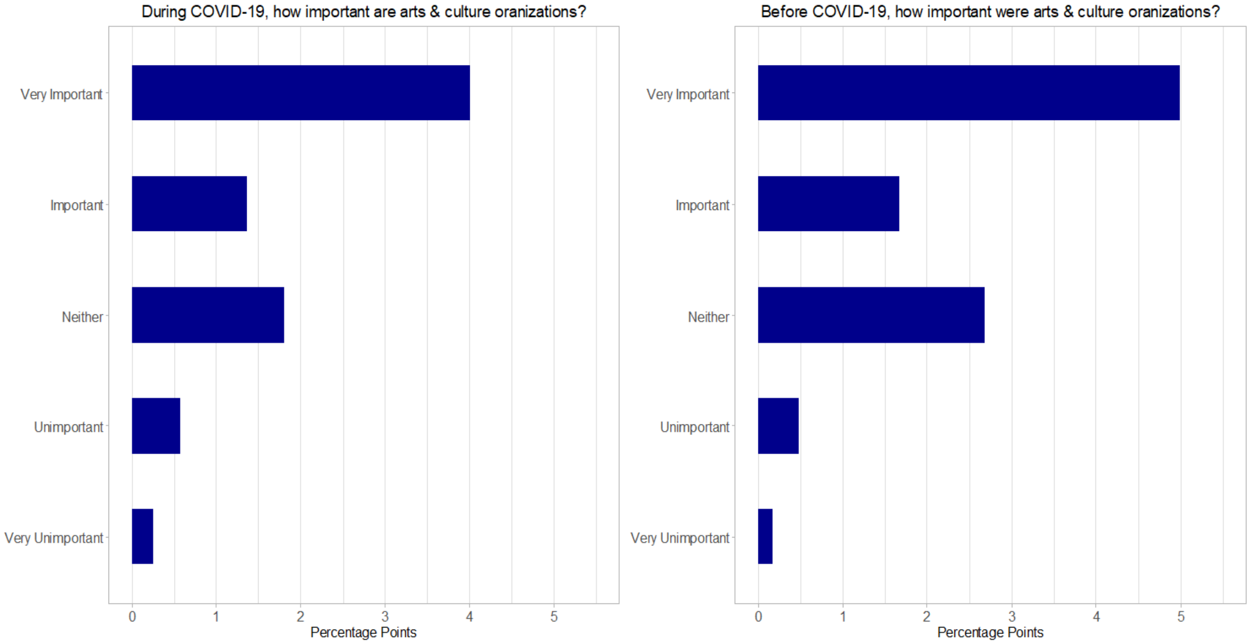

Figure 2: Absolute Known Bias of Attitudinal Variables

For more information about the construction of the simulation data, please see Benoit-Bryan and Mulrow (2021).

4. Simulation Results

Each of the two estimation methods is used to develop a set of combined weights for the probability and nonprobability samples, and these weights are used to derive combined sample estimates for the evaluation variables. In this section, we evaluate the properties of these estimates, focusing on estimation bias and variance evaluations.

4.1 Bias Evaluation

For each estimate associated with an evaluation variable, let $\hat{p}_m$ denote the estimated proportion for iteration $m = 1, \dots, 1{,}000$, $\bar{\hat{p}}_m$ denote the average of $\hat{p}_m$ across all iterations, and $P = P_{Frame\ 1}$ denote the known population proportion.

The simulated absolute relative bias (ARB) for each estimate is computed as

$$ARB(\hat{p}) = \left|\frac{\bar{\hat{p}}_m - P}{P}\right|.$$

We first compare the ARB under each estimation method for the 22 estimated proportions associated with the 12 behavioral variables and the two attitudinal variables.

