Nonparticipation (Unit Nonresponse) In Surveys:

A Practitioner’s Guide to the Conceptualization, Impact of, and Adjustment for Unit Nonresponse

Pamela Woitschach, Ph.D.

Post-Doctoral Research Fellow, BC Office of Patient-Centred Measurement

BC SUPPORT Unit

Post-Doctoral Research Fellow, UBC-Paragon Research Initiative

University of British Columbia

Bruno D. Zumbo, Ph.D.

Professor & Distinguished University Scholar

Tier 1, Canada Research Chair in Psychometrics and Measurement; &

Paragon UBC Professor of Psychometrics and Measurement

University of British Columbia

How to Cite This Report

Woitschach, P., & Zumbo, B.D. (2021, March 31). *Nonparticipation (Unit Nonresponse) In Surveys: A Practitioner's Guide to the Conceptualization, Impact of, and Adjustment for Unit Nonresponse [webinar]*. British Columbia Patient-Centred Measurement.

We would like to thank Dr. Beth Griffin, RAND Corporation, for her support and advice in the use of the TWANG package.

Nonparticipation (Unit Nonresponse) In Surveys:
A Practitioner’s Guide to the Conceptualization, Impact of, and Adjustment for Unit Nonresponse

Webinar Part 3
Approaches to Address Unit Nonresponse

Pamela Woitschach, Ph.D.

Post-Doctoral Research (part-time) Fellow, Patient-Centred Measurement Methods Cluster

BC SUPPORT Unit

Post-Doctoral Research Fellow, UBC-Paragon Research Initiative

University of British Columbia

Outline

- Adjusting for Nonresponse
  I. Defining Unit Non-Response (Recap from Webinar Part 1).
  II. Adjusting for Nonresponse in Survey Research.
  III. Nonresponse Guidelines (Statistics Canada, 2020).
- Imputation & Weighting: Important concepts
  I. Multiple Imputation.
  II. Weighting.
  III. Multiple Imputation vs. Inverse Propensity Weighting.
- Propensity Score Weighting Approach
  I. PSW definition.
  II. PSW estimators.
  III. PSW software package.
- Demonstration of PSW computation.
- Framework to approach Unit Nonresponse.
- Concluding Remarks.

"No issue in survey research is more misunderstood or controversial than nonresponse."

(Dixon & Tucker, 2010)

Non-Participation

participation and non-participation health care
non-response bias mailed health survey
survey participation health expectancies
national health survey participation barriers
health survey non-response
health survey participation of households
reasons for non-participation population survey
non-participation and mortality prospective study

Representativeness

National surveys
Cross-sectional health survey

Missing data

Unit missing data

Item missing data

partially missing data health scores
missing data large surveys
missing data mechanisms max study
imputation methods for missing data
imputation algorithm health surveys

Non-response

Unit non-response

Partial/total nonresponse

predictor of unit nonresponse
unit nonresponse weighting adjustmentsunit
nonresponse measurement error
response and nonresponse bias health surveys
unit nonresponse simulation study
health survey non response
variance of survey weighting for nonresponse
nonresponse in sample surveys
health interview surveys
income nonresponse consumer survey
nonresponse bias mail survey

Reminder of some key ideas from earlier Parts of the Webinar

Defining Unit Non-Response (From Webinar Part I).

Unit Nonresponse: Occurs when no survey data are collected for an element selected for the sample. Nonresponse results from refusals to participate in the survey, noncontacts, and other reasons such as language barrier, among others (Brick et al. 1996, p.1).

Ignorable Nonresponse: MCAR and MAR

Non-Ignorable Nonresponse:The survey response mechanism depends on a variable of interest measured within the same survey and observed for only part of the sample (MNAR).

Two models of response: Deterministic (Respondents and Nonrespondents) and Stochastic (Response Propensity).

Defining Unit Non-Response (From Webinar Part I).

	IDENT	Variable_1	Variable_2	Variable_3
1	1.00	1.03	59.27	.53
2	2.00	-.13	48.80	-1.17
3	3.00	.66	55.93	.94
4	4.00	-.23	47.90	-1.41
5	5.00	.77	56.97	-.41
6	6.00	.93	58.38	.42
7	7.00	.59	55.30	.63
8	8.00	.95	58.51	1.15
9	9.00	.59	55.32	-.10
10	10.00	-.35	46.89	-.58
11	11.00	.58	55.25	.46
12	12.00	-.18	48.42	1.20
13	13.00	.70	56.28	.36
14	14.00	-1.60	35.60	-.04
15	15.00	-1.07	40.33	-.99
16	16.00	.77	56.89	-1.26
17	17.00	-.67	43.94	.06
18	18.00	.91	58.20	1.41
19	19.00	1.03	59.26	2.78
20	20.00	-.20	48.22	-.18

Ideal case- no missing data

	IDENT	Variable_1	Variable_2	Variable_3
1	1.00	1.03	59.27	.53
2	2.00	-.13	.	-1.17
3	3.00	.66	55.93	.94
4	4.00	-.23	47.90	-1.41
5	5.00	.77	56.97	.
6	6.00	.	58.38	.42
7	7.00	.59	55.30	.63
8	8.00	.95	.	1.15
9	9.00	.59	55.32	-.10
10	10.00	.	46.89	-.58
11	11.00	.58	.	.46
12	12.00	-.18	48.42	.
13	13.00	.70	56.28	.36
14	14.00	-1.60	35.60	-.04
15	15.00	.	.	-.99
16	16.00	.77	56.89	-1.26
17	17.00	-.67	43.94	.06
18	18.00	.	58.20	1.41
19	19.00	1.03	59.26	.

Item nonresponse

	IDENT	Variable_1	Variable_2	Variable_3
1	1.00	1.03	59.27	.53
2	2.00	.	.	.
3	3.00	.66	55.93	.94
4	4.00	-.23	47.90	-1.41
5	5.00	.77	56.97	-.41
6	6.00	.93	58.38	.42
7	7.00	.59	55.30	.63
8	8.00	.95	58.51	1.15
9	9.00	.	.	.
10	10.00	.	.	.
11	11.00	.58	55.25	.46
12	12.00	-.18	48.42	1.20
13	13.00	.70	56.28	.36
14	14.00	.	.	.
15	15.00	-1.07	40.33	-.99
16	16.00	.77	56.89	-1.26
17	17.00	-.67	43.94	.06
18	18.00	.	.	.

