UNIVERSITY OF Southampton

Southampton Statistical Sciences Research Institute

Theoretical Properties of Partial Indicators for Representative Response

Version 2: January 2013

Natalie Shlomo
University of Southampton, United Kingdom

Barry Schouten
Centraal Bureau voor de Statistiek, Netherlands

Theoretical Properties of Partial Indicators for Representative Response

1. Introduction

The RISQ project (Representativity Indicators for Survey Quality), funded by the European 7th Framework Programme (FP7), was a joint effort of the NSI's of Norway, the Netherlands and Slovenia, and the Universities of Leuven and Southampton to develop quality indicators for survey response. These indicators measure the degree to which the group of respondents of a survey resembles the complete sample. When this is the case, the response is called representative. We measure the contrast between respondents and the full sample based on auxiliary information that is known to both respondents and non-respondents in the survey. It was the objective of the RISQ project to translate auxiliary information to Representativity Indicators, to develop these quality indicators, to explore their characteristics and to show how to implement and use them in a practical data collection environment.

Two types of indicators were developed under the RISQ project: R-indicators and partial R-indicators. R-indicators provide a single value between zero and one that measures the closeness to representative response. Representativity is defined in terms of the response propensities of different sample units given their values on a specified set of auxiliary variables. Response is said to be representative if all the response propensities in the sample are equal (and none are equal to zero). Our definitions of R-indicators will be most effective in capturing non-response bias in a survey estimate when the auxiliary variables are, in combination, strong predictors of the survey item(s) upon which the estimate is based.

Partial R-indicators are defined in terms of a single specified auxiliary variable and in terms of the categories of this variable when it is categorical. They are designed to measure the impact of the specified variable on deviations from representative response. We also make a distinction between unconditional and conditional partial R-indicators.

The aim of this paper is to investigate the statistical properties of these partial R-indicators. More information on the definition of the partial R-indicators and their use in different settings are in Deliverable 4 from the RISQ project by Shlomo, Skinner, Schouten, Carolina and Morren (2009) and in the journal article by Schouten, Shlomo and Skinner (2011).

Following the end of the RISQ project in May 2010, a further research contract was set up between Statistics Netherlands and the University of Southampton to investigate the theoretical properties of the partial R-indicators. In particular, we examine if a bias correction is necessary for the partial R-indicator and develop analytical expressions for their variance.

Section 2 provides a brief summary of the definitions of R-indicators and corresponding partial R- indicators as published in the deliverables of the RISQ project and subsequent journal articles. Section 3 covers statistical properties of the partial R-indicators. Section 3.1 describes two proposed methods for bias corrections and Section 3.2 provides theoretical analytical expressions for estimates of the variance at the categorical level of the partial R-Indicators. The performance of the proposed bias corrections and analytical expressions of the variance are presented in Section 4 through a simulation study based on real samples drawn from an extract of the 1995 Israel Census microdata. Section 5 contains a conclusion, recommendations and further work.

2. Definition of R- indicators and Partial R-Indicators

We use the notation and definition of response propensities as set out in the previous RISQ deliverable Shlomo, Skinner, Schouten, Bethlehem, Zhang (2008), and the journal paper in Schouten, Cobben and Bethlehem (2009) and Shlomo, Skinner and Schouten (2012). We let U denote the set of units in the population and s the set of units in the sample. We define a response indicator variable R_i which takes the value 1 if unit i in the population responds and the value 0 otherwise. The *response propensity* is defined as the conditional expectation of R_i given the vector of values x_i of the vector X of auxiliary variables: $\rho_X(x_i) = E(R_i = 1 | X = x_i) = P(R_i = 1 | X = x_i)$ and also denote this response propensity by ρ_X. We assume that the values x_i are known for all sample units, i.e. for both respondents and non-respondents, and can include both specified variables and survey fieldwork conditions.

2.1 Definition of R-Indicator

We define the R-indicator as: $R(\rho_X) = 1 - 2S(\rho_X)$. The estimation of the propensities is typically based on a logistic regression model: $\log[\rho_X / (1-\rho_X)] = x'\beta$ where β is a vector of unknown parameters to be estimated, and x may involve the transformation of the original auxiliary variables (e.g. by including interaction terms) for the purpose of model specification. The estimator of the response propensity is: $\hat{\rho}_X = \dfrac{\exp(x'\hat{\beta})}{\exp(x'\hat{\beta}) + 1}$ where $\hat{\beta}$ is the estimator of β based on the model. The estimator of the variance of the response propensities: $\hat{S}^2(\hat{\rho}_X) = \dfrac{1}{N-1}\sum_s d_i(\hat{\rho}_X(x_i) - \hat{\bar{\rho}}_X)^2$ where $d_i = \pi_i^{-1}$ is the design weight or inclusion weight and $\hat{\bar{\rho}}_X = \dfrac{1}{N}\sum_s d_i \hat{\rho}_X(x_i)$. We estimate the R-indicator: $\hat{R}(\hat{\rho}_X) = 1 - 2\hat{S}(\hat{\rho}_X)$.

2.2 Definitions of partial R indicators

In this section we define the unconditional and conditional partial indicators.

Unconditional partial indicators measure the distance to representative response for single auxiliary variables and are based on the between variance given a stratification with categories of Z. The variable Z may or may not be included in the covariates of the model X for estimating the response propensities.

Given a stratification based on a categorical variable Z having categories $k = 1,2,\dots,K$, the variable level unconditional partial R-indicator is defined as $P_u(Z,\rho_X) = S_B(\rho_X \mid Z)$ where $S_B^2(\rho_X \mid Z) = \frac{1}{N-1}\sum_{k=1}^{K} N_k(\bar{\rho}_{X,k} - \bar{\rho}_X)^2 \cong \sum_{k=1}^{K}\frac{N_k}{N}(\bar{\rho}_{X,k} - \bar{\rho}_X)^2$ where $\bar{\rho}_{X,k}$ is the average of the response propensity in stratum k. This between variance is estimated by: $\hat{S}_B^2(\hat{\rho}_X \mid Z) = \sum_{k=1}^{K}\frac{\hat{N}_k}{N}(\hat{\bar{\rho}}_{X,k} - \hat{\bar{\rho}}_X)^2$ where $\hat{\bar{\rho}}_{X,k} = \frac{1}{N_k}\sum_{s_k} d_i \rho_X(x_i)$, $\hat{N}_k = \sum_{s_k} d_i$ is the estimated population size of stratum k and s_k is the set of sample units in the stratum.

At the category level *Z=k,* the unconditional partial indicator is defined as: $P_u(Z,k,\rho_X) = S_B(\rho_X \mid Z=k)\frac{(\bar{\rho}_{X,k} - \bar{\rho}_X)}{|\bar{\rho}_{X,k} - \bar{\rho}_X|} = \sqrt{\frac{N_k}{N}}(\bar{\rho}_{X,k} - \bar{\rho}_X)$ and is estimated by: $\hat{S}_B(\hat{\rho}_X \mid Z=k) = \sqrt{\frac{\hat{N}_k}{N}}(\hat{\bar{\rho}}_{X,k} - \hat{\bar{\rho}}_X)$.

Conditional partial R- indicators measure the remaining variance due to variable *Z* within sub-groups formed by all other remaining variables, denoted by X^-. In contrast to the unconditional partial R- indicator, the variable Z must be included in the model for estimating response propensities.

Let δ_k be the 0-1 dummy variable that is equal to 1 if $Z = k$ and 0 otherwise. Given a stratification based on all categorical variables except Z, denoted by X^- and indexed by *j*, *j=1...J*, the conditional partial R-indicator is based on the within variance and is defined as: $P_c(Z,\rho_X) = S_W(\rho_X \mid X^-)$ where $S_w^2(\rho_X \mid X^-) = \frac{1}{N-1}\sum_{j=1}^{J}\sum_{i\in U_j}(\rho_X(x_i) - \bar{\rho}_{X,j})^2$ and is estimated by $\hat{S}_w^2(\hat{\rho}_X \mid X^-) = \frac{1}{N-1}\sum_{j=1}^{J}\sum_{i\in s_j} d_i(\hat{\rho}_X(x_i) - \hat{\bar{\rho}}_{X,j})^2$.

At the categorical level of $Z=k$, we restrict the within variance to population units in stratum k and obtain: $P_c(Z,k,\rho_X)=\sqrt{\frac{1}{N-1}\sum_{j=1}^{J}\sum_{i\in U_j}\delta_{k,i}(\rho_X(x_i),-\bar{\rho}_{X,j})^2}$ and estimated by: $\hat{P}_c(Z,k,\hat{\rho}_X)=\sqrt{\frac{1}{N-1}\sum_{j=1}^{J}\sum_{i\in s_j}d_i\delta_{k,i}(\hat{\rho}_X(x_i),-\hat{\bar{\rho}}_{X,j})^2}$.