In Figure 3, each comparison is represented by the distance between two dots: the orange dot denotes ARB under the MP method, the green dot denotes the ARB under the MP-SM method, and the horizontal axis is in terms of percentage points. With one exception in the middle of the graph (*q1_6*), which is virtually a tie, ARB is consistently smaller under MP-SM than under MP, indicating that MP-SM performs better in terms of bias reduction. On the other hand, note that ARB under MP is also small. Under both methods, ARB is smaller than 1.5 percentage points for all behavioral variables, indicating that both methods are effective in reducing bias. ARB is a little larger for a few estimates, notably *q17_1*, *q18_1*, *q18_2*, and *q18_3*. These estimates are associated with attitudinal variable response categories that were rarely chosen by the respondents, so the denominator *P* in the relative bias formula tends to be very small, leading to large relative bias even when the absolute bias is negligibly small.

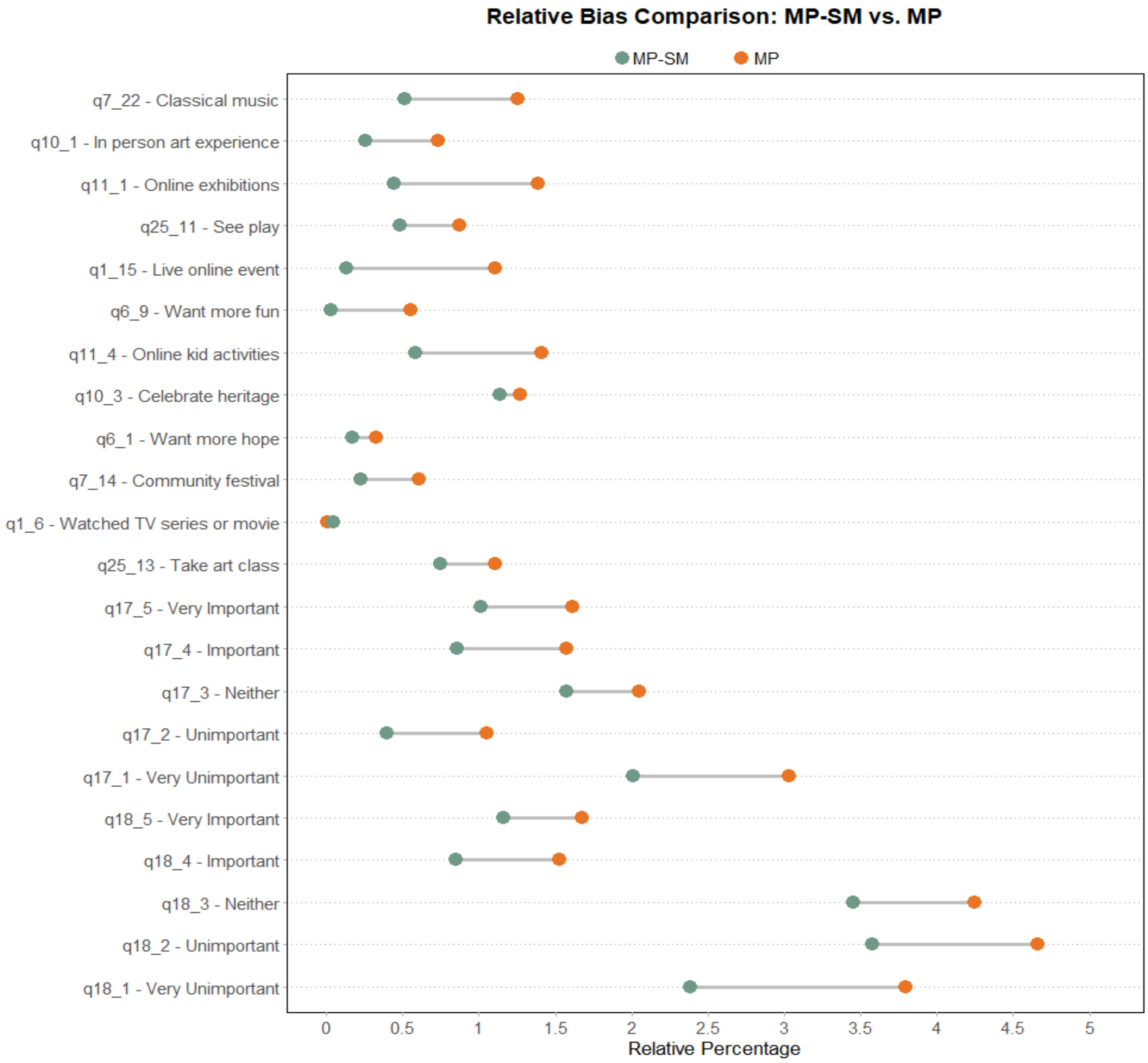

Figure 3: Absolute Relative Bias Comparison across Behavioral and Attitudinal Variables