Unit nonresponse

Nonresponse rates in probability samples are increasing worldwide
(Lohr et al., 2016, p.195).

Adjusting for Nonresponse in Survey Research

Imputation and weighting are the general approaches used to adjust for nonresponse in survey research. As a rule of thumb, *imputation* is mainly used to adjust for *item nonresponse*.

However, in some situations, imputation is also an approach used to address unit nonresponse. Those adjustment methods make use of covariates that are available for both respondents and nonrespondents.

Imputation and weighting share a common goal: to reduce the nonresponse bias and control the nonresponse variance (Gelein et al., 2018).

Statistics Canada Quality Guidelines: Response and Nonresponse .

- Setting an anticipated response rate.
- Reducing nonresponse.

- Including nonresponse follow-up procedures.
- Assessing potential nonresponse bias.
- Determining the response mechanism.

- **Deciding how to handle nonresponse.**
- Evaluating and disseminating nonresponse rates.
- Identifying and analyzing reasons for nonresponse.

- Quality indicators.
    - Evaluating response and nonresponse rates.
    - Evaluating nonresponse variance.
    - Examining bias.

Statistics Canada (2020). Statistics Canada Quality Guidelines: Response and Nonresponse https://www150.statcan.gc.ca/n1/pub/12-539-x/2009001/response-reponse-eng.htm

Adjusting for Nonresponse in Survey Research

Some questions need to be answered before considering which one is the most effective approach to adjust for unit nonresponse.

- Can we assume the ignorability of the nonresponse mechanism?
- Are there meaningful differences between the groups of respondents and nonrespondents?
- What are the claims we want to make from our data/results?
- How many cases we have in each strata/class?
- Can we justify the use of one or the other approach?

Treating it as a "missing data" problem and using imputation

Multiple imputation is an approach that involves filling missing values in variables using a selected imputation method and repeating the process multiple times, creating several data sets (Peytchev, 2012, p. 217).

Main approach	Tool	Description	Data required	Challenges	Bias
Imputation (Pike, 2007)	Hot-deck imputation	Data from a survey respondent is copied to represent data from a nonrespondent who has characteristics that are similar to those of the respondent.	Sociodemographic characteristics from the sampling frame.	- Similarity between groups. - Comparability between groups. - With low response rates, a small number of respondents may be used to represent a large number of nonrespondents.	Hot-deck imputation Increases statistical power, and it can also increase variance (i.e., decrease the precision) of the estimator.

Nonresponse Line of Research – MRC/CSO Social and Public Health Sciences Unit, University of Glasgow

Soc Psychiatry Psychiatr Epidemiol (2016) 51:155–157
DOI 10.1007/s00127-015-1153-8

CrossMark

COMMENTARY (INVITED)

The importance of post hoc approaches for overcoming non-response and attrition bias in population-sampled studies

Linsay Gray[1]

Received: 30 October 2015 / Accepted: 8 November 2015 / Published online: 28 November 2015

Abstract Population-based health studies are critical resources for monitoring population health and related factors such as substance use, but reliable inference can be compromised in various ways. Non-response and attrition are major methodological problems which reduce power and can hamper the generalizability of findings if individuals who participate and who remain in a study differ systematically from those who do not. In this issue of SPPE, McCabe et al. studied participants of the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions, comparing attrition in Wave 2 across participants with different patterns of substance use at Wave 1. The implications of differential follow-up and further possibilities for addressing selective participation are discussed.

Keywords Attrition · Non-response · Bias · Population based · Substance use

(or “non-participation”)—where people who have been selected for inclusion have not participated—and “item missingness” where participants have not provided data for all individual variables. In longitudinal studies, “attrition”—the loss of follow-up of cohort members over time—is yet another facet of the missing data problem. All three aspects lead to subsequent loss of power, increasing the chances of both type I (false positive) and type II (false negative) errors. The potential for bias is also elevated if certain sub-groups of individuals are systematically missing: under these conditions, associations among variables which may not be true can arise, and vice versa (internal validity is compromised) and the extent to which results are generalizable to the population (external validity) is threatened. The occurrence of missing data leads to bias in analysis unless the underlying mechanism of missing data is ‘missing completely at random’ (MCAR). MCAR relies on the probability of participation/of not answering a particular question/of dropping-out of follow-up being

(**Research team main publications:** Gorman et al., 2014; Gorman et al., 2017; Gray, 2016; Gray et al., 2013; McMinn et al., 2020; McMinn et al., 2018)

