Encipher Nonprob:

Improving Weighting for Nonprobability-only Samples

In a previous paper, we introduced the SSRS Encipher Hybrid solution. Encipher Hybrid is designed for true hybrid samples—those that blend respondents from parallel probability and nonprobability samples. As discussed in that paper, hybrid designs help strike a balance between the greater accuracy of probability samples and the lower cost of nonprobability samples.

This paper now introduces Encipher Nonprob, a calibration solution for nonprobability-only samples. It includes the results of a validation study demonstrating that, though hybrid designs remain preferable when feasible, **Encipher Nonprob reduces the median selection bias by over 30% relative to standard demographic-only weighting of nonprobability-only samples**.

Hybrid vs. Nonprobability-only Samples

To understand the differences and tradeoffs between hybrid and nonprobability-only methodologies, it is useful to understand the role played by the probability sample in our Encipher Hybrid solution.

The distinguishing feature, and key advantage, of a hybrid design is that the entire study questionnaire is administered to a parallel probability sample in addition to the nonprobability sample. The SSRS Encipher Hybrid solution relies on this parallel probability sample to obtain fully probability-based estimates for two types of questionnaire items:

- **Calibrators**: *Non-demographic* items that are not necessarily of substantive interest in the study but are important for explaining differences between probability and nonprobability samples. Calibrators are used to weight the hybrid sample alongside more traditional weighting demographics. A core component of Encipher Hybrid is the SSRS **Calibration Item Bank**: a set of items that SSRS has experimentally validated as being useful for correcting selection bias in nonprobability samples. Because these are non-demographic items that lack externally available benchmarks, weighting benchmarks are instead estimated from the probability portion of the hybrid sample.

- **Outcomes**: The items of substantive interest to the study. Outcomes are not themselves used in weighting, but Encipher Hybrid does use them help identify the optimal weighting model. The **Stepwise Calibration algorithm**, another core component of Encipher Hybrid, searches for a calibration model that minimizes selection bias in the study's most important outcomes. Since outcomes usually do not have external benchmarks, Stepwise Calibration relies on the probability portion of the sample to measure selection bias in the nonprobability portion. This allows us to track the success of each potential weighting model at reducing selection bias in key study outcomes.

This is why full hybrid designs can help maximize data quality when working with nonprobability samples. The data provided by a parallel probability sample, even a small one, are highly valuable for identifying the best calibration model, correcting selection bias in the nonprobability sample, and ultimately obtaining the most accurate weighted estimates.

Sometimes, however, budgets or timelines do not permit the inclusion of even a small probability sample. In these situations, SSRS's *Encipher Nonprob* solution can recover some of the benefits of a full hybrid design.

The Encipher Nonprob Solution

Our Encipher Nonprob solution compensates for the lack of a parallel probability sample in two ways.

First, SSRS **periodically administers "benchmarking surveys" to produce probability-based benchmarks for the calibrators from our Encipher Calibration Item Bank**. These benchmarking surveys are administered to samples from the SSRS Opinion Panel, a mixed-mode probability panel recruited via nationally representative address-based sampling (ABS) and random digit dialing (RDD). This gives our weighting team access to probability-based benchmarks for the items in our Calibration Item Bank. These benchmarks are used as weighting targets for nonprobability-only studies that do not themselves include a parallel probability sample.

The second component of Encipher Nonprob is a **modified version of Stepwise Calibration referred to as Model-Assisted Stepwise Calibration (MASC)**. Since outcome variables are specific to a given study, SSRS's benchmarking surveys can provide probability-based benchmarks for calibrators, but not for outcomes. The Stepwise Calibration algorithm used for Encipher Hybrid cannot be directly applied to nonprobability-only samples because, as described in the hybrid white paper, Stepwise Calibration relies on the internal probability sample to estimate selection bias in outcomes. When we lack a probability sample in which outcomes are observed, there is no way to *directly* measure selection bias in the outcomes.

The Encipher Nonprob MASC method instead uses the nonprobability sample (in which respondents answered the outcome items) to build random forest models predicting a few key outcomes from the study. These models are then used to impute *predicted* responses to the outcome items into the probability-based benchmarking sample (in which respondents did not answer the outcome items). We then use the Stepwise Calibration algorithm to identify the calibration model that minimizes bias (differences between the probability- and nonprobability-based estimates) in the means of these *predicted* outcomes.