Figure 4 shows the ARB comparisons for the same set of estimates within six domains defined by race and ethnicity. The domain estimates have less ARB under MP-SM than under MP, but, again, both methods seem to perform well in bias reduction because relative bias tends to be small for all estimates under both methods. The same pattern is observed consistently across other 1-way and 2-way domains that are not shown here. Across the hundreds of domains we examined, the MP-SM method led to smaller relative bias about 85% of the time. Based on these results, our bias evaluation reveals a clear and consistent pattern: MP-SM does a better job than MP in bias reduction, but the difference between the methods is very small.

Given that, we chose not to conduct any significance test of the differences in ARB as they are considered practically unimportant even though some of these may be statistically significant.

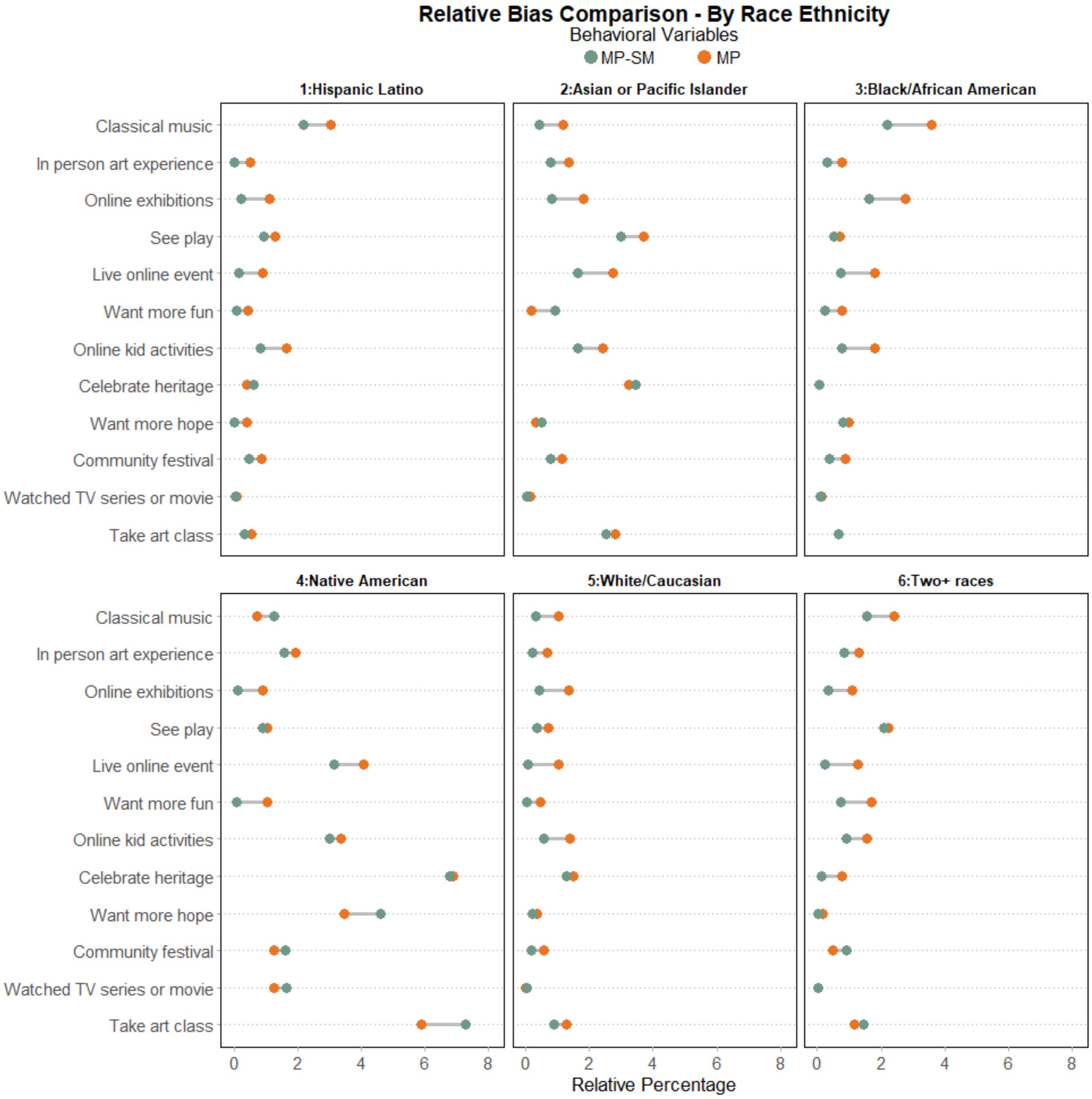

Figure 4: Absolute Relative Bias Comparison across Behavioral Variables

4.2 Variance Evaluation

For both the MP and MP-SM estimators, we computed their Monte Carlo simulation variance, which we consider the true variance of these estimators. The variance of the hybrid estimators based on Monte Carlo simulations, taken across 1,000 iterations, is,

$$v_{MC}(\hat{p}) = \frac{1}{1000-1}\sum(\hat{p}_m - \bar{\hat{p}})^2.$$

We first compare the true variance between the two estimators by taking their ratio, defined as the MP-SM variance divided by the MP variance, for each estimate.

As shown in Figure 5, the true variance ratios range from about 1.64 to 2.54 across the 22 estimates. The average of these ratios, represented by the dashed vertical line in the figure, is greater than 2, indicating that the true variance under MP-SM is much greater than under MP. Given that the MP-SM estimator has much larger variance and therefore is much less efficient, we concentrate on variance estimation for the MP method only in this investigation. As indicated in the concluding section, we intend to investigate variance estimation for the MP-SM estimator after that estimator is improved through future research.