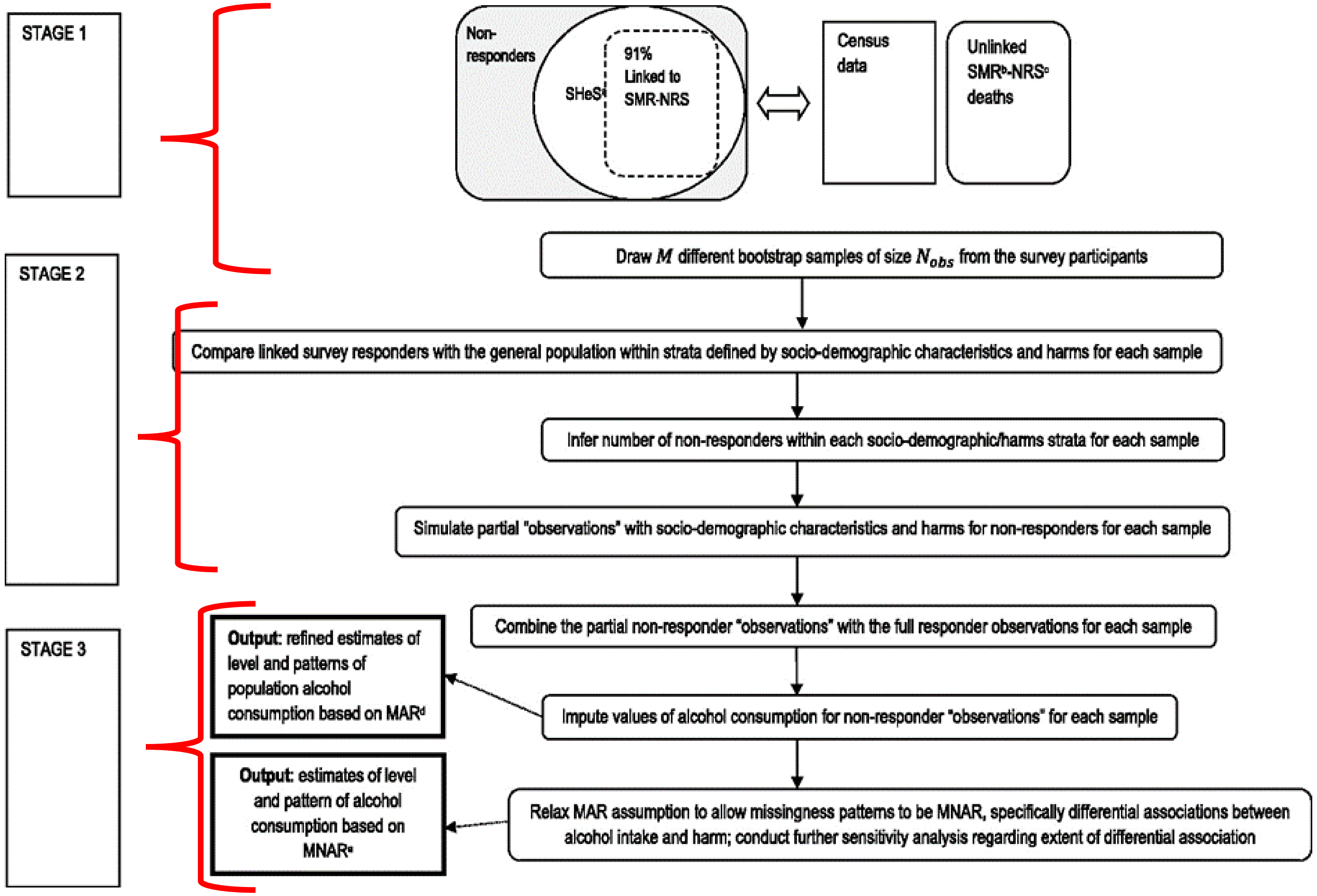
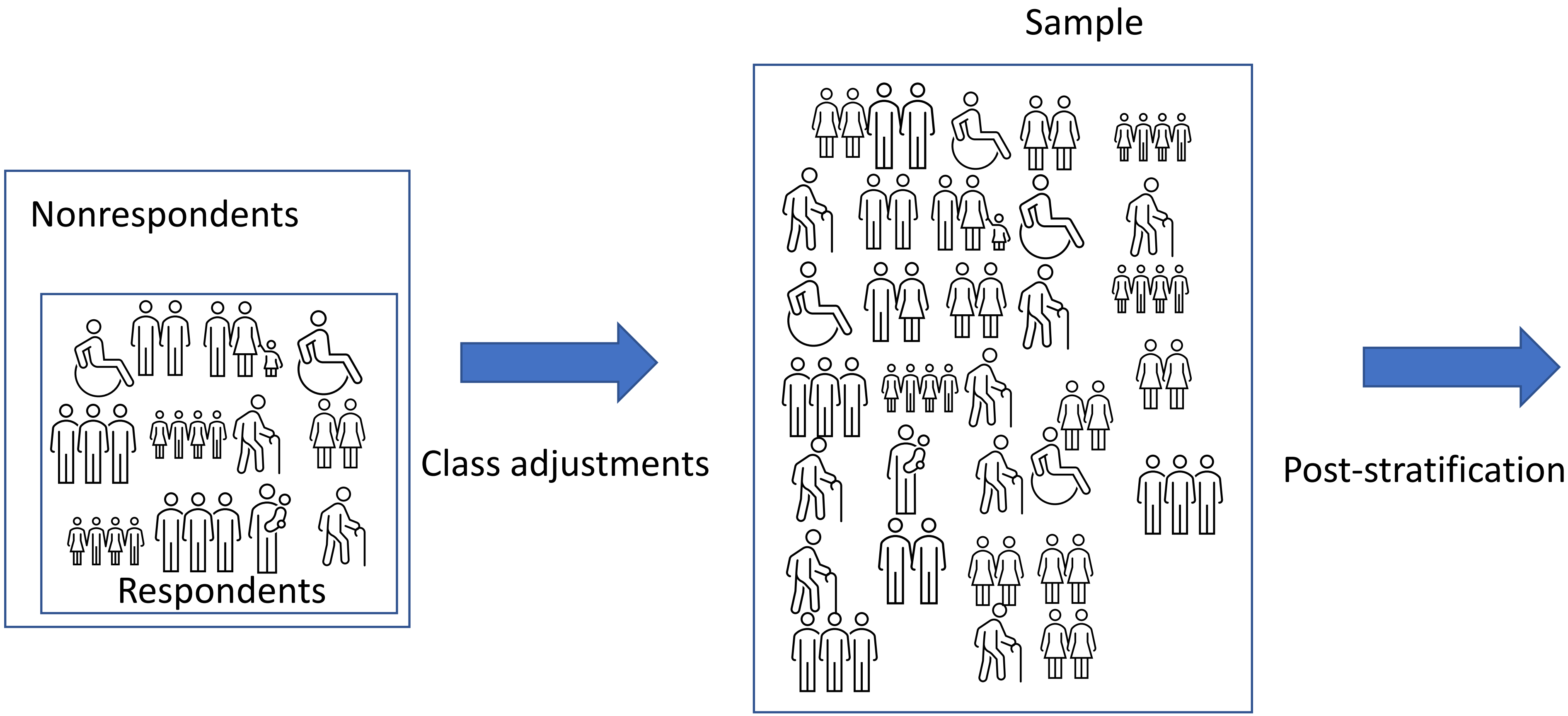
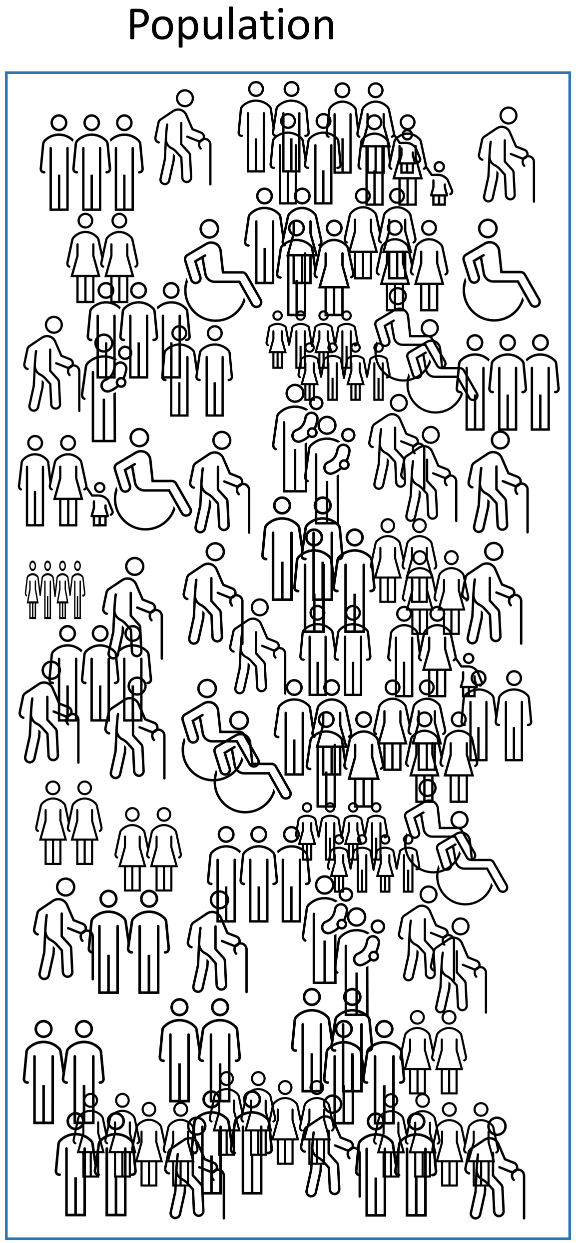


