

Proceedings of Statistics Canada Symposium 2024: The Future of Official Statistics

A Bias Evaluation for Probabilistic Web Panels at Statistics Canada

by Anne Mather, Cilanne Boulet, and Youngchang Ra

Release date: September 8, 2025



A Bias Evaluation for Probabilistic Web Panels at Statistics Canada

Anne Mather, Cilanne Boulet, and Youngchang Ra¹

Abstract

Three series of web panels were implemented at Statistics Canada from 2020 to 2024. Participants for these web panel series were recruited from respondents of large probabilistic social surveys (recruitment surveys), and subsequently were invited to complete a series of short online surveys. Estimates of recruitment survey variables were calculated using both recruitment survey weights and web panel weights, and these were compared; differences signal the possibility of residual bias that was not corrected by the web panel weighting process. This investigation found more significant differences than would be expected if the web panel estimator fully corrected for the bias resulting from the web panel response process. Questions related to certain topics such as politics and voting, sense of belonging, and media consumption were found to have the most significant differences between web panel estimates and recruitment survey estimates.

Key Words: Web panel; Non-response bias.

1. Introduction

1.1 Context

Web panels are a collection method that Statistics Canada began to implement in 2020. (For a discussion of the initial implementation, see Baribeau, 2020.) In this collection method, respondents of a Statistics Canada survey (referred to in this context as a “recruitment survey”) are invited to opt into an additional series of short online surveys by providing contact information. When each panel survey (“wave”) is conducted, those who opted in are contacted and invited to participate. These web panels are designed to allow for timely, efficient, and cost-effective collection of data to produce high-level estimates that can be used to address emerging needs.

This study focuses on the first three web panels conducted at Statistics Canada. The first two were the Canadian Perspectives Survey Series (CPSS) and the Portrait of Canadian Society (PCS), with responses collected exclusively online; the bias evaluation described in this paper is covered in greater detail in a Statistics Canada Methodology working paper (Mather and Boulet, 2024). The third panel, the Survey Series on People and their Communities (SSPC), differed somewhat in that it ran its own recruitment survey which framed the request for contact information differently. Furthermore, though its collection for the subsequent short surveys was primarily through online self-response, it also included some follow-up and collection by telephone. More collection details are available in MacIsaac, Boulet, and Thomas (2025). Despite the telephone collection on SSPC, the term web panel is used throughout this article since the primary mode of collection for all three panels was through online self-response.

1.2 Web Panel Response Rates

The three web panels in this study are outlined in Table 1.2-1. As was expected for this collection method, the web panels had lower response rates than are typical of Statistics Canada surveys that use traditional collection methods. When recruitment survey response rates, web panel opt-in rates, and web panel response rates are all taken into account, cumulative response rates were between 3.7% and 15.4%.

¹Anne Mather, Statistics Canada, 150 Tunney’s Pasture Driveway, Ottawa, Ontario, K1A 0T6, anne.mather@statcan.gc.ca; Cilanne Boulet, Statistics Canada, 150 Tunney’s Pasture Driveway, Ottawa, Ontario, K1A 0T6, cilanne.boulet@statcan.gc.ca; Youngchang Ra, Statistics Canada, 150 Tunney’s Pasture Driveway, Ottawa, Ontario, K1A 0T6, youngchang.ra@statcan.gc.ca.

Table 1.2-1
Participation rates for the first three probabilistic web panels implemented at Statistics Canada

Panel	Recruitment Survey (Collection Period)	Recruitment Survey Response Rate	Provided Contact Information	Web Panel Response Rates (Range Across Waves)	Cumulative Response Rates (Range Across Waves)
Canadian Perspectives Survey Series	Labour Force Survey (April-July 2019)	87%	23%	54%-64%	10.7%-12.6%
Portrait of Canadian Society	General Social Survey – Social Identity (2020)	40%	22%	41%-44%	3.7%-3.9%
Survey Series on People and their Communities	Survey Series on People and their Communities (2022-2023)	53%	97%	24%-28%	12.9%-15.4%

This current study explores the impact of these low response rates by measuring the extent to which estimates obtained using panel respondents mirror the equivalent estimates using the recruitment surveys. Ultimately, we found that weighting can be helpful for mitigating differences, but that many remain statistically significant.

Section 2 outlines the weighting methods employed in web panels. Sections 3 and 4 describe the methodology of the bias evaluation and its results. Finally, section 5 includes a discussion of the results, including lessons learned and future directions this work could take.