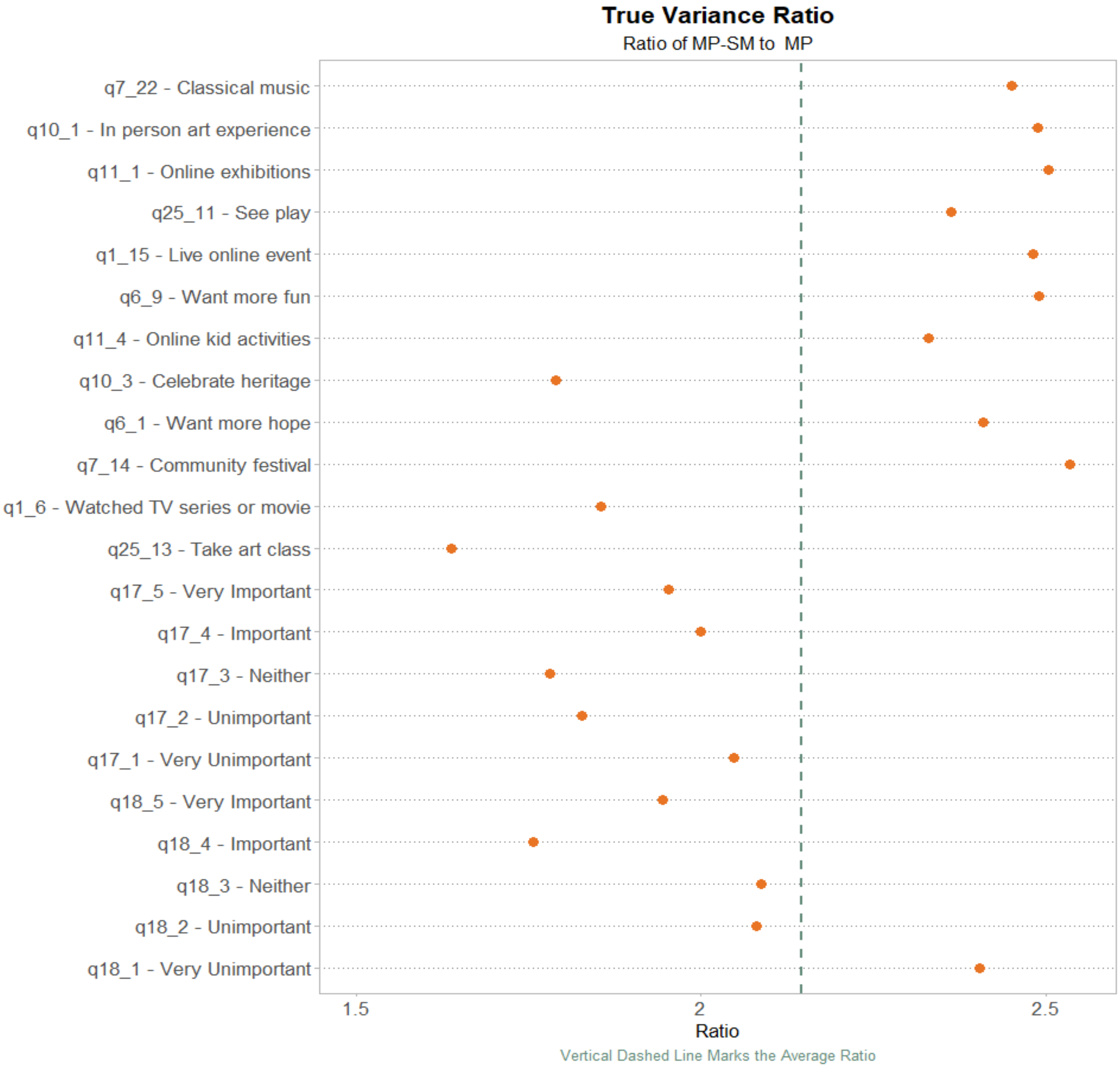

Figure 5: True Variance Ratio between MP-SM and MP Methods

For the MP method, we investigated the Taylor Series variance estimator as applied to a ratio estimate, with some custom modifications, as described below.

Let A denote the probability sample, with sample weights w_k^A known for $k \in A$. Let B denote the nonprobability sample, with sample weights w_k^B estimated from some propensity model. Let $M_k = 1$ if sample unit $k \in A$ is matched to some sample unit in B, and $M_k = 0$ otherwise. Let λ denote some combination factor, $0 < \lambda < 1$. Then, for a characteristic y, the MP estimator of the population total may be written as

$$\hat{T}_y = \sum_{k \in A} y_k (1 - M_k) w_k^A + \lambda \sum_{k \in A} y_k M_k w_k^A + (1 - \lambda) \sum_{k \in B} y_k w_k^B.$$

The MP estimator of the population total has three components: (1) the total estimated from the unmatched portion of the probability sample, (2) the total estimated from the matched probability sample, and (3) the total estimated from the nonprobability sample. The last two components are combined through a composite estimator where $\lambda = 0.5$ for estimates reported in this study.

The MP estimator for the population proportion is

$$\hat{p} = \frac{\hat{T}_y}{\sum_{k \in A} \{(1 - M_k) + \lambda M_k\} w_k^A + \sum_{k \in B} (1 - \lambda) w_k^B} = \frac{\hat{T}_y}{\hat{T}_1},$$

a ratio estimator with denominator equal to the sum of the weights across the three sample components defined by the matching results (where $\hat{T}_1$ is obtained by replacing y_k by 1 in $\hat{T}_y$, as the notation suggests).