Figure . Summary of methodological strategy for addressing survey non-representativeness and refining alcohol consumption estimates Extracted from: Gray et al. (2019). Correcting for non-participation bias in health surveys using record-linkage, synthetic observations and pattern mixture modelling. Statistical Methods in Medical Research, 29(4), 1212–1226. https://doi.org/10.1177/0962280219854482

Weighting approach in Survey Research





As Brick and Kalton (1996) described, weights are developed in a series of stages to compensate for:

Unequal selection probabilities.

Nonresponse.

Noncoverage.

Sampling fluctuations.

Weighting approach in Survey Research

The auxiliary information can be classified as:

- Design variables.
- Variables used to construct the imputation models and variables also related to the response probabilities.
- Calibration or benchmark variables.

Covariates example

- Age
- Ethnicity
- Marital status
- Gender
- Education level
- Family income
- Employment status
- Health status
- Urban/rural residence

Weighting approach in Survey Research Adjusting for non-response

Weighting Adjustments

- Class Weighting
- Response Propensity Weighting

Approach	Description	Data required	Challenges or Benefits
Sample weighting	SW weight responses within classes so that the profile of respondents across classes is equivalent to the profile of the entire survey sample. The total sample across all weighting classes can be viewed as the first-phase sample. Respondents within the various weighting classes represent the second-phase sample.	Data about the profile of respondents and nonrespondents, but it does not require info of the population.	It cannot compensate for non coverage. One important difference between sample weighting and two-phase sampling designs is that the second-phase elements are selected at random in a two-phase sample, but the respondents in sample weighting are self selected.
Population weighting	The respondent sample is weighted so that the weighted sample distribution is the same as the distribution of the population across classes. Is applied to the scores of individual respondents in order to weight up the respondent sample to the population.	Data about the distribution of the population and the distribution of respondents across weighting classes. Data about the distribution of nonrespondents is not required.	It allows to compensate for noncoverage and nonresponse.
Post-stratification	It is a later stage and consist in adjust sample weights conform to known population values for some key variables. It forces the sample joint distribution of certain variables to match the known population joint distribution.	Population data is required.	It compensate for noncoverage and improve survey estimates. It can also be used to compensate for nonresponse. Poststratification is intended to compensate for minor sampling fluctuations.
Calibration	It means that the weights were made to agree with the known population totals for each margin (Kolenikov, 2016).	Population data is required.	Allows to increase the precision of the population parameters, using the known auxiliary information. It does not attempt to perfectly align with the population as Post-stratification would.

Weighting approach in Survey Research

What that would look like in data set

iew Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

IDENT	Mental	depress	agecat	sex	visits	responded	wgt
1.00	39.00	.00	5.00	1.00	4.00	1.00	5.34
2.00	.	.	3.00	2.00	.00	.00	1.17
3.00	.	.	3.00	2.00	5.00	.00	1.16
4.00	.	.	4.00	2.00	7.00	.00	1.18
5.00	56.00	.00	6.00	2.00	7.00	1.00	4.58
6.00	.	.	2.00	2.00	.00	.00	1.11
7.00	.	.	6.00	2.00	5.00	.00	1.28
8.00	.	.	3.00	2.00	.00	.00	1.17
9.00	45.00	.00	1.00	2.00	2.00	1.00	11.86
10.00	54.00	.00	6.00	2.00	8.00	1.00	4.70
11.00	49.00	.00	5.00	2.00	6.00	1.00	5.27
12.00	36.00	1.00	4.00	2.00	8.00	1.00	6.19

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

survey1

Filter

	IDENT	Mental	depress	agecat	sex	visits	responde	wgt
1	1	39	0	5	1	4	1	5.342116
2	2	NA	NA	3	2	0	0	1.170568
3	3	NA	NA	3	2	5	0	1.157193
4	4	NA	NA	4	2	7	0	1.179748
5	5	56	0	6	2	7	1	4.584825
6	6	NA	NA	2	2	0	0	1.113644
7	7	NA	NA	6	2	5	0	1.278192
8	8	NA	NA	3	2	0	0	1.170568

Showing 1 to 9 of 44,135 entries, 8 total columns

Treating it as a “missing data” problem and using imputation vs Weighthing

Alanya, Wolf, and Sotto (2015) compared multiple imputation (MI) and inverse propensity score weighting (IPSW) in unit-nonresponse adjustments in simulated data under MAR assumptions. To evaluate the effectiveness of MI and PSW to adjust for unit nonresponse bias adjustment. They consider how robust each method is against misspecification, such as omitted interactions and nonlinear terms.