2. Web Panel Weighting

2.1 Weighting Process

A weighting process was applied to each web panel wave in order to assign a weight to each respondent, to make the respondents more representative of the target population. Thus, a different set of weights is produced for each wave. The weighting processes attempt to account for the sampling designs of and the non-response to both the recruitment surveys and the web panels themselves. In particular, the weights encompass both parts of non-response to the web panels: not opting in and not responding after having opted in.

Broadly, the web panel weighting process consists of:

1. *Assigning a design weight:* The design weight for the web panels is the starting point for the weighting process. Weights from the recruitment survey – which reflect its design and have been adjusted to account for its non-response – are taken as the design weights for the web panel.
2. *Adjusting this weight for web panel non-response:* The weights of web panel respondents are adjusted to account for web panel non-response (amongst recruitment survey respondents). The non-response adjustment can leverage many variables from the recruitment survey.
3. *Calibrating:* The weights are calibrated so that their sums match known or projected target population totals.

Of particular note are the variables available for the non-response adjustment step. Given that these web panel series were recruited from respondents to large Statistics Canada surveys, many variables (including socio-demographic traits and subject matter content) were available for both respondents and non-respondents. Additionally, the Canadian Census of Population was used as a frame for the SSPC, providing even more variables in this case. The richness of this data allowed for insight into which characteristics were over- or under-represented amongst web panel respondents; for example, it was observed that younger individuals tended to be less likely to respond and that individuals in households with higher incomes were more likely to respond. These variables were available for use in non-response adjustments in the web panel weighting process.

For each web panel wave, the recruitment survey’s bootstrap replicate weights were adjusted following the same steps as the web panel weighting process. These bootstrap weights allow for estimations of variance and were used in this study.

2.2 Design Effect and Effective Sample Size

The design effect reflects the ratio of the variance of an estimator under a sampling plan and the variance that would hypothetically be expected had the sample been a simple random sample of the same size. While the design effect depends on the estimator selected and is different for each parameter, for a non-clustered design, it can be estimated based on the dispersion of the weights (Kish, 1965). Notably, it can be used to estimate the effective sample size, calculated as the number of respondents divided by the design effect. Table 2.2-1 shows estimates of the effective sample sizes for the recruitment surveys and web panel waves.

Table 2.2-1
Effective sample sizes of recruitment surveys and web panels

Panel	Recruitment Survey Effective Sample Size	Web Panel Effective Sample Size (Range Across Waves)
Canadian Perspectives Survey Series	36,568	1,100-1,570
Portrait of Canadian Society	9,152	490-580
Survey Series on People and their Communities	7,705	1,750-2,000

The effective sample sizes of web panels are substantially lower than those of their respective recruitment surveys. Thus, a question asked on a wave of a web panel can be considered to have only a small fraction of the statistical power of the recruitment survey. This limited power will have an impact on any analysis carried out.

3. Methodology of Bias Investigation

3.1 Methodology

Given their low cumulative response rates and the use of online self-response as the only or main response mode, the risk of bias on web panels must be investigated. The main objective of the evaluation that follows is to quantify the degree to which web panel weights are able to adjust for differences between respondents and non-respondents.

The way in which this was done was to work with variables from the recruitment surveys and measure the differences observed when generating estimates with recruitment survey weights and web panel weights. These differences correspond to the extent to which web panel estimates differ from recruitment survey estimates of the same variables. Variables available for analysis include all questions asked on the recruitment surveys. Most of these variables are categorical and it is on this type of variable that the analysis below is undertaken. The distributions of these categorical variables were estimated using the respondents and weights from the recruitment surveys. For each wave of the web panels, these estimates were then compared to (1) estimates using the recruitment survey weights for web panel respondents only (referred to below as the initial web panel estimates) and (2) estimates using the final web panel weights. Comparisons were made using a Satterthwaite adjusted F-test (Shah, 1976, as cited in Research Triangle Institute, 2012, pp. 216-218); estimates were deemed to be different if the p-value of the test was below a threshold of $\alpha = 0.05$.

3.2 Categorization

Based on F-tests, each estimate from a web panel wave was categorized into one of the following groups:

- A. *Neither the initial nor final web panel distribution is significantly different from the recruitment survey distribution:* The difference between the recruitment survey estimate and that of the initial web panel estimate is not statistically significant, and neither is the difference between the recruitment survey estimate and the final web panel estimate.