To approximate the variance of $\hat{T}_y$, we assume that the A probability sample is independent of the B nonprobability sample. Ignoring variation due to statistical matching and due to estimating the parameters of the propensity model, the variance of $\hat{T}_y$ is approximately

$$\mathrm{Var}\left(\sum_{k \in A} y_k \{(1 - M_k) + \lambda M_k\} w_k^A \right) + \mathrm{Var}\left(\sum_{k \in B} y_k (1 - \lambda) w_k^B \right),$$

which is the variance of a stratified sample with weights w_k^A and response $y_k\{(1 - M_k) + \lambda M_k\}$ in the first stratum, A, and weights w_k^B and response $(1 - \lambda) y_k$ in the second stratum, B. The same variance approximation holds for $\hat{T}_1$, replacing y_k throughout by 1.

Accordingly, approximate variance estimators for the estimated totals $\hat{T}_y$ and $\hat{T}_1$and for their ratio $\hat{p} = \hat{T}_y / \hat{T}_1$ can be obtained by combining the two data sets, A and B, constructing custom response variables as indicated above, and treating the combined data as a stratified sample with two artificial strata, A and B, according to the sample that generated the observation. Both A and B may contain substrata based on the original sample design. In stratum A, the weights are the probability sample weights, w_k^A, and in stratum B, the weights are the estimated nonprobability sample weights, w_k^B. With this custom data set and custom response variables, the combined probability/nonprobability point estimate and its estimated variance can be computed using standard survey software as a survey-weighted ratio and its Taylor Series variance estimate.

Figure 6 shows the ratio of the Taylor Series variance estimate to the true variance for all 22 estimates.