3. Statistical Properties of Partial indicators

3.1 Bias Adjustments

As shown in Shlomo, Skinner, Schouten, Bethlehem and Zhang (2008) and Shlomo, Schouten, Skinner (2012), estimated R-indicators have a sample size dependent bias. When the sample size decreases, the bias increases. For this reason a bias adjustment was proposed for $\hat{R}(\hat{\rho}_X)$. When the sampling design is a simple random sample without replacement the bias-adjusted R-indicator has the form:

$$\hat{R}_B(\hat{\rho}_X)=1-2\sqrt{(1+\frac{1}{n}-\frac{1}{N})\hat{S}^2(\hat{\rho}_X)-\frac{1}{n}\sum_{i\in s}z_i^T\left[\sum_{j\in s}z_jx_j^T\right]^{-1}z_i}\ , \tag{1}$$

with $z_i=\nabla h(x_i^T\hat{\beta})x_i$ and h the link function in the model for response propensities. For linear regression, h is the linear function and for a logistic regression it is the logit function.

From the observation that the R-indicator is biased for small sample sizes, we can conclude directly that the proposed partial R-indicators will also be biased since they are based on the same variance or components of that variance. Hence, a bias adjustment would be needed to avoid false conclusions about the impact of single variables. We first investigate the bias of the variable level conditional and unconditional partial R-indicators.

Variable Level Partial R-indicators:

We propose two simple and pragmatic approaches to adjust the bias for the estimates of the variable level partial R-indicators and compare the two methods in a simulation study in Section 4.

Method 1: The R-indicators $\hat{R}_B(\hat{\rho}_X)$ are based on a bias adjusted variance of the response propensities as shown in (1). The variable level unconditional partial R-indicator $P_u(Z,\rho_X)$ is the between variance given the stratifying variable *Z*. The variable level conditional partial R-indicator $P_c(Z,\rho_X)$ is the within variance given the stratifying variable X^- (all auxiliary variables except Z). By calculating the complementary between and within variance for each of the stratifying variables, we implement a heuristic of pro-rating the overall bias correction of the variance of estimated response propensities that is shown in (1). We pro-rate the bias correction term between the decomposed variance components which define the respective partial R-indicators and obtain bias corrections at the variable level for both partial R-indicators.

Method 2: The unconditional partial R-indicator $P_u(Z,\rho_X)$ is the decomposed between variance based on a stratification on variable *Z* of the overall variance of the response propensities. Therefore, the bias correction for the unconditional partial R-indicator would be the same as the bias correction for the overall R-indicator in (1) if we use only the auxiliary variable Z in the logistic regression model to estimate the response propensities. Similarly, the conditional partial R-indicator $P_c(Z,\rho_X)$ is the decomposed within variance based on a stratification on variables X^- of the overall variance of the response propensities. Therefore, the bias correction would be approximately the same as the bias correction proposed for the overall R-indicator in (1) if we use only the auxiliary variables X^- in the logistic regression model to estimate the response propensities. The bias correction would be exactly the same only in the case of a saturated model on X^- to estimate the response propensities. Since we generally are using main effects models to estimate the response propensities, we obtain an approximation of the bias correction for the conditional partial R-indicator.

Categorical Level Partial R-indicators:

For the categorical level conditional and unconditional partial R-indicators, we use the heuristic of pro-rating as described in Method 1. We pro-rate the variable level bias correction term across the categorical level partial R- indicators and assess their performance in the simulation study in Section 4.

3.2 Confidence Intervals of Partial indicators

Since partial R- indicators are random variables, they will have a certain precision that depends on the size of the sample. Through the calculation of standard errors and confidence intervals, we can develop statistical tests for the significance of variables contributing to the representativity of the sample. In this section, we provide approximate analytical expressions for the variance estimates at the categorical level of the partial R-indicators: $P_u(Z,k,\rho_X)$ and $P_c(Z,k,\rho_{X,Z})$. The analytical expressions of standard errors for the variable level partial R-indicators involves more complexity and possible covariance terms and therefore is a topic for future research.

Let X^- be the auxiliary variables, taking values $j=1,2,...,J$ and Z a categorical variable for which the partial indicator is calculated with categories $k=1,2,\ldots,K$.

Approximate analytical expression of the standard error for the unconditional category level partial R-indicator:

The variance of the estimated unconditional partial R-Indicator: $\hat{P}_u(Z,k,\hat{\rho}_X)=\sqrt{\frac{\hat{N}_k}{N}}\left(\hat{\bar{\rho}}_{X,k}-\hat{\bar{\rho}}_X\right)$ can be written as:

$$Var(\hat{P}_u(Z,k,\hat{\rho}_X)) = \frac{\hat{N}_k}{N} Var\left(\hat{\bar{\rho}}_{X,k} - \hat{\bar{\rho}}_X\right) = \frac{\hat{N}_k}{N}\left[Var(\hat{\bar{\rho}}_{X,k}) + Var(\hat{\bar{\rho}}_X) - 2Cov(\hat{\bar{\rho}}_{X,k}, \hat{\bar{\rho}}_X)\right]$$

assuming that N_k is the number of units with Z=k and is known, $\hat{\bar{\rho}}_{X,k} = \sum_{i \in s} d_i \hat{\rho}_i \delta_i^k / \hat{N}_k$ where $\delta_i^k = 1$ if $Z = k$ and $\delta_i^k = 0$ otherwise, and $\hat{\bar{\rho}}_X = \sum_{i \in s} d_i \hat{\rho}_i / N$. In general N_k may not be known and we may need to estimate it by the sample-based estimator $\hat{N}_k = \sum_{s_k} d_i$. This will introduce a small additional loss of precision.

Since $\hat{\bar{\rho}}_X = \frac{\hat{N}_k}{N} \hat{\bar{\rho}}_{X,k} + \left(1 - \frac{\hat{N}_k}{N}\right) \hat{\bar{\rho}}_{X,k^C}$ where $\hat{\bar{\rho}}_{X,k^C} = \sum_{i \in s} d_i \hat{\rho}_i (1 - \delta_i^k)/(N - \hat{N}_k)$ we have that: $Cov(\hat{\bar{\rho}}_{X,k}, \hat{\bar{\rho}}_X) = \frac{\hat{N}_k}{N} Var(\hat{\bar{\rho}}_{X,k})$

and

$$Var(\hat{P}_u(Z,k,\hat{\rho}_X)) = \frac{\hat{N}_k}{N}\left[Var(\hat{\bar{\rho}}_{X,k})(1 - \frac{2\hat{N}_k}{N}) + Var(\hat{\bar{\rho}}_X)\right] \quad (2)$$

In addition, $Var(\hat{\bar{\rho}}_X) = \left(\frac{\hat{N}_k}{N}\right)^2 Var(\hat{\bar{\rho}}_{X,k}) + \left(1 - \frac{\hat{N}_k}{N}\right)^2 Var(\hat{\bar{\rho}}_{X,k^C})$ and we obtain:

$$Var(\hat{P}_u(Z,k,\hat{\rho}_X)) = \frac{\hat{N}_k}{N}\left[\left(1 - \frac{\hat{N}_k}{N}\right)^2 Var(\hat{\bar{\rho}}_{X,k}) + \left(1 - \frac{\hat{N}_k}{N}\right)^2 Var(\hat{\bar{\rho}}_{X,k^C})\right] \quad (3)$$

We restrict ourselves to a first-order approximation and approximate $Var(\hat{\bar{\rho}}_{X,k})$ by a standard design based variance estimator $\sum_{i \in s} d_i \hat{\phi}_i$, where $\hat{\phi}_i = \delta_i^k \hat{\rho}_i / \hat{N}_k$ and approximate $Var(\hat{\bar{\rho}}_{X,k^C})$ by a standard design based variance estimator $\sum_{i \in s} d_i \hat{\upsilon}_i$, where $\hat{\upsilon}_i = (1 - \delta_i^k)\hat{\rho}_i / (N - \hat{N}_k)$. The standard error is obtained by taking the square root of the expression in (3).

Approximate analytical expression of the standard error for the conditional category level partial R-indicator:

Early work on the analytical expression for the standard error of the categorical level conditional partial R-indicator suggested that a second-order term was needed. Therefore, we review how the variance is calculated for the overall R-indicator as shown in Shlomo, Skinner, Schouten, Bethlehem and Zhang (2008) and Shlomo, Schouten, Skinner (2012) and adapt it to the case of the conditional partial R-indicator which is stratified by the auxiliary variables X^{-}.