- B. *The final distribution is not significantly different from the recruitment survey distribution:* There is a statistically significant difference between the recruitment survey estimate and the initial web panel estimate, but no statistically significant difference between the recruitment survey estimate and the final web panel estimate.
- C. *The weighting process improves the similarity of the final distribution to the recruitment survey distribution, but insufficiently:* There is a statistically significant difference both between the recruitment survey estimate and the initial web panel estimate, and between the recruitment survey estimate and the final web panel estimate. However, the latter p-value is greater than the former, indicating a closer fit to the recruitment survey estimate.
- D. *Others:* The final web panel estimate is significantly different from the recruitment survey estimate, and the weighting process does not improve the similarity to the recruitment survey estimate.

In all of these comparisons, statistical difference was measured at $p < 0.05$. Categories A and B contain variables whose estimates at the final stages of weighting are not statistically different when compared to the recruitment survey. On the other hand, categories C and D contain variables where the final estimates are statistically different.

4. Results

The proportion of variables in each category was analyzed by topic on each survey series. Figures 4-1 to 4-3 show the categorizations of estimates from each web panel series, with all waves of each series grouped together and then displayed by the topics of the questions. Note that only variables with an effective sample size of at least 300 respondents in the web panel estimate were kept; this threshold was selected as one that allowed for sufficient power to detect differences. The results below include 29-31 Labour Force Survey variables for each wave of CPSS, 374-385 General Social Survey – Social Identity variables for each wave of PCS, and 167-168 variables from the SSPC recruitment survey and 140-141 variables from the Census 2021 long-form survey (available on the SSPC recruitment survey frame) for each wave of SSPC.

When using the threshold of $\alpha = 0.05$, the proportion of variables categorized as A or B would be expected to be at 95% if the estimator used were unbiased; however, it is quite frequently below this level. Of particular note are a few topics at the top of each figure (e.g., questions related to politics and voting in Figure 4-2, or to sense of belonging in both 4-2 and 4-3) on which estimates between the recruitment surveys and web panels are very frequently found to be significantly different. These variable groupings, which have a lower proportion of variables categorized as A or B compared to the other variable groupings, are topics on which the web panel respondents differed from the non-respondents, and on which the weighting process did not sufficiently adjust for the differences. Manual investigation of some variables in these categories revealed that web panel estimates were more likely than recruitment survey estimates to point to, for example, higher levels of homeownership, more frequent consumption of news, and greater engagement in political activities.

Figure 4-1
Percentage of estimates in each category, using weights from the Canadian Perspectives Survey Series

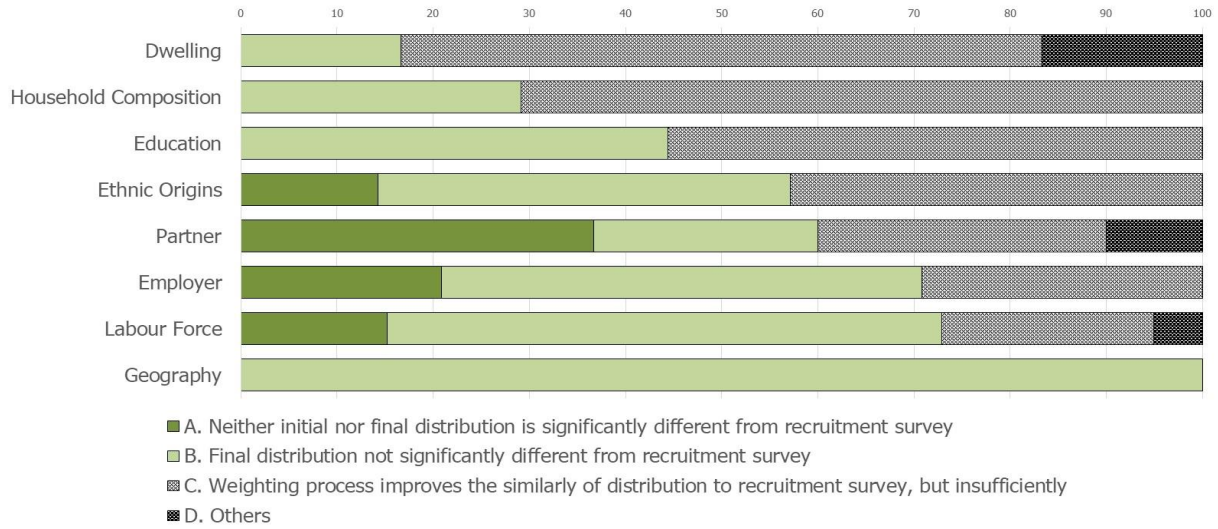


Figure 4-2
Percentage of estimates in each category, using weights from the Portrait of Canadian Society