- MI generally yields to lower RMSE when auxiliary variables are strongly associated with the response. However, MI is not consistently better than IPSW.

- MI requires caution, especially while generating global nonresponse weights. It also requires more effort and expertise in model specification.

- MI can handle item missing data in auxiliary variables and unit missing data in one step. However, MI can also be used as a first step before estimating propensity scores to complete missing auxiliary information.

- Regarding to IPSW, the extreme weights result from logistic regression can be addressed by the use of an alternative method such as Generalized Boosted Methods.

Propensity Score

<u>Key points:</u>

- Randomized and Nonrandomized experiments.
- Causal effects.
- Propensity score as a balancing score.
- Average treatment effect (ATE).
- Average treatment effect for the treated (ATT).

<u>Assumptions:</u>

- Strongly ignorable treatment assignment.
- Stable unit treatment value assumption (SUTVA).

Biometrika (1983), 70, 1, *pp.* 41–55
Printed in Great Britain **41**

The central role of the propensity score in observational studies for causal effects

BY PAUL R. ROSENBAUM

Departments of Statistics and Human Oncology, University of Wisconsin, Madison, Wisconsin, U.S.A.

AND DONALD B. RUBIN

University of Chicago, Chicago, Illinois, U.S.A.

SUMMARY

The propensity score is the conditional probability of assignment to a particular treatment given a vector of observed covariates. Both large and small sample theory show that adjustment for the scalar propensity score is sufficient to remove bias due to all observed covariates. Applications include: (i) matched sampling on the univariate propensity score, which is a generalization of discriminant matching, (ii) multivariate adjustment by subclassification on the propensity score where the same subclasses are used to estimate treatment effects for all outcome variables and in all subpopulations, and (iii) visual representation of multivariate covariance adjustment by a two-dimensional plot.

Some key words: Covariance adjustment; Direct adjustment; Discriminant matching; Matched sampling; Nonrandomized study; Standardization; Stratification; Subclassification.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55. https://doi.org/10.1093/biomet/70.1.41

Propensity Score (Definitions and Uses)

A propensity score of response in surveys is essentially the conditional probability that a person or household responds given the covariates (Wun, et.al 2014, p. 4626).

The propensity score is the probability that a particular case would be assigned or exposed to a treatment condition (Ridgeway et al., 2020, p.1).

Response propensities are unknown. In fact, they are **latent variables** and cannot be observed directly – we observe only the binary outcome of response or nonresponse (Brick, 2013 p. 339)

Propensity Score is used for:

- Matching
- Stratify or Subclassify
- Weighting

Propensity Score Weighting Approach

- Alanya, Wolf, and Soto (2015) described PSW as the commonplace and popular model-based survey research technique for adjusting for unit-nonresponse bias.

- *PSW approach computes propensity scores by modeling the probability/likelihood of the response indicator conditional on auxiliary information (e.g., sample frame information,* paradata, or nonresponse surveys), then assigning each respondent a weight that is equal to the inverse of his/her estimated propensity score (Alanya, Wolf, & Soto, 2015, p.636).

Propensity Score Weighting Approach

PSW takes a differential amount of information from each participant depending on the participant's conditional probability of receiving treatment (responding or not to the questionnaire).

The method directly exploits the inverse of estimated propensity scores as weights in an outcome analysis, and to a large extent, it shares similarities with weighted analysis using unequal sampling weights.

Guo, S., & Fraser, M. (2015). *Propensity Score Analysis* (2nd ed.). Sage.

Propensity Score Weighting Approach (Steps)

- Estimate propensity scores
  - Logistic Regression.
  - Generalized Boosted Models.
- Calculate the weights
  - ATE
  - ATT
- Specify the weight in an outcome.
  - Using the weights as sampling weights and becoming a propensity score weighted analysis.
  - Only the respondents group plus their weight is included in outcome analysis.

Guo, S., & Fraser, M. (2015). *Propensity Score Analysis* (2nd ed.). Sage.

Propensity Score Weighting (Estimators)

Gelein et. al (2018, p.3) referes that nonparametric methods are usualy preferred as they protect against the misspecification of the nonresponse model.

Logistic regression is a case of parametric models and some issues are related to the use of a parametric model.

(1) They are not robust to model misspecification,

(2) They are not robust to the non-inclusion of interactions or predictors that account for curvature,

(3) They may yield in very small estimated response probabilities, resulting in very large nonresponse adjustment factors, and in consequence potentially unstable estimates.

Propensity Score Weighting (Estimators)

The use of a Machine Learning Technique to estimate propensity scores has been impulse by McCaffrey et al. (2013) as a strategy that outperform over simple logistic regression models.

Generalized Boosted Model (GBM) estimation involves an iterative process with multiple regression trees to capture complex and nonlinear relationships between treatment assignment and pre-treatment covariates without overfitting the data (McCaffrey et al., 2013 p. 3).

GBM works with a larger number of continuous or discrete covariates and its iterative estimation procedure can be specified to find the propensity score model leading to the optimal balance between groups.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Statistics in medicine, 32(19), 3388–3414. https://doi.org/10.1002/sim.5753

Propensity Score Weighting Package

The toolkit for weighting and analysis of non-equivalent groups (TWANG) is a propensity score R-package developed in 2004 to support causal modeling of observational data by estimating and evaluating propensity scores and associated weights. TWANG uses the GBM approach for the estimation of the propensity score weights (Ridgeway et al., 2020).

PSW has also been extensively used in education, policing and criminal justice, drug treatment evaluation, and military workforce issues by statisticians at RAND Corporation (Ridgeway & McCaffrey, 2007).

For more information visit: https://cran.r-project.org/web/packages/twang/index.html

Propensity Score Weighting Software Package
(Ridgeway et al., 2020).

PSW aims to reduce bias that could potentially result from excluding incomplete cases. This reduction is achieved by <u>weighting complete cases with nonresponse weights and making them look like the entire sample</u> (respondents & nonrespondents).