Thus, MASC is "model-assisted" in the sense that it relies on predictive models to obtain a "best guess" as to the direction and magnitude of selection bias in the outcomes. The idea of minimizing bias in the mean of a *modeled* outcome, when bias in the *actual* outcome cannot be observed, is adapted from the Standardized Measure of Unadjusted Bias (SMUB) statistic proposed by Little et al. (2020).

In this way, **by leveraging the SSRS Opinion Panel and advanced predictive modeling techniques, Encipher Nonprob allows our Calibration Item Bank and a version of our Stepwise Calibration algorithm to be applied to nonprobability-only samples**. As with Encipher Hybrid, Encipher Nonprob requires that a small selection of calibrators from the Calibration Item Bank—customized to the topic of the study—be included on the study questionnaire.

Validation Study of the SSRS Encipher Nonprob Approach

Study Design

To validate the Encipher Nonprob methodology, we used several surveys that were fielded simultaneously to (1) probability samples selected from the SSRS Opinion Panel and (2) nonprobability samples obtained from an opt-in panel vendor.

These surveys included calibrators from the Calibration Item Bank as well as some selected outcome items covering various topic areas. All items were administered to both the probability and nonprobability samples, allowing us to assess the effectiveness of calibration at reducing selection bias in the outcomes. In the results that follow, we compare estimated outcomes between four hypothetical designs:

- *Probability*: Using only the probability sample from the SSRS Opinion Panel, weighted on standard demographics. This provides the benchmark against which we measure selection bias in the nonprobability designs.
- *Nonprobability – demo*: Using only the nonprobability sample, weighted only on standard demographics (i.e., not using Encipher Nonprob).
- *Nonprobability – Encipher*: Using only the nonprobability sample, applying Encipher Nonprob to calibrate both on standard demographics and on non-demographic calibrators.
- *Pseudo-hybrid*: Using only the nonprobability sample but applying the hybrid version of Stepwise Calibration—that is, optimizing on the *actual* bias in outcomes, rather than on the model-based *prediction* of bias. We call this a "pseudo-hybrid" design because (unlike in a true hybrid design) we still produce the final estimates only using the nonprobability sample; but we use the probability-based estimates to help identify the calibration model that minimizes bias in the nonprobability estimates.[1] We include this design as a point of comparison to assess how well the *Encipher – Nonprob* solution (which assumes that we do *not* have probability-based data for outcomes) replicates the informational advantage afforded by a hybrid design (in which we *do* have probability-based data for outcomes).

[1] To allow for a clean comparison, we do not include a true hybrid design in this comparison. All else equal, a true hybrid sample (using the full hybrid sample, including probability-based completes, to produce estimates) would always show less selection bias than a nonprobability-only sample, because the probability sample is the benchmark used to assess selection bias.

