Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Comments on “Exploring the assumption that commercial online nonprobability survey respondents are answering in good faith”

by Aditi Sen

Release date: June 25, 2024

Statistics Canada Statistique Canada

Canadä

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, June 2024
Vol. 50, No. 1, pp. 29-32
Statistics Canada, Catalogue No. 12-001-X

Comments on "Exploring the assumption that commercial online nonprobability survey respondents are answering in good faith"

Aditi Sen[1]

Abstract

This discussion summarizes the interesting new findings around measurement errors in opt-in surveys by Kennedy, Mercer and Lau (KML). While KML enlighten readers about "bogus responding" and possible patterns in them, this discussion suggests combining these new-found results with other avenues of research in nonprobability sampling, such as improvement of representativeness.

Key Words: Opt-in survey; Measurement error; Data quality; Data integration; Inverse propensity score weighting.

KML in their seminal research focus on the important aspect of measurement error in nonprobability surveys, especially commercial online ones, referred to as "opt-in-surveys". Advanced methods of estimation of population characteristics from nonprobability surveys are commonly developed under the assumption of accuracy of survey responses. In the presence of inaccurate responses, where this assumption is violated, those methods may be inadequate. Thus, when KML question the accuracy of individual survey responses in opt-in surveys, our attention is drawn to this serious issue that calls for further research on the problem.

The most popular type of web survey is based on the so-called "opt-in" or volunteer panel. Unlike probability surveys where a sample, representative of the population, is drawn from a frame, opt-in panels are not constructed using a probability-based design. In such a panel, volunteers are recruited through various convenient but nonprobability methods like quota sampling, snowball sampling etc. and people often join to receive some kind of incentive. Issues like increasing cost and declining response rates of probability surveys are well talked about. In the age of big data and fast and efficient computer programming capabilities, opt-in surveys are receiving much interest. These surveys cost less and panel recruitment as well as receiving responses from volunteers can be achieved quickly. However, there is no guarantee that such samples properly represent the target population. In addition, as KML point out, there is a concern that responses may not be genuine. There is a chance that some respondents might be driven by the incentive or may intentionally provide nonsensical responses. KML highlight these issues of opt-in surveys and find interesting pathways for future research in nonprobability surveys.

Groves and Lyberg (2010) discuss the total survey error framework starting from error typology from Deming's 1944 American Sociological Review article where bias components of error versus variance components of error are clearly noted. It is also noted that earlier sampling theories and methods are most applicable when nonsampling errors are small. Nonsampling errors include nonresponse errors,

1. Aditi Sen, University of Maryland College Park, USA. E-mail: asen123@umd.edu.

measurement errors and so on that are not related to the sample selection. Hansen, Hurwitz and Bershad (1961) in their paper mention that the collection and processing operations of sample surveys constitute the measurement process and are the sources of measurement error. KML quantify measurement error in individual responses by comparing the average size of errors of opt-in surveys with probability surveys and thus connect two areas of methodological research around data quality in opt-in panels.

KML's paper focuses on the not-so-talked about area of nonsensical responses in opt-in-surveys, which are termed as "bogus responses". The use of benchmarking helps to formulate the hypothesis that these answers to survey questions are nonsensical. This hypothesis is further tested using a "follow-up survey" where questions based on rare events are asked, to which an affirmative response is highly unlikely. In total they work with six surveys; three of them are commercial opt-in surveys where vendors used quota sampling (with the 2019 American Community Survey, ACS, as target) to select the samples. The other three are probability surveys where panelists are recruited using address-based sampling (ABS). In the benchmarking study responses to 25 questions common to all these surveys (treated as estimates) are compared with those from government surveys (treated as true values) like the ACS, the Current Population Survey (CPS) and the National Health Interview Survey (NHIS) in terms of Mean Absolute Error (MAE). A thought along the lines of the follow-up survey is as follows: suppose that the main survey questionnaire could be planned in such a way as to include a set of special "detective" questions (in addition to the main questions). Responses to these questions would be used to measure "probabilities of a bogus response" given covariates. These probabilities of a bogus response could be used to downweight individual responses. For example, we provide the extreme values of the weight adjustment, which is between 0 and 1, using the following conditions: if the probability that a response is bogus is 1 then the weight adjustment is 0, again if the response is well trusted then the weight adjustment is 1. A method could be developed to incorporate these probabilities along with the usual response participation probability weighting, to simultaneously improve representativeness and account for the probability of a bogus answer.