The package's main workhorse is the ps() function, which implements GBM to estimate the propensity scores. The package aims to:

- <u>Compute</u> from the data estimates of the <u>propensity scores</u>, which yield accurate causal effect estimates,
- <u>Check the quality</u> of the resulting propensity score weights by assessing whether they have the balancing properties expected.
- <u>Use them in computing</u> treatment <u>effects</u>.

Propensity Score Weighting Software Package
(Ridgeway et al., 2020).

- PSW is easy to implement and does not require additional programming efforts.
- In theory, the two types of weights-PSW and sampling weights-are probability-typed quantities, and as such, it is not invalid to incorporate the two types of weights into one by multiplication.

Demonstration (1)

By using simulated data, we aim to compare the results of complete case analysis vs. weighted analysis. We demonstrate the use of Logistic Regression and PSW-GBM for estimating survey weights to answer the following research questions.

We focused on the estimation of the population mean of a continuous variable. Complete case analysis and weighted data analysis were performed.

Demonstration (2)

Accordingly, a hypothetical survey dataset was generated with a sample size of 8,717 where the survey outcome of interest is Y (Mental), and the binary unit-response indicator is Responded (Response=1, Nonresponse=0).

Four auxiliary variables, Sex, Age category, Confidence in Health Providers, and Numbers of hospital visits, were included.

Demonstration (3)

We began by estimating weights by using the Logistic Regression approach on IBM SPSS Software version 23. Descriptive Statistics were calculated with weights from the Logistic Regression using the Complex Survey tool on IBM SPSSv23.

As a second step, ATE weights were calculated using the TWANG Package on R. Later, descriptive statistics were calculated using the ATE weights in R by using Survey package.

The software syntax for each approach is attached to this presentation, and a video demonstration is included.

Demonstration: Propensity Score Weighting – Logistic Regression

1- Locate the **UnitNonresponse** folder provided for this webinar in your “C” drive.

2- You will find a sub-folder named “DemoLG” in which you will see the Dataset and the SPSS syntax needed to run the analysis described in the video tutorial.

3- Open the SPSS syntax and follow each Step of the syntax.

3.1. Open the dataset.

3.2. Estimate the predictive probabilities with Logistic Regression.

3.3. Calculate the inverse of the predictive probabilities from step 3.2.

3.4. Create a new dataset that will contain only participants who responded to the survey and their associated nonresponse weights.

3.5. Use this dataset for outcome analysis in SPSSComplex Samples Analysis.

Demonstration: Propensity Score Weighting – GBM TWANG

1- Locate the **UnitNonresponse** folder provided with this webinar materials in your "C" drive.

2- You will find a sub-folder named "DemoTwang" in which you will see the Dataset and the R syntax needed to run the analysis described in the video tutorial.

3. Open the R Studio and follow each Step of the syntax.

  3.1. Set working directory.

  3.2. Install Packages.

  3.3. Import dataset.

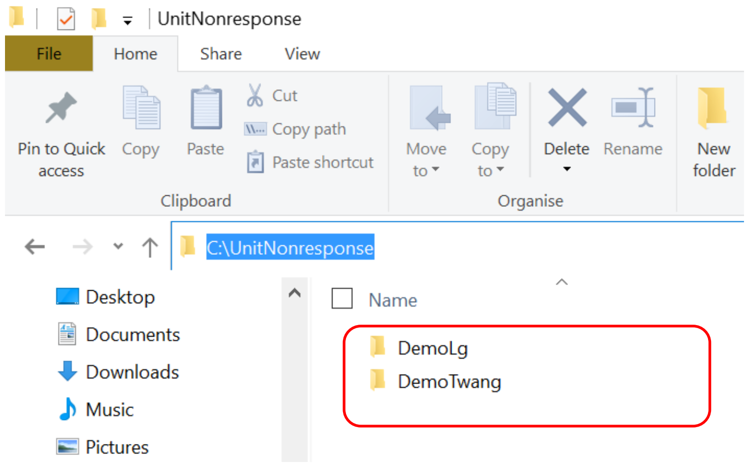
  3.4. Estimate ATE weights.

  3.5. Evaluate balance of the weighted and unweighted sample.

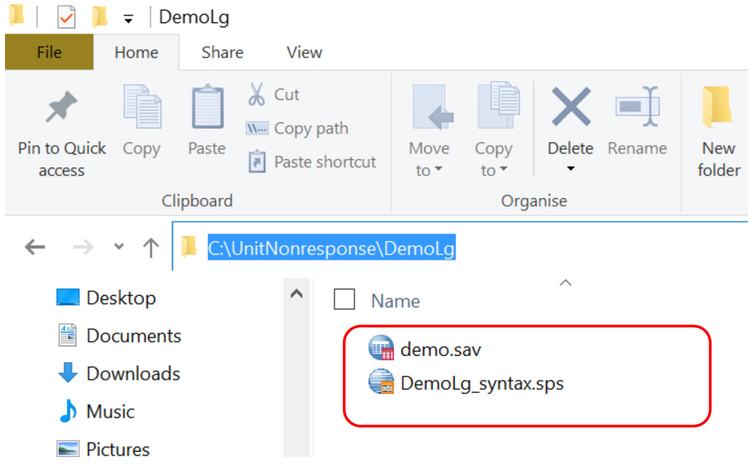
  3.6. Create a new dataset that will contain only participants who responded to the survey and their associated weights.

  3.7. Use this dataset for outcome analysis using Survey package in R.

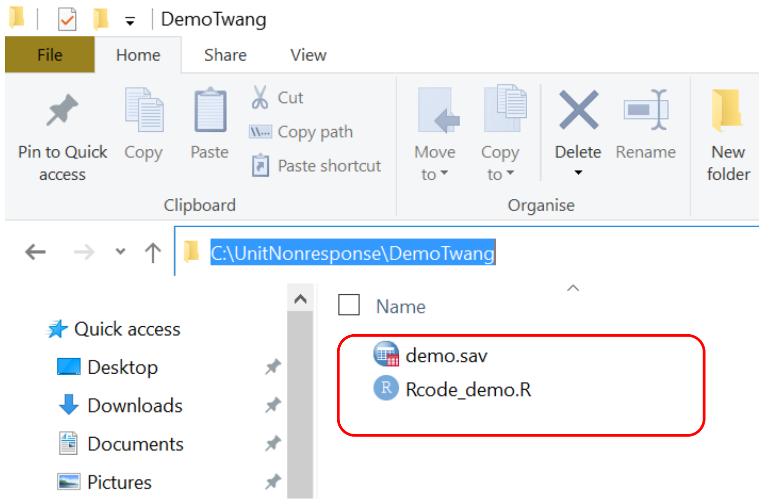
1- Locate the UnitNonresponse folder provided for this webinar in your “C” drive.