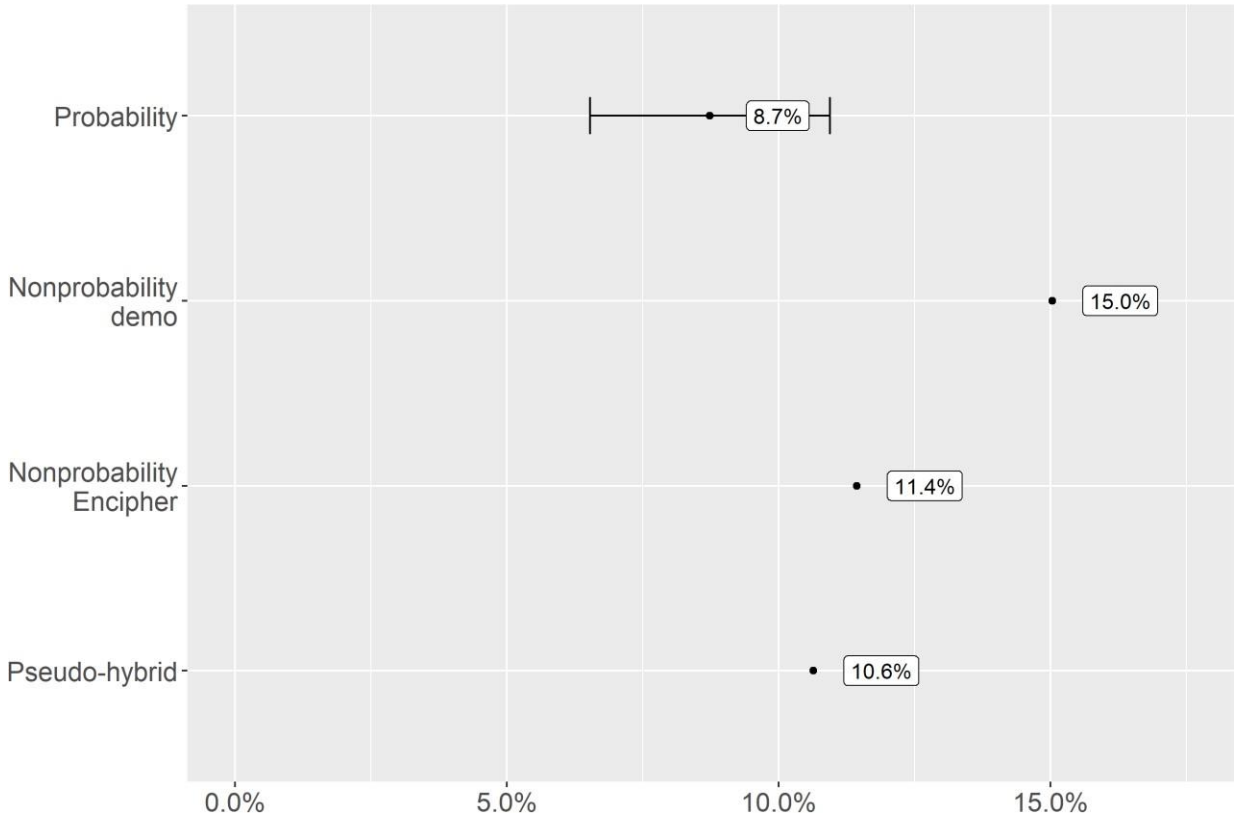
Results

Example Outcome: E-cigarette Use

Figure 1 shows one of the outcome estimates—the percent of adults who use e-cigarettes or other vaping products—under each of the four designs.

If we relied solely on standard demographic weighting, a nonprobability sample (*Nonprobability – demo*) would overestimate this percentage by about 6 percentage points, relative to the probability-based benchmark. **Encipher Nonprob reduces this selection bias to about 3 percentage points, offering a substantial improvement over simple demographic-only weighting** of the nonprobability sample.

Figure 1: Percent using e-cigarettes, by sampling/weighting methodology



NOTE: Error bars show the 95% confidence interval around the Probability estimate.

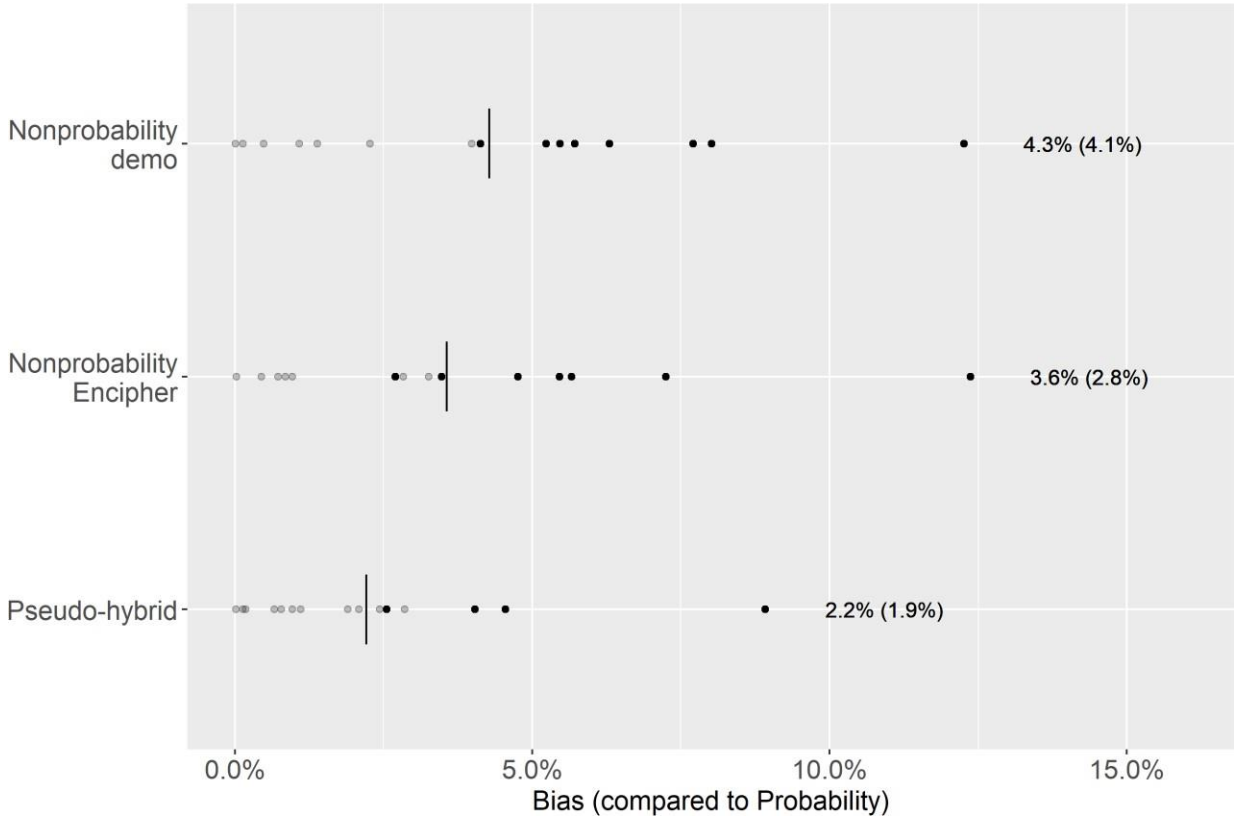
As illustrated by the *Pseudo-hybrid* results, we could further reduce selection bias by using the hybrid version of our calibration procedure, which optimizes on the actual rather than predicted bias. This additional bias reduction illustrates the value of the parallel probability sample within a hybrid design (relative to a nonprobability-only design). Specifically, we can develop a more effective weighting model when we have a probability-based benchmark that allows us to measure the actual selection bias in the outcomes.