To approximate $\text{var}(\hat{S}_{\rho}^{2})$ in the variance calculation of the overall R-indicator $\hat{R}(\hat{\rho}_X)$, we decompose $\hat{S}_{\rho}^{2}$ into the part induced by the sampling design for a fixed value of $\hat{\boldsymbol{\beta}}$ and the part induced by the distribution of $\hat{\boldsymbol{\beta}}$. We take the latter to be $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \boldsymbol{\Sigma})$, where:

$$\boldsymbol{\Sigma} = \mathbf{J}(\boldsymbol{\beta})^{-1} \text{var}\{\textstyle\sum_s d_i [R_i - h(\mathbf{x}_i{}'\boldsymbol{\beta})]\mathbf{x}_i\}\mathbf{J}(\boldsymbol{\beta})^{-1} \tag{4}$$

and $\mathbf{J}(\boldsymbol{\beta}) = E\{\mathbf{I}(\boldsymbol{\beta})\}$ is the expected information rather than the observed information in (4) and write:

$$\text{var}(\hat{S}_{\rho}^{2}) = E_{\hat{\boldsymbol{\beta}}}[\text{var}_s(\hat{S}_{\rho}^{2})] + \text{var}_{\hat{\boldsymbol{\beta}}}[E_s(\hat{S}_{\rho}^{2})], \tag{5}$$

where the subscript $\hat{\boldsymbol{\beta}}$ denotes the distribution induced by $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \boldsymbol{\Sigma})$, which may be interpreted as arising from the response process. Following usual linearization arguments we obtain:

$$\text{var}_s(\hat{S}_{\rho}^{2}) \approx \text{var}_s[N^{-1}\sum_{i \in s} d_i(\rho_i - \bar{\rho}_U)^2]\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}$$

where we denote $\rho_i \equiv \rho(x_i)$ and, given the consistency of $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ (and for standard kinds of sampling designs), we have approximately:

$$E_{\hat{\boldsymbol{\beta}}}[\text{var}_s(\hat{S}_{\rho}^{2})] \approx \text{var}_s[N^{-1}\sum_{i \in s} d_i(\rho_i - \bar{\rho}_U)^2]. \tag{6}$$

Turning to the second component in (5), we write:

$$E_s(\hat{S}_{\rho}^{2}) \approx N^{-1}\sum_{i \in U}(\rho_i - \bar{\rho}_U)^2\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}.$$

As a linear approximation we have $\hat{\rho}_i \approx \rho_i + \mathbf{z}_i'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ where $\mathbf{z}_i = \nabla h(\mathbf{x_i}'\boldsymbol{\beta})\mathbf{x}_i$. Hence

$$\sum_{i\in U}(\rho_i - \bar{\rho}_U)^2\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \approx \sum_{i\in U}(\rho_i - \bar{\rho}_U)^2 + 2\sum_{i\in U}(\rho_i - \bar{\rho}_U)(\mathbf{z}_i - \bar{\mathbf{z}}_U)'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$$
$$+ \sum_{i\in U}(\mathbf{z}_i - \bar{\mathbf{z}}_U)'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\mathbf{z}_i - \bar{\mathbf{z}}_U)$$

where $\bar{\mathbf{z}}_U = N^{-1}\sum_U \mathbf{z}_i$.

In large samples, we assume that $\hat{\boldsymbol{\beta}}$ is normally distributed so that $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is uncorrelated with $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'$. Hence, we have

$$\text{var}_{\hat{\boldsymbol{\beta}}}[E_s(\hat{S}_\rho^2)] \approx 4\mathbf{A}'\boldsymbol{\Sigma}\mathbf{A} + \text{var}_{\hat{\boldsymbol{\beta}}}\{tr[\mathbf{B}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})']\}, \tag{7}$$

where $\mathbf{A} = N^{-1}\sum_{i\in U}(\rho_i - \bar{\rho}_U)(\mathbf{z}_i - \bar{\mathbf{z}}_U)$, $\mathbf{B} = N^{-1}\sum_{i\in U}(\mathbf{z}_i - \bar{\mathbf{z}}_U)(\mathbf{z}_i - \bar{\mathbf{z}}_U)'$ and $\boldsymbol{\Sigma}$ is defined in (4). The second term involves the fourth moments of $\hat{\boldsymbol{\beta}}$ which can also be expressed in terms of $\boldsymbol{\Sigma}$ since $\hat{\boldsymbol{\beta}}$ is assumed normally distributed.

The variance of $\hat{S}_\rho^2$ can be estimated by the sum of the estimated components of (5). The first of these appears in (6) and can be estimated by a standard design-based estimator of $\text{var}_s[\sum_{i\in s} d_i(\rho_i - \bar{\rho}_U)^2]$, where this is treated as the variance of a linear statistic $\text{var}_s[\sum_{i\in s} u_i]$ and u_i is replaced by $d_i(\hat{\rho}_i - \hat{\bar{\rho}}_U)^2$ in the expression for the variance estimator. The second component of the variance appears in (7). To estimate this term requires estimating $\mathbf{A}$, $\mathbf{B}$ and $\boldsymbol{\Sigma}$. First, $\mathbf{z}_i$ may be estimated by $\hat{\mathbf{z}}_i = \nabla h(\mathbf{x}_i'\hat{\boldsymbol{\beta}})\mathbf{x}_i$. Then $\mathbf{A}$ may be estimated by $\hat{\mathbf{A}} = N^{-1}\sum_{i\in s} d_i(\hat{\rho}_i - \hat{\bar{\rho}}_U)(\hat{\mathbf{z}}_i - \hat{\bar{\mathbf{z}}}_U)$, $\mathbf{B}$ may be estimated by $\hat{\mathbf{B}} = N^{-1}\sum_{i\in s} d_i(\hat{\mathbf{z}}_i - \hat{\bar{\mathbf{z}}}_U)(\hat{\mathbf{z}}_i - \hat{\bar{\mathbf{z}}}_U)'$, where $\hat{\bar{\mathbf{z}}}_U = N^{-1}\sum_s d_i\hat{\mathbf{z}}_i$, and $\boldsymbol{\Sigma}$ may be estimated by a standard estimator of the covariance matrix of $\hat{\boldsymbol{\beta}}$.

Now for the conditional categorical level partial R-Indicator $P_c(Z,k,\rho_X) = \sqrt{\frac{1}{N-1}\sum_{j=1}^{J}\sum_{U_k}\delta_{k,i}(\rho_X(x_i), -\bar{\rho}_{X,j})^2}$ we use the same methodology as explained above but we add in the stratification variable as follows:

The first term in (6) may be estimated by a standard design-based estimator of $\text{var}_s[\sum_{j=1}^{J}\sum_{i\in s_k} d_i(\hat{\rho}_i - \hat{\bar{\rho}}_{X=j})^2]$ where this is treated as the variance under a stratified sample design of a linear statistic $\text{var}_s[\sum_{j=1}^{J}\sum_{i\in s_k} u_{ji}]$ and u_{ji} is replaced by $d_i(\hat{\rho}_i - \hat{\bar{\rho}}_{X=j})^2$ in the expression for the variance estimator. For the second term in (7), we replace the **A** and **B** under a stratified design by: $\mathbf{A} = N^{-1}\sum_{j=1}^{J}\sum_{U_k} \delta_i^k(\rho_i - \bar{\rho}_{X=j})(z_i - \bar{z}_{X=j})$ and $\mathbf{B} = N^{-1}\sum_{j=1}^{J}\sum_{U_k} \delta_i^k(z_i - \bar{z}_{X=j})(z_i - \bar{z}_{X=j})'$ and estimated by: $\hat{\mathbf{A}} = N^{-1}\sum_{j=1}^{J}\sum_{s_k} d_i\delta_i^k(\hat{\rho}_i - \hat{\bar{\rho}}_{X=j})(\hat{z}_i - \hat{\bar{z}}_{X=j})$ and $\hat{\mathbf{B}} = N^{-1}\sum_{j=1}^{J}\sum_{s_k} d_i\delta_i^k(\hat{z}_i - \hat{\bar{z}}_{X=j})(\hat{z}_i - \hat{\bar{z}}_{X=j})'$.

To assess the performance of the analytical expressions of the variance of the category level partial R-indicators, we carry out a simulation study described in Section 4.

4. Simulation Study

In this section we investigate the theoretical properties of the partial R-indicators as shown in Section 3 using simulated survey data. The aim of the simulation is to analyze the effectiveness of the two methods for bias adjustments, the analytical expressions of the variance of the partial R-indicators and assess the dependence of these properties on the sample size.

For the simulation study, we use a dataset from the 1995 Israel Census Sample of Individuals aged 15 and over (N=753,711). Population response propensities were calculated using a 2-step process:

1. Probabilities of response were defined according to variables: child indicator, income from earnings groups, age group, sex, number of persons and locality type.
2. Using the response indicator as the dependent variable, we fit a logistic regression model on the population using the above explanatory variables. The predictions from this model serve as the 'true' response propensities for our simulation.

The overall response rate generated in the simulated dataset was 69.2%. Table 1 presents the response rates for the variables defining the population response propensities using the logistic regression model. High non-response rates in categories are likely to cause the sub-group in the population to be under-represented according to the partial R-indicators.