2- In sub-folder “DemoLG” you will find the dataset and the SPSS syntax needed to estimate the weights through Logistic Regression.



3- In sub-folder “DemoTwang” you will find the dataset and the R syntax needed to estimate the weights through the TWANG package in R.



Run the PSW analysis in each specific software

PAUSE, please go to the videos for the calculation of the PSW.

1- Video 1: Logistic Regression in IBM SPSS.

2- Video 2: GBM Twang in R.

Demonstration: Results from CCA, LogReg and GBM

	Complete Case analysis – ignoring nonresponse		**Logistic Regression Weights - Complex Survey tool IBM SPSS**		**Twang Weights – Survey package in R**	
	Mean MCS	*Standard error of the mean*	*Mean MCS*	*Standard error of the mean*	*Mean MCS*	*Standard error of the mean*
Sample _Demo	50.827	0.163	48.358	0.204	48.484	0.203

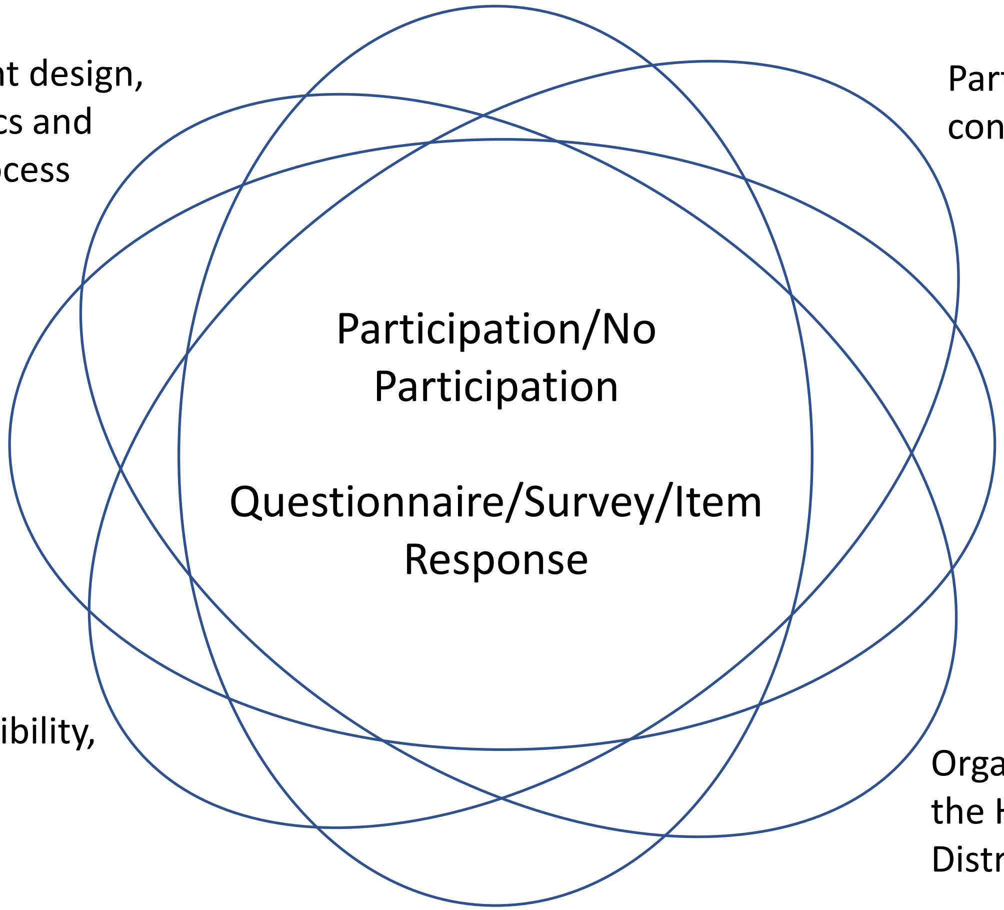
Demonstration: Conclusions

- Weighting for non-response is analogous to the role of sampling weights.
- Is the nonresponse Ignorable or Nonignorable.
- Who has not responded and what do we know (covariates) about them.
- PSW as an individual score and not in cells weighting approaches.
- Model based (Logistic Regression) or machine learning approach (GBM).
- Response mechanism, estimation of the probability of response.
- Remove the nonrespondents from the dataset.
- We are not imputing. We are only working with the respondents+weight.
- SPSS complex design tool and R Survey Package for outcomes analysis.

Framework to understand and conceptualize Unit Nonresponse

Ecological Model of Item and Test Responding (Woitschach, 2018)

- **One instantiation**
- **Ecological**
- **Embedded**



- Zumbo, B.D., Liu, Y., Wu, A.D., Shear, B.R., Astivia, O.L.O. & Ark, T.K. (2015). A Methodology for Zumbo's Third Generation DIF Analyses and the Ecology of Item Responding. Language Assessment Quarterly, 12, 136-151.
- Chen, M.Y., & Zumbo, B.D. (2017). Ecological framework of item responding as validity evidence: An application of multilevel DIF modeling using PISA data. In B. D. Zumbo and A.M. Hubley (Eds.), Understanding and Investigating Response Processes in Validation Research (pp. 53-68). New York, NY: Springer.
- Woitschach, P. (2018). Large-scale assessments in Latin America: Terce. [Doctoral dissertation, Complutense University of Madrid] https://eprints.ucm.es/id/eprint/55315/

Webinar Part I, II and III.

The whole webinar (Part I, II and III) series was prepared to help data analyst and practitioners to understand how:

- Missing data may complicate the interpretation and analysis of survey data.