Most interestingly, thankfully to the research of KML, we learn about the presence of patterns in such bogus responding. As the authors indicate, theirs is the first paper that uses benchmarking to identify subgroups that have a high probability of bogus responses. When sub-grouped by demographic variables of respondents, one at a time, primarily age and race-ethnicity, it is observed from the three opt-in-surveys that young (age 18-29 years) and Hispanic respondents are more prone to such trends. Of course, these groupings are subjective; there are other variables like gender, education that are not found to be very impactful in distinguishing those stark differences. It is understandable that considering interaction and subdivision into multiple domains would decrease sample size considerably. In such scenarios one can think of applying small area techniques. Ghosh (2020) provides a great review of different small area models and methods. Here a question can be raised: can a statistical tool be developed, using machine learning methods like regression trees and such, that would help find significant variables to identify bogus responding? This might help to discover the interaction effect between variables unlike considering one at a time, as done by KML. In the context of grouping using machine learning, Loh (2011) reviews some widely available algorithms on classification and regression trees.

Literature on nonprobability surveys has focused on the improvement of representativeness, i.e., to reduce selection bias to make the sample more representative of the population. Much work has been done on the inverse propensity score weighting (IPSW) methods where the propensity is defined by the participation probability of population units in the nonprobability sample. Valliant and Dever (2011), Chen, Li and Wu (2020), Wang, Valliant and Li (2021), Savitsky, Williams, Gershunskaya, Beresovsky and Johnson (2023) derive methods involving combining/stacking nonprobability surveys with probability/reference surveys to estimate participation probabilities, which are otherwise unknown for nonprobability surveys due to their unknown selection mechanism. Here assumptions about ignorability and strict positivity of propensity scores need to be made. These methods generally define a log-likelihood upon creating an indicator variable defining success if a unit is present in the nonprobability sample. The information about the whole finite population being unknown, the subsequent modification of the likelihood into a pseudo-likelihood and use of the reference survey depend on the combining methodology. Researchers thus estimate the propensity score with the help of different data integration techniques and support the performances of estimators in terms of bias and variance with the help of simulation studies and real-life datasets. To be specific, Chen, Li and Wu (2020) compute a doubly robust estimator for finite population means where the name doubly robust comes from two models: one is the propensity score model and the other is the outcome regression model.

To put all these into the context of discussion of the paper by KML, it would be of interest to see the effect of using the ideas developed in the aforementioned papers to estimate the weights for opt-in surveys. Currently the weights in question are developed using calibration to match with population characteristics, but will such estimation procedures involving data integration affect the measurement errors due to bogus responding? KML's paper throws light on the fact that it is not advisable to directly use the responses from opt-in surveys. Researchers should check the credibility of such responses with the help of available probability surveys, like benchmarking with government surveys. For survey statisticians it would be beneficial to know how to properly combine opt-in-surveys with other sources which help validate their credibility and help improve responses or eliminate bogus responding. In essence, the outstanding ideas developed by KML throw the readers into a unique direction of thought, focusing on nonsampling errors. This, combined with the recent development of methodologies on improved representativeness of nonprobability samples, provide us with innovative research outcomes to look forward to.

References

Chen, Y., Li, P. and Wu, C. (2020). Doubly robust inference with nonprobability survey samples. *Journal of the American Statistical Association*, 115(532), 2011-2021.

Ghosh, M. (2020). Small area estimation: Its evolution in five decades. *Statistics in Transition*, New Series, Special Issue on Statistical Data Integration, 1-67.

Groves, R.M., and Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74, 5, 849-879. Doi: https://doi.org/10.1093/poq/nfq065.

Hansen, M., Hurwitz, W. and Bershad, M. (1961). Measurement errors in censuses and surveys. *Bulletin of the International Statistical Institute*, 32nd Session, 38, Part 2, 359-74.

Loh, W.-Y. (2011). Classification and regression trees. *WIREs Data Mining Knowl Discov*, 1, 14-23. Doi: https://doi.org/10.1002/widm.8.

Savitsky, T.D., Williams, M.R., Gershunskaya, J., Beresovsky, V. and Johnson, N.G. (2023). Methods for combining probability and nonprobability samples under unknown overlaps. Doi: https://doi.org/10.48550/arXiv.2208.14541.

Valliant, R., and Dever, J.A. (2011). Estimating propensity adjustments for volunteer web surveys. *Sociological Methods and Research*, 40, 105-137.

Wang, L., Valliant, R. and Li, Y. (2021). Adjusted logistic propensity weighting methods for population inference using nonprobability volunteer-based epidemiologic cohorts. *Statistics in Medicine*, 40(4), 5237-5250.