Table 1: Percent response generated in the simulation population dataset according to auxiliary variables

Variable	Category	Percent Response
Children	None	67.3
	1+	75.0
Type of Locality	Large cities	64.4
	Type 1 cities	70.2
	Type 1 towns	69.1
	Type 2 cities	73.6
	Type 2 towns	72.7
Age group	15-17	79.8
	18-21	61.8
	22-24	59.2
	25-34	62.9
	35-44	68.1
	45-54	72.8
	55-64	71.3
	65-74	77.4
	75+	82.0
Persons in Household	1	69.2
	2	63.9
	3	75.2
	4	79.4
	5	65.9
	6+	59.4

From this population, we draw 1000 samples under three sample fractions: 1:50 (sample size is 15,074), 1:100 (sample size is 7,537) and 1:200 (sample size is 3,768) using simple random sampling.

We begin with examining the two methods for bias corrections described in Section 3.1: method 1 based on prorating the bias correction term calculated for the overall R-indicator across the decomposed between and within variances, and method 2 based on adapting the response model according to the respective stratification variables and calculating the overall bias correction term in (1) according to the revised response model. We denote method 1 by the 'prorating model' and method 2 by the 'proposed model'. Tables 2 through 4 presents the unadjusted and the two bias correction methods for the variable level conditional partial R-indicator under the different sample sizes: 1:50, 1:100 and 1:200.

Table 2: Variable level conditional partial R-Indicator under the 1:50 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0614	0.0614	0.0035	0.13%	0.0609	0.0036	-0.78%	0.0607	0.0035	-1.12%
Locality Type	0.0370	0.0374	0.0037	1.25%	0.0367	0.0038	-0.75%	0.0370	0.0037	-0.01%
Age Group	0.0667	0.0673	0.0037	0.85%	0.0665	0.0038	-0.34%	0.0664	0.0037	-0.41%
Child in HH	0.0120	0.0115	0.0032	-4.03%	0.0109	0.0034	-9.37%	0.0114	0.0032	-5.22%
Avg. Rel. Error				0.90%			1.19%			0.90%
RRMSE				2.15%			4.72%			2.68%
RMSE				0.0004			0.0006			0.0005

Table 3: Variable level conditional partial R-Indicator under the 1:100 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0614	0.0622	0.0048	1.34%	0.0611	0.0049	-0.47%	0.0607	0.0048	-1.14%
Locality Type	0.0370	0.0379	0.0053	2.59%	0.0364	0.0056	-1.44%	0.0370	0.0053	0.10%
Age Group	0.0667	0.0678	0.0051	1.62%	0.0662	0.0053	-0.76%	0.0661	0.0051	-0.86%
Child in HH	0.0120	0.0118	0.0045	-1.97%	0.0106	0.0048	-11.91%	0.0115	0.0044	-4.36%
Avg. Rel. Error				1.75%			1.56%			1.03%
RRMSE				1.94%			6.01%			2.30%
RMSE				0.0008			0.0008			0.0005

Table 4: Variable level conditional partial R-Indicator under the 1:200 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0614	0.0624	0.0067	1.70%	0.0602	0.0070	-1.95%	0.0594	0.0067	-3.13%
Locality Type	0.0370	0.0389	0.0072	5.22%	0.0359	0.0079	-2.93%	0.0370	0.0070	0.23%
Age Group	0.0667	0.0688	0.0068	3.22%	0.0657	0.0072	-1.53%	0.0656	0.0068	-1.69%
Child in HH	0.0120	0.0117	0.0059	-2.58%	0.0099	0.0058	-17.27%	0.0111	0.0057	-7.22%
Avg. Rel. Error				3.07%			3.03%			2.26%
RRMSE				3.44%			8.84%			4.03%
RMSE				0.0015			0.0014			0.0012

From Tables 2 to 4, it is clear that the estimated variable level conditional partial R-indicators have an increase in bias as the sample sizes decreases. Neither of the two methods reduces the bias entirely for the estimated conditional partial R-indicators and the results vary depending on the variable. The bias of the estimated conditional partial R-indicator based on Age Group improves under the proposed method, the estimated conditional partial R-indicator based on Locality Type improves under the pro-rating method, the estimated conditional partial R-indicator based on whether there is a child in the household shows no improvement for either method and the estimated conditional partial R-indicator based on the number of persons in the household wavers between no bias adjustment and the pro-rating method. On average, the method of pro-rating slightly outperforms the proposed method (which is based on adapting the non-response model) as can be seen by the absolute relative errors in Tables 2 to 4. However, we see that correcting for the bias increases the variation as shown in the relative root mean square errors of the Tables. Table 5 provides a summary of the average of the errors (true value minus the estimate) for the variable level conditional partial R-indicators across the 1000 samples for each of the sample sizes.

Table 5: Average error of the variable level conditional partial R-Indicator comparing method 1 (pro-rating) with method 2 (proposed) according to sample sizes

	1:50			1:100			1:200		
	No Bias Adjustment	Proposed Model	Prorating Model	No Bias Adjustment	Proposed Model	Prorating Model	No Bias Adjustment	Proposed Model	Prorating Model
Persons in HH	0.00008	0.00048	0.00069	0.00082	0.00029	0.00070	0.00105	0.00120	0.00192
Locality Type	0.00046	0.00028	0.00000	0.00096	0.00053	0.00004	0.00193	0.00108	0.00008
Age Group	0.00057	0.00023	0.00027	0.00108	0.00051	0.00057	0.00215	0.00102	0.00112
Child in HH	0.00048	0.00113	0.00063	0.00024	0.00143	0.00052	0.00031	0.00207	0.00087

In Figures 1 to 4 we present box plots of the estimated conditional partial R-indicators across the 1000 samples for each of the variables according to the different sample sizes. The horizontal line in each box plot represents the true value. From the figures, it is clear that the variation of the estimated conditional partial R-indicators increases as the sample sizes decrease and that the variance has a larger impact and dominates over the bias corrections.

Figure 1: Conditional partial R-indicator for persons in household comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

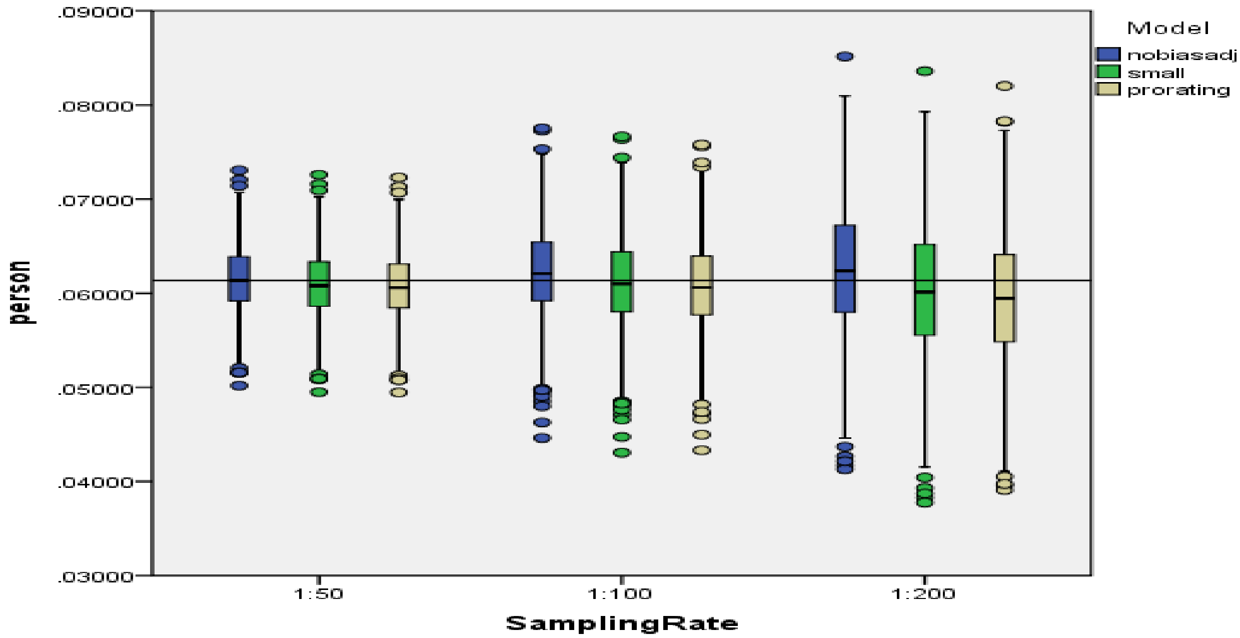


Figure 2: Conditional partial R-indicator for locality type comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