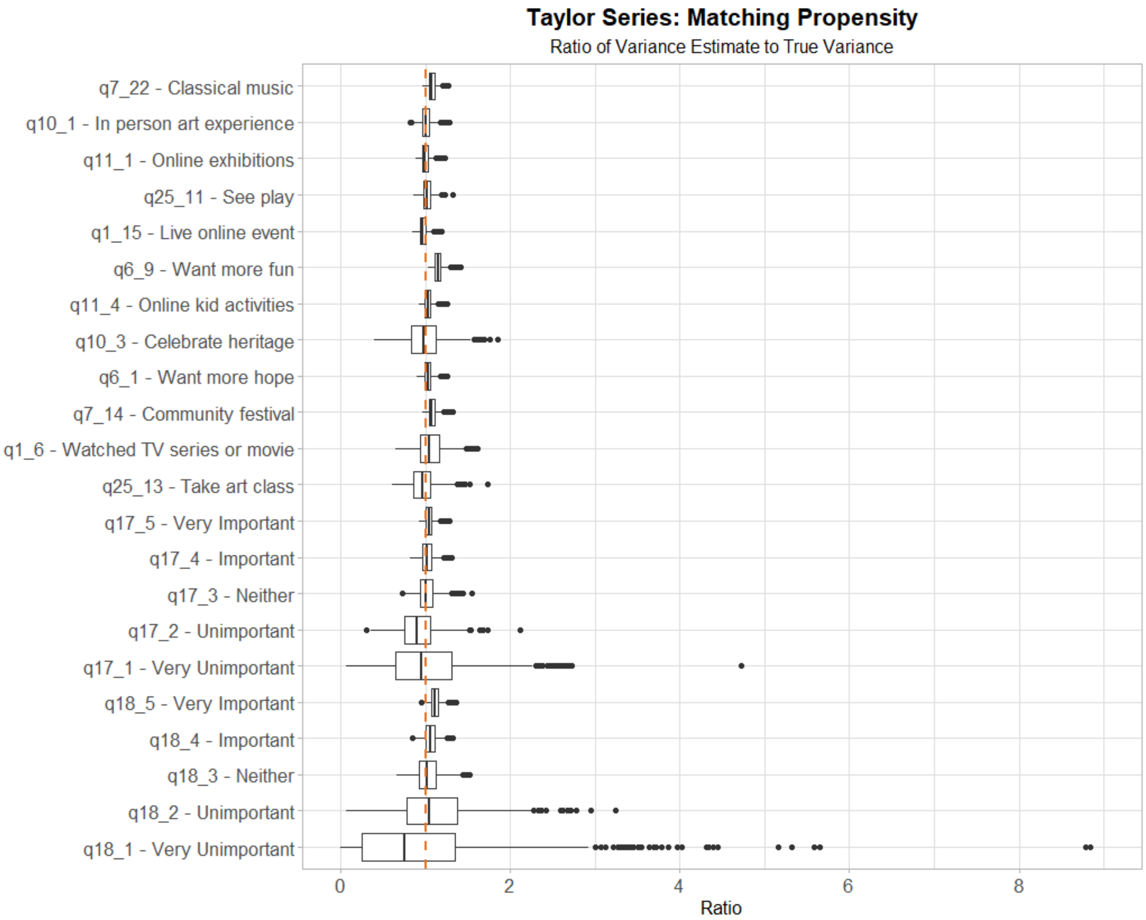

Figure 6: Ratio of Taylor Series Variance Estimate to True Variance under MP

The boxplots show the distribution of the ratio of the Taylor Series variance to the true variance for each estimate, with the dashed red line corresponding to the reference value of 1. The Taylor Series variance estimator performs well for all evaluation variables. With rare exceptions, the median of the Taylor Series variance is close to the true variance. The estimated variances for the attitudinal variables, which are the bottom 10 boxplots in Figure 6, are less stable but still quite close to the true variance on average. Overall, the Taylor Series variance estimator is approximately unbiased for most evaluation variables studied.

5. Summary and future research

Using Monte Carlo simulations, we evaluated two hybrid estimation or weighting methods for combining probability and nonprobability samples: Matching Propensity (MP) and Matching Propensity with Small Area Modeling (MP-SM). The evaluations show that, while MP-SM leads to slightly smaller estimation bias, both methods are effective in reducing bias associated with the nonprobability sample. Relative bias for the evaluation variables is so small that both estimators may be considered approximately unbiased for practical purposes.

The MP-SM method consistently leads to slightly more bias reduction relative to MP, but that greater bias reduction is achieved at the cost of much greater variance. The simulations show that the true variance under MP-SM, based on 1,000 iterations, is much greater than under MP, which is consistent with observations from our earlier investigations that calibration to small area estimates is most effective in reducing bias, but the resulting weights tend to be much more variable. Overall, from a total error perspective, the MP method performs better than the MP-SM method in the context of this simulation study. Therefore, no variance estimator is considered for MP-SM in the current investigation.