- Unit nonresponse may alter/distort your conclusions from your analyses, and by how much, and

- To introduce a framework to understand the phenomenon of Unit Nonresponse in a broader view, as well as to demonstrate a workable statistical tool that can go a long way toward adjusting for unit nonresponse.

Thank you

Nonparticipation (Unit Nonresponse) In Surveys: A Practitioner's Guide to the Conceptualization, Impact of, and Adjustment for Unit Nonresponse

Pamela Woitschach, Ph.D.
Post-Doctoral Fellow, Patient-Centred Measurement Methods
Cluster BC SUPPORT Unit
Post-Doctoral Research Fellow, UBC-Paragon Research Initiative
University of British Columbia

Bruno D. Zumbo, Ph.D.
Professor & Distinguished University Scholar
Tier 1, Canada Research Chair in Psychometrics and Measurement; &
Paragon UBC Professor of Psychometrics and Measurement
University of British Columbia

Financial Support from:

bc pcm
patient-centred measurement

BC SUPPORT Unit
Advancing Patient-Oriented Research

The UBC-Paragon Research Initiative and, Professor Bruno D. Zumbo, Paragon UBC Professor of Psychometrics and Measurement

UBC
a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA
Faculty of Education

TESTING ENTERPRISES
Paragon

References

Alanya, A., Wolf, C., & Sotto, C. (2015). Comparing multiple imputation and propensity-score weighting in unit-nonresponse adjustments: A simulation study. *Public Opinion Quarterly, 79*(3), 635-661. https://doi.org/10.1093/poq/nfv029

Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics, 29*(3), 329-353. https://doi.org/10.2478/jos-2013-0026

Brick, J. M., & Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research, 5*, 215-238. https://journals.sagepub.com/doi/pdf/10.1177/096228029600500302

Gelein, B., Haziza, D., & Causeur, D. (2018). Propensity weighting for survey nonresponse through machine learning. *13es Journées de méthodologie statistique de l'Insee Jun 2018*.

Gorman, E., Leyland, A., McCartney, G., White, I. R., Katikireddi, S. V., Rutherford, L., Graham, L., & Gray, L. (2014). Assessing the representativeness of population-sampled health surveys through linkage to administrative data on alcohol-related outcomes. *American Journal of Epidemiology*. https://doi.org/10.1093/aje/kwu207

Gorman, E., Lyland, A., McCartney, G., Katikireddi, S. V., Rutherford, L., Graham, L., & Robinson, J. C. (2017). Adjustment for survey non-representativeness using record-linkage: refined estimates of alcohol consumption by deprivation in Scotland. *Addiction Methods and Techniques, 112*, 1270–1280. https://doi.org/10.1111/add.13797

Gray, L. (2016). The importance of post hoc approaches for overcoming nonresponse and attrition bias in population-sampled studies. *Soc Psychiatry Psychiatr Epidemiol, 51*, 155–157. https://doi.org/10.1007/s00127-015-1153-8

Gray, L., Gorman, E., White, I. R., Katikireddi, S. V., McCartney, G., Rutherford, L., & Leyland, A. (2019). Correcting for non-participation bias in health surveys using record-linkage, synthetic observations and pattern mixture modelling. *Statistical Methods in Medical Research, 29*(4), 1212–1226. https://doi.org/10.1177/0962280219854482

Gray, L., McCartney, G., White, I. R., Katikireddi, S. V., Rutherford, L., Gorman, E., & Lyland, A. (2013). Use of record-linkage to handle non-response and improve alcohol consumption estimates in health survey data: a study protocol. *BMJ Open*. https://doi.org/10.1136/bmjopen-2013002647

References

Guo, S., & Fraser, M. (2015). *Propensity score analysis* (2nd ed.). Sage.

Kolenikov, S. (2016). Post-stratification or non-response adjustment. *Survey Practice, 9*(3), 1-12. https://doi.org/10.29115/SP-2016-0014

Lohr, S., Riddles, M., & Morganstein, D. (2016). Test for evaluating nonresponse bias in surveys. *Survey Methodology, 42*(2), 195-218.

McCaffrey, D., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med, 32*(19), 3388–3414. https://doi.org/10.1002/sim.5753

McMinn, M., Gray, L., Härkänen, T., Tolonen, H., Pitkänen, J., Molaodi, O., Leyland, A., & Martikainen, P. (2020). Alcohol-related outcomes and all cause mortality in the health 2000 survey by participation status and compared with the finnish population. *Epidemiology, 31*(4), 534-541. https://doi.org/10.1097/EDE.0000000000001200

McMinn, M., Martikainen, P., Gorman, E., Rissanen, H., Härkänen, T., Tolonen, H., Leyland, A., & Gray, L. (2018). Validation of non-participation bias methodology based on record-linked Finnish register-based health survey data: a protocol paper. *BMJ Open*. https://doi.org/10.1136/ bmjopen-2018-026187

Peytchev, A. (2012). Multiple imputation for unit nonresponse and measurement error. *The Public Opinion Quarterly, 76*(2), 214-237. https://doi.org/10.1093/poq/nfr065

Pike, G. (2007). Adjusting for nonresponse in surveys. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. XXII, pp. 441-449). Springer. https://doi.org/https://doi.org/10.1007/978-1-4020-5666-6_8

Ridgeway, G., & McCaffrey, D. (2007). Demystifying double robustness; A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science, 22*(4), 540-543. https://doi.org/10.1214/07-STS227

Ridgeway, G., McCaffrey, D., Morral, A., Burgette, L., & Griffin, B. A. (2020). *Toolkit for weighting and analysis of nonequivalent groups: A tutorial for the twang package*. RAND.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55. https://doi.org/10.1093/biomet/70.1.41

Statistics Canada. (2020). Statistics Canada quality guidelines: Response and nonresponse. https://www150.statcan.gc.ca/n1/pub/12-539-x/2009001/response-reponse-eng.htm

Woitschach, P. (2018). *Large scale educational assessments in Latin America: TERCE* Complutense University of Madrid]. https://eprints.ucm.es/55315/

Wun, L. M., Ezzati-Rice, T., Baskin, R., Greenblatt, J., & Zodet, M. (2014). Using propensity score to adjust weights to compensate for dwelling unit level nonresponse in the medical expenditure panel survey. ASA Section on Survey Research Methods,