However, if this is not an option, the Encipher Nonprob solution, optimizing on the model-based prediction of bias, still improves upon simple demographic-only weighting.]

All Outcomes

Figure 2 generalizes these results, plotting the observed selection bias (i.e., the difference from the probability-based benchmark) for all outcomes collected in the validation study. The figure reports the mean and (due to the presence of one outlier) median selection bias across all outcomes. Overall, although not quite as effective as a hybrid design, **Encipher Nonprob reduces the median selection bias by over 30% relative to standard demographic-only weighting**.

Figure 2: Selection bias in validation study outcomes, by sampling/weighting methodology



NOTE: Vertical line shows the mean bias across outcomes. Data labels report the mean and (in parentheses) median bias across outcomes. Shading indicates statistical significance: light-shaded estimates are within the 95% confidence bounds of the corresponding Probability estimate, while dark-shaded estimates are outside the 95% confidence bounds.

Conclusions

The validation study reported here demonstrates that, by leveraging SSRS benchmarking samples and advanced predictive modeling techniques, **Encipher Nonprob can improve upon standard demographic-only weighting for studies that rely entirely on nonprobability samples**.

That said, this study also reinforces that **hybrid designs should usually be preferred over nonprobability-only designs where feasible**, since the availability of probability-based estimates for key study outcomes allows for the development of more effective calibration models. In a real-world study, without a parallel probability sample in which outcome items were asked, we ultimately cannot know whether the selected calibration model succeeded at meaningfully reducing selection bias.

SSRS's team of methodologists and data scientists can provide consultation to help researchers consider the tradeoffs among the available options for working with nonprobability samples. For many studies, the SSRS Opinion Panel is a feasible source of affordable probability-based samples, making a full hybrid design accessible at a reasonable price point. For researchers who do not wish to administer a full questionnaire to a parallel probability sample, but still want to take advantage of the benefits of a hybrid design, another option would be to run a small sub-selection of critical outcomes and topic-customized calibration items on the SSRS Opinion Panel Omnibus.

References

Little, R.J.A., West, B.T., Boonstra, P.S., and Hu, J. (2020). "Measures of the Degree of Departure from Ignorable Sample Selection." *Journal of Survey Statistics and Methodology 8*(5): 932 – 964.
This paper introduces the new SSRS Encipher Hybrid methodology for blending probability and nonprobability samples. Encipher is a unique and sophisticated method that leverages study-specific outcomes, advanced modeling techniques, and customized non-demographic measures to produce weighting margins that are optimized for reducing selection bias in key study outcomes. In a validation study whose results are reported here, Encipher:

- Reduced bias in topline estimates by nearly 60% relative to a nonprobability-only sample.
- Reduced bias in subgroup estimates (including breakouts by gender, age, educational attainment, and race) by similar amounts.
- Substantially increased effective sample sizes relative to a probability-only sample.

For more information about how Encipher Nonprob could be useful for your study, Visit https://ssrs.com/encipher-nonprob/.