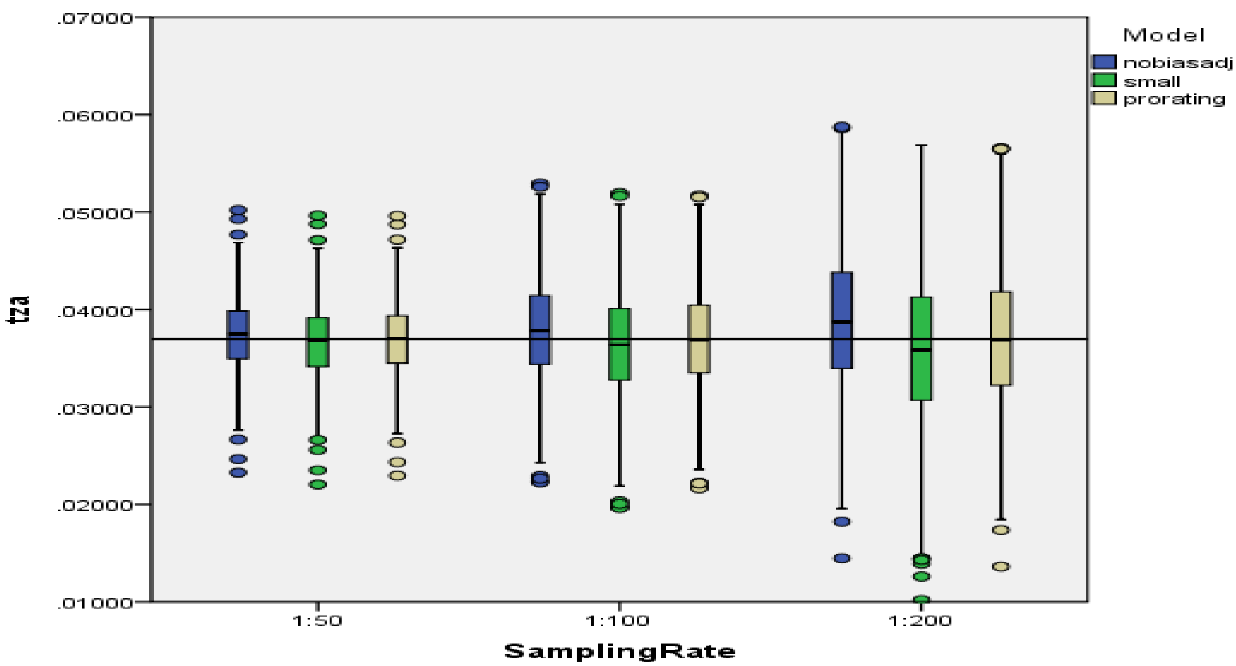


Figure 3: Conditional partial R-Indicator for age-group comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

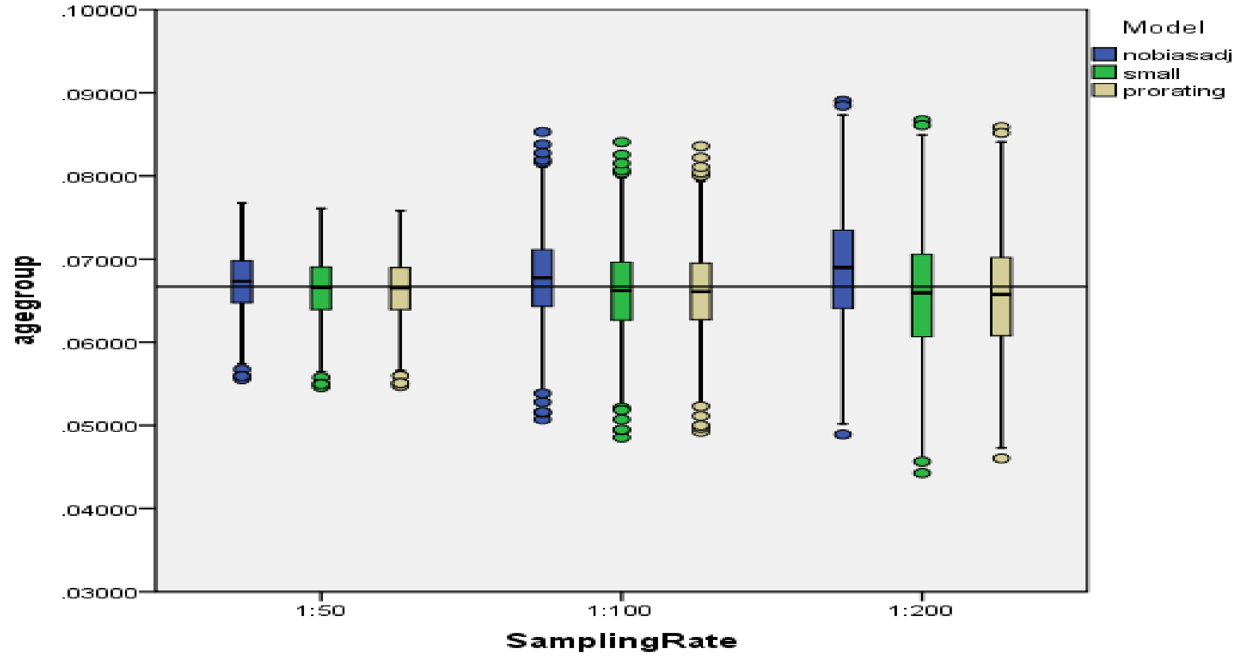
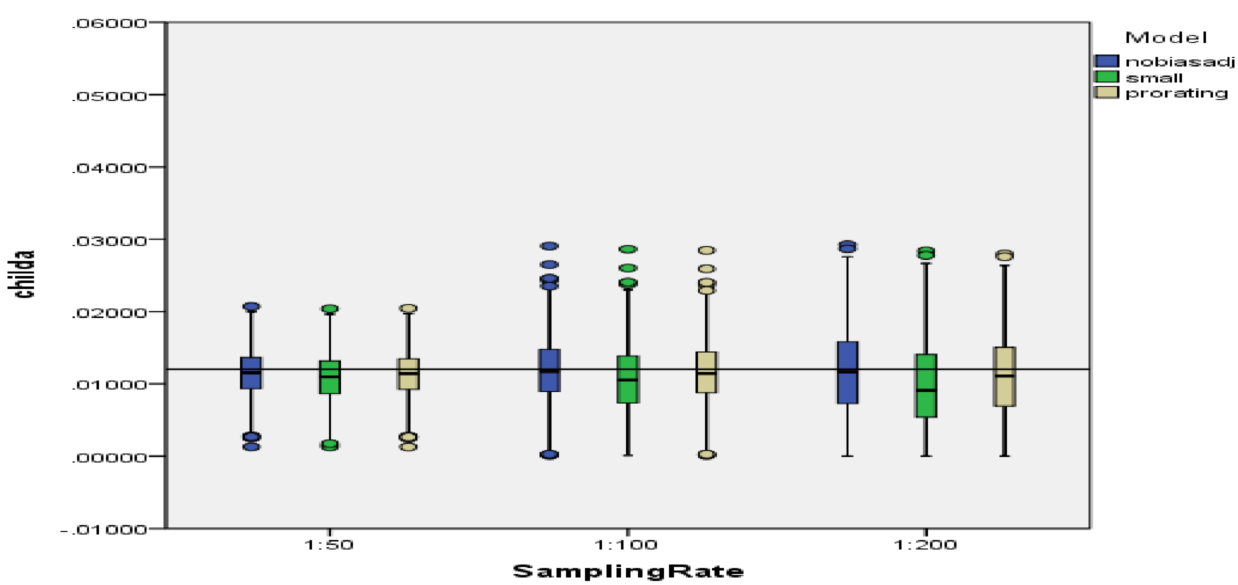


Figure 4: Conditional partial R-Indicator for child indicator comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes



We now carry out an evaluation of the results for the estimated variable level unconditional partial R-indicators in Tables 6 through 8 according to the varying sample sizes.

Table 6: Variable level unconditional partial R-indicator under the 1:50 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0654	0.0656	0.0037	0.35%	0.0650	0.0037	-0.63%	0.0648	0.0037	-0.90%
Locality Type	0.0281	0.0289	0.0038	2.77%	0.0276	0.0040	-1.67%	0.0285	0.0037	1.49%
Age Group	0.0670	0.0676	0.0037	0.95%	0.0667	0.0038	-0.39%	0.0668	0.0037	-0.31%
Child in HH	0.0329	0.0327	0.0035	-0.53%	0.0323	0.0036	-1.82%	0.0323	0.0035	-1.78%
Avg. Rel. Error				0.94%			0.90%			0.93%
RRMSE				1.50%			1.29%			1.25%
RMSE				0.0005			0.0005			0.0005

Table 7: Variable level unconditional partial R-indicator under the 1:100 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0654	0.0666	0.0049	1.86%	0.0653	0.0050	-0.09%	0.0650	0.0049	-0.62%
Locality Type	0.0281	0.0297	0.0053	5.61%	0.0272	0.0059	-3.41%	0.0290	0.0052	3.04%
Age Group	0.0670	0.0683	0.0052	1.94%	0.0665	0.0053	-0.74%	0.0666	0.0052	-0.54%
Child in HH	0.0329	0.0333	0.0050	1.15%	0.0324	0.0051	-1.42%	0.0324	0.0049	-1.31%
Avg. Rel. Error				2.31%			1.02%			1.06%
RRMSE				3.16%			1.88%			1.71%
RMSE				0.0012			0.0006			0.0006