For variance estimation under MP, we investigated the Taylor Series method as applied to a stratified ratio estimator constructed by redefining the original response variable. In our simulation, the Taylor Series method produces approximately unbiased estimates of the variance for the 14 evaluation variables. This result is practically important because the Taylor Series variance estimator has the advantage of simplicity and can be implemented with standard survey software packages.

Based on the results reported here, we would like to pursue two major future research topics, both related to the MP-SM method. Small Area Modeling is introduced primarily to enable more reliable small domain estimation that is important for many studies. Under its current formulation, the MP-SM estimator has large variances. However, some modified version of MP-SM may still be a valuable option. The current small area models include random and fixed bias terms. If statistical matching and propensity weighting already have the ability to remove much of the bias, as shown in the current simulation, these bias terms may be dropped from the model, which might help to reduce the MP-SM variance. Therefore, we would like to consider small area models without the bias terms in future studies.

The Taylor Series variance estimator performs well under the MP method. If a modified MP-SM method proves to be a viable option, a different variance estimation method may be needed. Under the MP-SM method, calibration to estimated control totals may introduce additional non-negligible sources of error that warrant further investigation. This might be relevant, for example, if the nonprobability sample weights are calibrated only to the part of the population that the nonprobability sample is thought to represent.

References

J. Benoit-Bryan and Edward Mulrow (2021) “Exploring Nonprobability Methods with Simulations from a Common Data Source: Culture and Community in a Time of Crisis.” *Joint Statistical Meetings 2021 Proceedings.*

Bethlehem J. (2015) “Solving the nonresponse problem with sample matching?” *Social Science Computer Review*, Vol. 34, Issue 1, pp. 59 – 77.