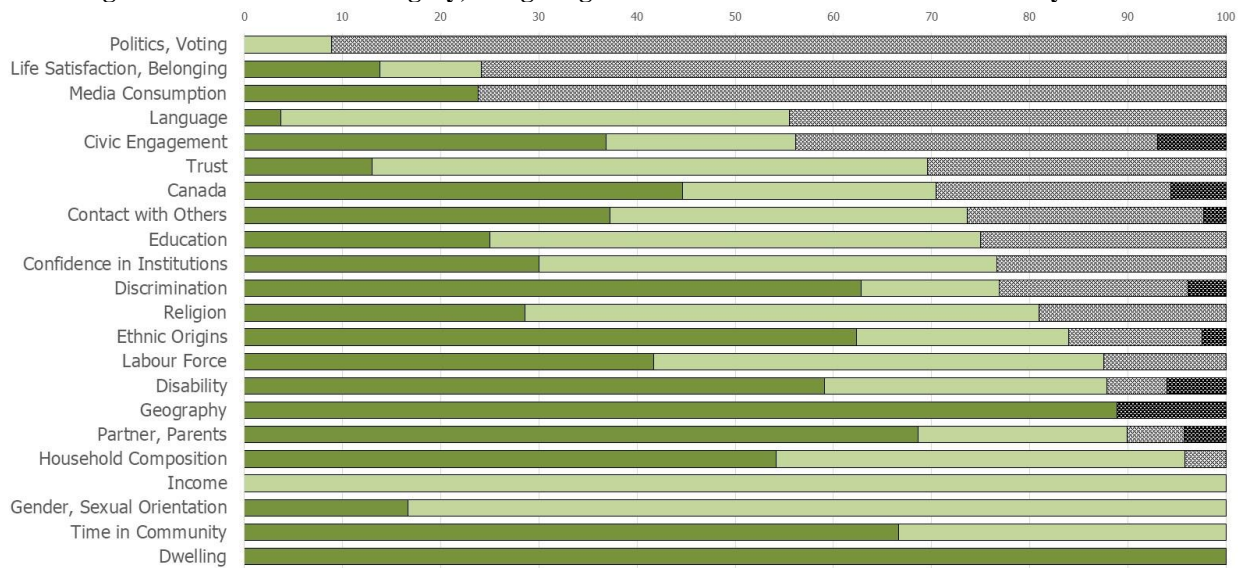
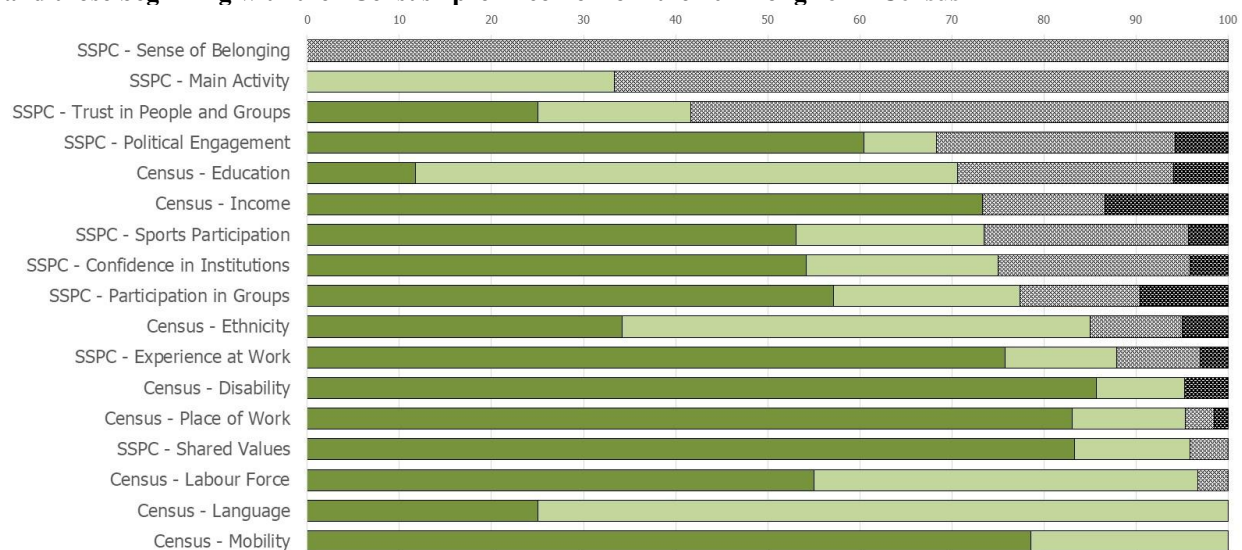