Table 8: Variable level unconditional partial R-indicator under the 1:200 sample comparing method 1 (pro-rating) with method 2 (proposed)

	True Value	No Bias Adjustment			Proposed Model			Prorating Model		
		indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
Persons in HH	0.0654	0.0670	0.0070	2.51%	0.0645	0.0073	-1.41%	0.0638	0.0069	-2.36%
Locality Type	0.0281	0.0312	0.0072	10.88%	0.0261	0.0085	-7.03%	0.0297	0.0070	5.62%
Age Group	0.0670	0.0695	0.0070	3.85%	0.0660	0.0075	-1.50%	0.0662	0.0070	-1.09%
Child in HH	0.0329	0.0329	0.0069	0.01%	0.0311	0.0074	-5.43%	0.0313	0.0066	-4.75%
Avg. Rel. Error				3.77%			2.94%			2.80%
RRMSE				5.91%			4.56%			3.90%
RMSE				0.0022			0.0015			0.0014

From Tables 6 to 8, it is clear that the estimated variable level unconditional partial R-indicators have an increase in bias as the sample sizes decreases. Similar to the conditional partial R-indicators, neither of the two methods reduces the bias entirely and the results vary depending on the variable. The estimated unconditional partial R-indicator based on Age Group improves under the prorating method although it favoured the proposed method for the conditional R-indicator. The estimated unconditional partial R-indicator based on Locality Type wavers with the prorating method reducing the bias more effectively for the smaller sample sizes. The estimated unconditional partial R-indicator based on the child indicator shows no improvement for either method as was the case for the conditional partial R-indicator. The estimated unconditional partial R-indicator based on the number of persons in the household also wavers but the pro-rating method reduces the bias the most for smaller sample sizes. On average, the proposed method slightly outperforms the prorating method according to the absolute relative errors in Tables 6 to 8 for the larger sample sizes but for the smallest sample size, the prorating method achieves the most reduction in bias. Contrary to what was seen in Tables 2 to 4 with respect to the variation and the relative root mean square errors, we see that the prorating method provides the lowest relative root mean square errors. Table 9 provides a summary of the average of the errors (true value minus the estimate) for the estimated variable level unconditional partial R-indicators across the 1000 samples for each of the sample sizes.

Table 9: Average error of the variable level unconditional partial R-indicator comparing method 1 (pro-rating) with method 2 (proposed) according to sample sizes

	Unconditional Indicator								
	1:50			1:100			1:200		
	No Bias Adjustment	Proposed Model	Prorating Model	No Bias Adjustment	Proposed Model	Prorating Model	No Bias Adjustment	Proposed Model	Prorating Model
Persons in HH	0.00023	0.00041	0.00059	0.00122	0.00006	0.00041	0.00164	0.00092	0.00154
Locality Type	0.00078	0.00047	0.00042	0.00158	0.00096	0.00085	0.00306	0.00198	0.00158
Age Group	0.00064	0.00026	0.00021	0.00130	0.00050	0.00036	0.00258	0.00101	0.00073
Child in HH	0.00018	0.00060	0.00058	0.00038	0.00047	0.00043	0.00000	0.00179	0.00156

In Figures 5 to 8 we present box plots of the estimated unconditional partial R-indicators across the 1000 samples for each of the variables according to the different sample sizes. The horizontal line in each box plot represents the true value. As in the previous boxplots in Figures 1 to 4 for the conditional partial R-indicators, it is clear that the variation of the estimated unconditional partial R-indicators increases as sample sizes decrease and that the variance has a larger impact and dominates over the bias.

Figure 5: Unconditional partial R-indicator for persons in household comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

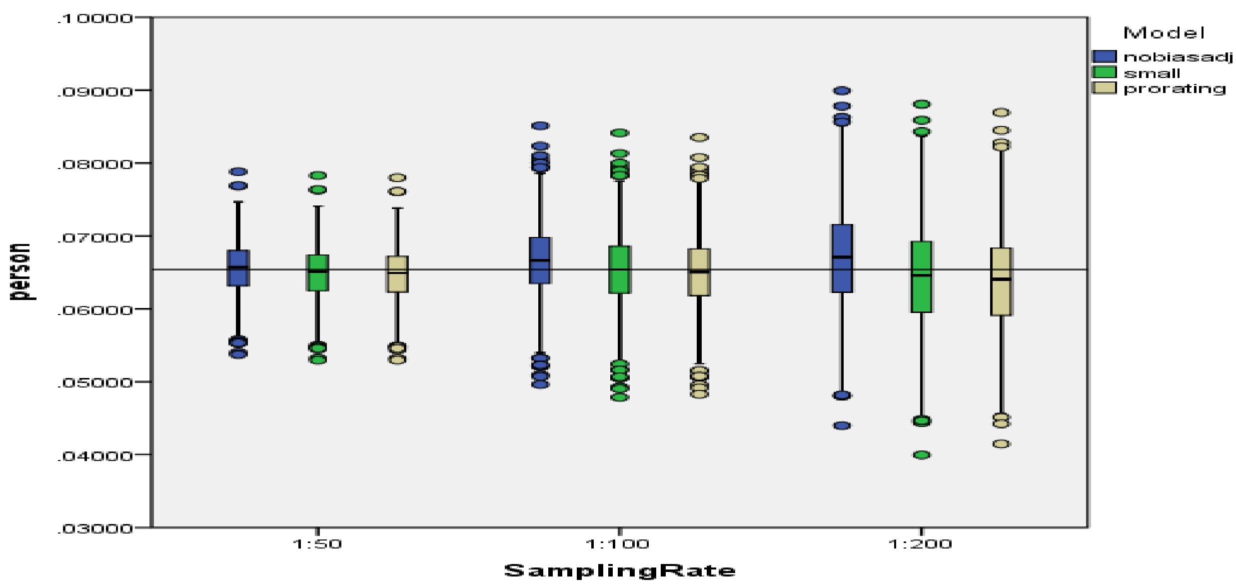


Figure 6: Unconditional partial R-indicator for locality type comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

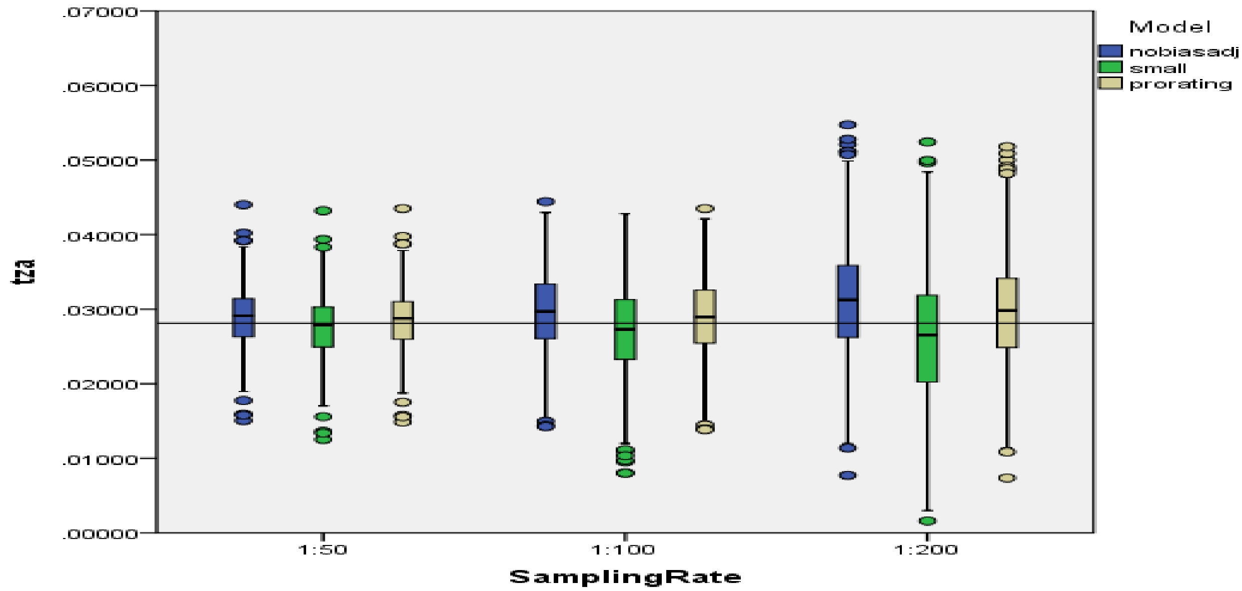


Figure 7: Unconditional partial R-indicator for age group comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes

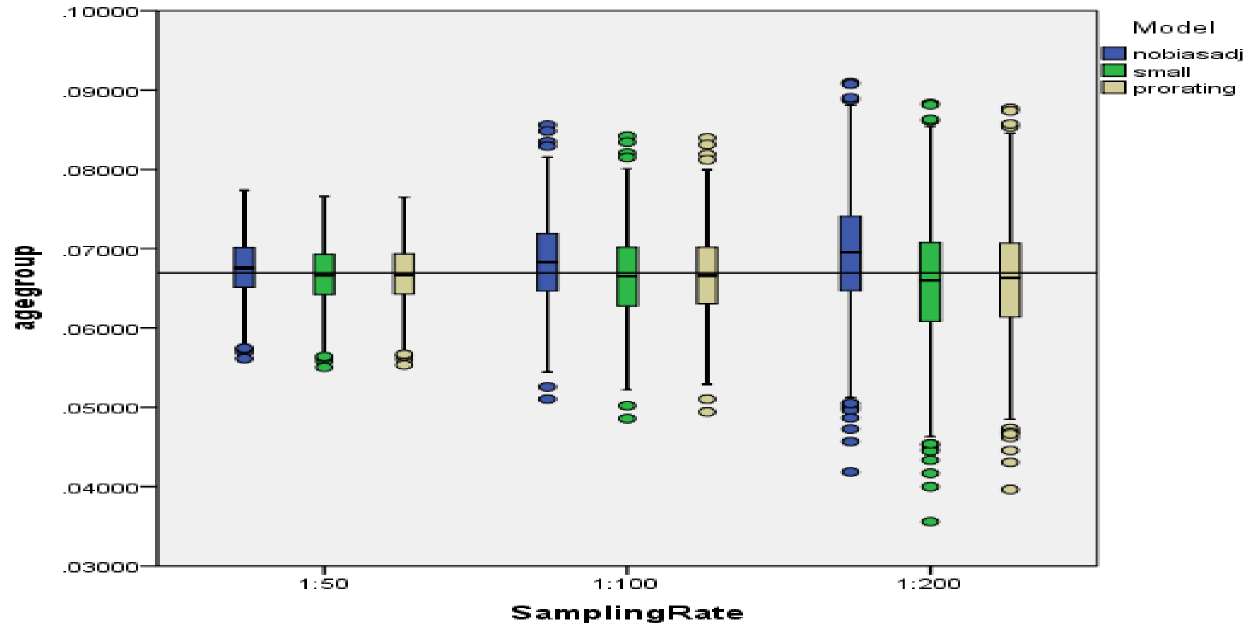
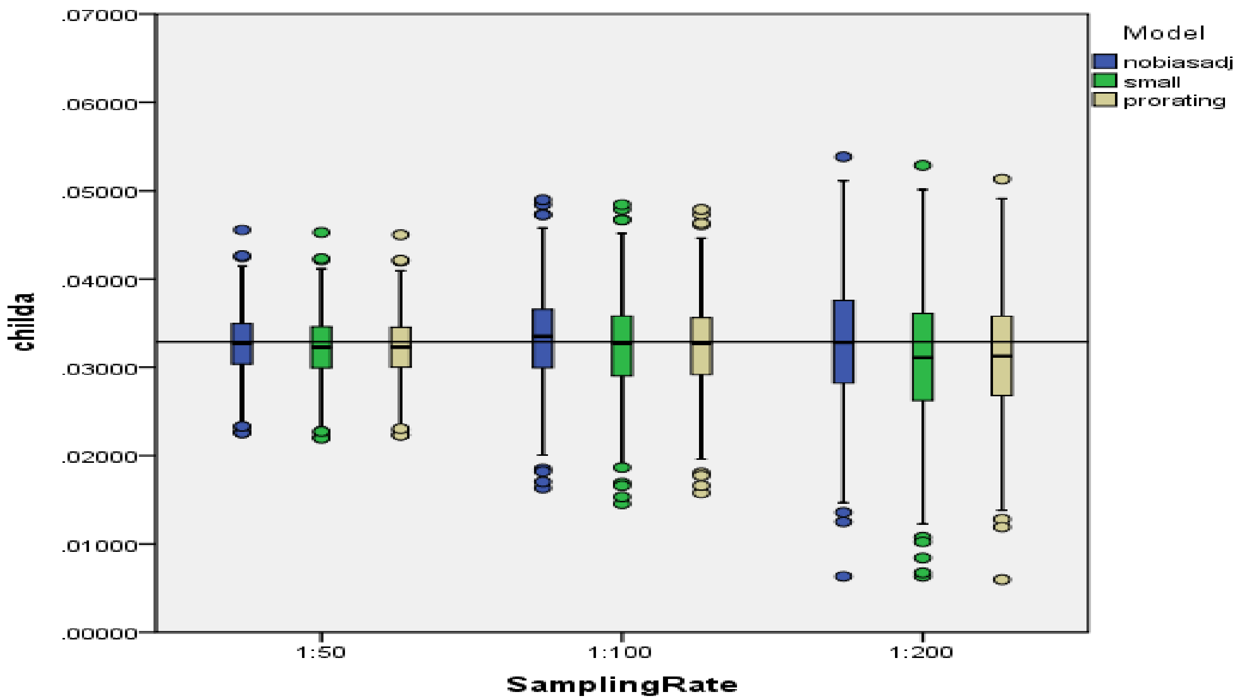


Figure 8: Unconditional partial R-Indicator for child indicator comparing method 1 (pro-rating) with method 2 (proposed) for varying sample sizes



Based on the results of the bias correction for the estimated variable level partial R-indicators, it is likely that any improvement in the bias according to the two methods may be related to the number of categories in the variable. We propose to not carry out any bias corrections if the sample sizes are over 15,000. For smaller sample sizes, we propose the method of prorating.

We next examine the method of prorating to obtain bias corrections for the estimated categorical level partial R-indicators where the bias term that is prorated arises either from Method 1 of prorating the overall bias correction term or Method 2 of adapting the response model according to the proposed method. We examine the results in Table 10 for the estimated categorical level unconditional partial R-indicator and Table 11 for the estimated categorical level conditional partial R-indicator at the smallest sample size with the 1:200 sample fraction.

Table 10: Category level unconditional partial R-indicator with prorating according to method 1 (pro-rating) and method 2 (proposed) for the 1:200 sample size

	Unconditional Indicator							
	True Unconditional Indicator	No Bias Adjustment			Prorating Proposed Model		Prorating Prorated Model	
		indicator (average)	indicator (stdev)	relative error	indicator (average)	relative error	indicator (average)	relative error
person1	0.00011	0.00012	0.00694	12.45%	-0.00033	-410.43%	-0.00031	-397.19%
person2	-0.03350	-0.03333	0.00578	-0.50%	-0.03321	-0.87%	-0.03195	-4.63%
person3	0.02647	0.02659	0.00660	0.45%	0.02651	0.15%	0.02551	-3.65%
person4	0.04278	0.04258	0.00604	-0.47%	0.04282	0.08%	0.04119	-3.73%
person5	-0.00941	-0.00952	0.00746	1.16%	-0.00964	2.39%	-0.00926	-1.55%
person6	-0.02307	-0.02328	0.00771	0.89%	-0.02319	0.49%	-0.02231	-3.33%
tza1	-0.02272	-0.02280	0.00680	0.33%	-0.02271	-0.07%	-0.01908	-16.04%
tza2	0.00781	0.00793	0.00478	1.58%	0.00797	2.14%	0.00669	-14.24%
tza3	-0.00021	-0.00064	0.00737	205.57%	-0.00077	269.60%	-0.00063	204.26%
tza4	0.01435	0.01459	0.00685	1.68%	0.01445	0.72%	0.01214	-15.40%
tza5	0.00271	0.00245	0.00725	-9.77%	0.00248	-8.69%	0.00210	-22.45%
agegroup1	0.02856	0.02865	0.00659	0.31%	0.02845	-0.38%	0.02698	-5.52%
agegroup2	-0.02330	-0.02344	0.00716	0.58%	-0.02307	-1.01%	-0.02188	-6.11%
agegroup3	-0.02602	-0.02573	0.00762	-1.12%	-0.02603	0.06%	-0.02470	-5.08%
agegroup4	-0.02795	-0.02792	0.00671	-0.09%	-0.02805	0.37%	-0.02661	-4.81%
agegroup5	-0.00455	-0.00445	0.00677	-2.20%	-0.00455	-0.11%	-0.00431	-5.37%
agegroup6	0.01371	0.01354	0.00661	-1.24%	0.01376	0.34%	0.01306	-4.79%
agegroup7	0.00682	0.00693	0.00726	1.51%	0.00687	0.66%	0.00651	-4.53%
agegroup8	0.02413	0.02410	0.00646	-0.11%	0.02364	-2.04%	0.02241	-7.10%
agegroup9	0.02881	0.02867	0.00618	-0.46%	0.02908	0.93%	0.02757	-4.29%
childa1	-0.01625	-0.01625	0.00344	0.01%	-0.01631	0.39%	-0.01542	-5.08%
childa2	0.02858	0.02859	0.00599	0.01%	0.02869	0.38%	0.02713	-5.10%
Average rel error				8.00%		9.81%		67.52%
RRMSE				43.97%		104.71%		95.58%
RMSE				0.00017		0.00023		0.00139