Chen, Yilin, Pengfei Li, Changbao Wu. (2019). “Doubly Robust Inference with Non-probability Survey Samples,” Paper presented at the Joint Statistical Meetings, Denver, Colorado.

Deville, Jean-Claude and Carl-Erik Särndal. (1992). “Calibration Estimators in Survey Sampling,” *Journal of the American Statistical Association*, Vol. 87, No. 418 (Jun., 1992), pp. 376-382.

D'Orazio, M. (2017). *StatMatch: Statistical Matching*. R package version 1.2.5. https://CRAN.R-project.org/package=StatMatch.

D'Orazio, M., Di Zio M., Scanu M. (2006) *Statistical matching: Theory and practice*. Wiley, Chichester.

Elliott, M.R. 2009. Combining data from probability and non-probability samples using pseudo-weights. *Survey Practice*, v. 2 (6).

Elliot, M. R., Valliant, R. (2017). “Inference for Nonprobability Samples,” *Statistical Science* 2017, Vol. 32, No. 2, 249–264.

Fay, R.E., and Herriot, R.A. 1979. "Estimates of income for small places: An application of James-Stein procedures to Census data," *Journal of the American Statistical Association*, v. 74 (366), pp. 269-277.

Ganesh, N., Pineau, V., Chakraborty, A., Dennis, J.M., (2017). "Combining Probability and Non-Probability Samples Using Small Area Estimation." *Joint Statistical Meetings 2017 Proceedings.*

Jiang, J., Lahiri, P., Wan S. and Wu, C. 2001. Jackknifing in the Fay-Herriot model with an example. Proceedings of the Seminar on Funding Opportunity in Survey Research, pp. 98-104.

Kim, J.K., and Haziza, D. (2014), "Doubly Robust Inference with Missing Data in Survey Sampling," *Statistica Sinica*, 24, 375-394.

Mulrow, Edward, Nada Ganesh, Vicki Pineau, and Michael Yang. (2020). "Using Statistical Matching to Account for Coverage Bias When Combining Probability and Nonprobability Samples," *Joint Statistical Meetings 2018 Proceedings.*

Mulrow, Edward. (2022), "Exploring Methods for Combining Probability and Nonprobability Samples with Simulations from a Common Data Source." OSF. https://osf.io/ygpzm/.

Rao, J.N.K. 2003. *Small Area Estimation*. Hoboken, NJ: John Wiley & Sons.

Rivers, D. (2007), "Sampling for Web Surveys", Proceedings of the Survey Research Methods Section, Joint Statistical Meetings, American Statistical Association, Alexandria, VA, 1-26.

Valliant, R. (2020). "Comparing Alternatives for Estimation from Nonprobability Samples," *Journal for Survey Statistics and Methodology*, vol. 8, 231-263.

Wang, W., Rothschild, D., Goel, S., and Gelman A. 2015. Forecasting elections with non-representative polls. *International Journal of Forecasting*, v. 31 (3), pp. 980-991.

Yang, Y. Michael, Nada Ganesh, Ed Mulrow, and Vicki Pineau. (2018). "Estimation Methods for Nonprobability Samples with a Companion Probability Sample," *Joint Statistical Meetings 2018 Proceedings.*

Yang, Y. Michael, Nada Ganesh, Ed Mulrow, and Vicki Pineau. (2019). "Evaluating Estimation Methods for Combining Probability and Nonprobability Samples through a Simulation Study," *Joint Statistical Meetings 2019 Proceedings.*

Yang, Y. Michael, Nada Ganesh, Ed Mulrow, and Vicki Pineau. (2020). "Variance Estimation for Combined Probability and Nonprobability Samples," *Joint Statistical Meetings 2020 Proceedings.*

Yang, Y. Michael, Nada Ganesh, Ed Mulrow, Evan-Herring Nathan, and Vicki Pineau. (2021). "Estimators with Combined Probability and Nonprobability Samples Using Small Area Models," *Joint Statistical Meetings 2021 Proceedings.*