Figure 4-3
Percentage of estimates in each category, using weights from the Survey Series on People and their Communities; question groupings beginning with the “SSPC” prefix come from the SSPC recruitment survey and those beginning with the “Census” prefix come from the 2021 long-form Census



5. Lessons Learned and Future Work

The low cumulative response rates of web panels introduce a risk of bias. This evaluation shows that substantially more significant differences were detected between recruitment survey estimates and web panel estimates than would be expected were the estimator unbiased. However, it should be noted that some of the significant differences detected were not particularly large, and the data could remain fit for purpose in certain circumstances. As laid out in the Statistics Canada Quality Guidelines (Statistics Canada, 2019), data quality is a multidimensional concept. Producing fit-for-purpose surveys can require trade-offs between these dimensions; in particular, the appropriate balance between accuracy and timeliness varies depending on the specific needs to be met by the survey and might be constrained by budget and operational considerations. It is recommended that, in general, web panel results be interpreted with caution and limitations be communicated to users.

This analysis also highlights the impact of the small effective sample sizes of web panels, resulting from both the relatively low number of respondents and the increase in design effects resulting from weighting adjustments. These small effective sample sizes limit the statistical power available to analyses using web panel data.

It is interesting to note the importance and impact of the web panel weighting process. The richness of the auxiliary information available from recruitment surveys allows for some powerful adjustments for non-response. This is evidenced by the large number of estimates that fell into categories B and C, for which the weighting process improved the similarity of the web panel estimate’s distribution to the recruitment survey estimate’s. However, this process was not consistently able to fully account for the differences between web panel respondents and non-respondents. Certain areas of particular concern were highlighted, such as politics and voting, media consumption, and sense of belonging.

Additionally, given that response rates have been falling for many social surveys, the conclusions drawn here may be applicable beyond the web panel context. This analysis points to non-response bias on web panels with low response rates. This could also be the case for other low-response-rate surveys, highlighting the importance of pushing to get as much response as possible on all types of surveys. It also suggests that certain themes may be at higher risk of non-response bias for surveys in general, such as those related to societal participation and engagement.

Finally, further analyses of the correlations between recruitment survey questions and web panel questions would be useful in understanding the extent to which the latter are at risk of bias. The results presented here did not examine

estimates of web panel questions, but it may be the case that these questions are correlated with some that were in the analysis. If web panel questions are highly correlated with recruitment survey questions on which differences were observed, the residual bias could be better understood and reported, even if not fully corrected.

6. Acknowledgements

The authors would like to acknowledge the contributions of many colleagues and reviewers. Special thanks to Beatrice Baribeau, Eliza Beckett, Nora Bohossian, Krista MacIsaac, Samuel MacIsaac, Isabelle Michaud, John Stardom, Claude Turmelle, and Yu Wang.

References

- Baribeau, B. (2020). Breaking with Tradition: Piloting a Probability Web Panel. Methodology Branch Working Paper SSMD-2020-02E. Statistics Canada.
- Kish, L. (1965), Survey Sampling, New York: Wiley.
- Mather, A. and Boulet, C. (2024). Analysis of Residual Bias on Web Panels. Methodology Branch Working Paper SSMD-2024-002E. Statistics Canada.
- MacIsaac, K., Boulet, C. and Thomas, M. (2025). “Recruitment and Collection of Web Panels at Statistics Canada”. *Proceedings of Statistics Canada Symposium 2024*.
- Research Triangle Institute. (2012). SUDAAN Language Manual, Volumes 1 and 2, Release 11. Research Triangle Park, NC: Research Triangle Institute.
- Statistics Canada. (2019). Statistics Canada Quality Guidelines: Sixth Edition – December 2019. Available at https://www150.statcan.gc.ca/n1/en/pub/12-539-x/12-539-x2019001-eng.pdf?st=0mOXi_LG