Table 11: Category level conditional partial R-indicator with prorating according to method 1 (pro-rating) and method 2 (proposed) for the 1:200 sample size

	Conditional Indicator									
		No bias adjustment			Prorating Proposed Model			Prorating Prorated Model		
	True Conditional Indicator	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error	indicator (average)	indicator (stdev)	relative error
person1	0.00632	0.00865	0.00319	36.75%	0.00834	0.00309	31.79%	0.00824	0.00304	30.21%
person2	0.02509	0.02503	0.00495	-0.21%	0.02414	0.00490	-3.78%	0.02384	0.00478	-4.95%
person3	0.02292	0.02323	0.00542	1.36%	0.02239	0.00533	-2.27%	0.02212	0.00521	-3.46%
person4	0.03989	0.03918	0.00562	-1.80%	0.03776	0.00565	-5.34%	0.03731	0.00547	-6.47%
person5	0.01582	0.01694	0.00507	7.09%	0.01633	0.00494	3.23%	0.01614	0.00485	1.99%
person6	0.02700	0.02671	0.00655	-1.10%	0.02575	0.00643	-4.65%	0.02544	0.00630	-5.80%
tza1	0.02399	0.02385	0.00641	-0.58%	0.02202	0.00650	-8.22%	0.02272	0.00616	-5.29%
tza2	0.00718	0.00906	0.00223	26.24%	0.00836	0.00228	16.40%	0.00863	0.00214	20.24%
tza3	0.00320	0.00674	0.00386	110.52%	0.00620	0.00362	93.85%	0.00642	0.00368	100.49%
tza4	0.02623	0.02606	0.00641	-0.66%	0.02404	0.00658	-8.36%	0.02482	0.00616	-5.38%
tza5	0.00639	0.00786	0.00506	23.07%	0.00726	0.00478	13.61%	0.00749	0.00483	17.25%
agegroup1	0.01147	0.01168	0.00528	1.84%	0.01115	0.00507	-2.83%	0.01113	0.00504	-3.00%
agegroup2	0.02971	0.02948	0.00656	-0.75%	0.02813	0.00640	-5.31%	0.02808	0.00631	-5.48%
agegroup3	0.02401	0.02369	0.00700	-1.35%	0.02260	0.00678	-5.87%	0.02256	0.00671	-6.04%
agegroup4	0.01731	0.01791	0.00500	3.50%	0.01709	0.00484	-1.26%	0.01706	0.00480	-1.42%
agegroup5	0.01162	0.01314	0.00420	13.04%	0.01253	0.00404	7.81%	0.01251	0.00401	7.66%
agegroup6	0.00925	0.01110	0.00323	20.03%	0.01059	0.00312	14.50%	0.01057	0.00309	14.32%
agegroup7	0.01117	0.01246	0.00471	11.56%	0.01189	0.00453	6.44%	0.01187	0.00450	6.25%
agegroup8	0.03218	0.03201	0.00619	-0.54%	0.03053	0.00607	-5.12%	0.03048	0.00597	-5.27%
agegroup9	0.03429	0.03394	0.00602	-1.02%	0.03237	0.00592	-5.58%	0.03232	0.00581	-5.73%
childa1	0.00818	0.00782	0.00399	-4.35%	0.00665	0.00393	-18.75%	0.00745	0.00381	-8.91%
childa2	0.00879	0.00870	0.00439	-1.09%	0.00738	0.00433	-16.02%	0.00828	0.00419	-5.81%
Average rel error				4.73%			7.36%			7.05%
RRMSE				26.62%			22.90%			23.79%
RMSE				0.00124			0.00151			0.00147

As can be seen in Tables 10 and 11, there is no reduction in the bias based on the prorating of either method. We propose to undertake no bias adjustment for the estimated categorical level partial R-indicators and leave them unadjusted.

We now proceed to examine the analytical expressions of the variance (standard deviation) of the estimated categorical level partial R-indicators according to the formulae in Section 3.2.

Tables 12 through 14 presents the comparison between the analytical expression of the variance (standard deviation) and the empirical variance (standard deviation) obtained from the repeated sampling for each of the sample sizes 1:50, 1:100 and 1:200. We calculate the analytical expression of the standard deviation in each of the 1000 samples and present the mean value and its standard deviation in the tables. We also calculate the empirical standard deviation of the estimated categorical level partial R-indicators across the 1000 repeated samples. If the analytical expressions of the standard deviations are correct, we should obtain that they are approximately equal to the empirical standard deviations as reflected by the relative bias also presented in the tables.

Table 12: Comparison of the empirical standard deviation (replication) and the average of the analytical standard deviations (with its standard deviation) for categorical level partial R-indicators under the 1:50 sample size

	Conditional R-Indicator				Unconditional R-Indicator			
	Replication	Analytical			Replication	Analytical		
	St. Dev	Average	St. Dev	Rel. Bias	St. Dev	Average	St. Dev	Rel. Bias
person1	0.00137	0.00212	0.00047	54.13%	0.00352	0.00494	0.00008	40.47%
person2	0.00243	0.00256	0.00002	5.23%	0.00279	0.003082	0.00002	10.55%
person3	0.00273	0.00278	0.00007	1.72%	0.00319	0.004601	0.00005	44.13%
person4	0.00280	0.00290	0.00003	3.76%	0.00301	0.004957	0.00005	64.91%
person5	0.00271	0.00281	0.00030	3.79%	0.00365	0.004778	0.00009	30.73%
person6	0.00343	0.00338	0.00005	1.39%	0.00401	0.004526	0.00012	12.96%
tza1	0.00348	0.00328	0.00003	5.68%	0.00358	0.003828	0.00004	6.97%
tza2	0.00098	0.00122	0.00022	23.78%	0.00257	0.002614	0.00002	1.80%
tza3	0.00185	0.00362	0.00030	95.43%	0.00354	0.005099	0.00010	43.89%
tza4	0.00350	0.00325	0.00004	7.26%	0.00370	0.005163	0.00007	39.69%
tza5	0.00338	0.00435	0.00147	28.69%	0.00374	0.005933	0.00037	58.79%
agegroup1	0.00273	0.00280	0.00004	2.82%	0.00317	0.005847	0.00009	84.48%
agegroup2	0.00367	0.00349	0.00003	4.76%	0.00388	0.004391	0.00009	13.04%
agegroup3	0.00380	0.00368	0.00005	3.25%	0.00395	0.004416	0.00011	11.69%
agegroup4	0.00276	0.00276	0.00021	0.19%	0.00351	0.003872	0.00005	10.45%
agegroup5	0.00236	0.00250	0.00041	5.97%	0.00332	0.004211	0.00005	26.79%
agegroup6	0.00180	0.00205	0.00048	13.69%	0.00333	0.004831	0.00006	45.00%
agegroup7	0.00271	0.00272	0.00040	0.61%	0.00364	0.005034	0.00008	38.21%
agegroup8	0.00329	0.00320	0.00004	2.72%	0.00334	0.005551	0.00008	66.20%
agegroup9	0.00305	0.00308	0.00007	0.83%	0.00312	0.006188	0.00010	98.36%
childa1	0.00222	0.00232	0.00007	4.22%	0.00185	0.002406	0.00002	30.29%
childa2	0.00245	0.00256	0.00009	4.41%	0.00324	0.004232	0.00004	30.69%

Table 13: Comparison of the empirical standard deviation (replication) and the average of the analytical standard deviations (with its standard deviation) for categorical level partial R-indicators under the 1:100 sample size

	Conditional R-Indicator				Unconditional R-Indicator			
	Replication	Analytical			Replication	Analytical		
	St. Dev	Average	St. Dev	Rel. Bias	St. Dev	Average	St. Dev	Rel. Bias
person1	0.00235	0.00368	0.00060	56.88%	0.00518	0.00698	0.00016	34.81%
person2	0.00357	0.00361	0.00005	1.01%	0.00415	0.00436	0.00004	5.12%
person3	0.00392	0.00391	0.00017	0.34%	0.00463	0.00651	0.00011	40.46%
person4	0.00429	0.00410	0.00005	4.39%	0.00457	0.00701	0.00011	53.42%
person5	0.00387	0.00391	0.00059	0.85%	0.00539	0.00676	0.00019	25.41%
person6	0.00469	0.00476	0.00010	1.42%	0.00540	0.00640	0.00025	18.56%
tza1	0.00404	0.00465	0.00005	15.10%	0.00426	0.00541	0.00009	27.13%
tza2	0.00133	0.00202	0.00036	51.41%	0.00343	0.00370	0.00004	7.75%
tza3	0.00277	0.00593	0.00095	113.84%	0.00516	0.00722	0.00021	40.02%
tza4	0.00440	0.00459	0.00008	4.39%	0.00487	0.00731	0.00015	50.09%
tza5	0.00411	0.00733	0.00393	78.25%	0.00500	0.00839	0.00073	67.94%
agegroup1	0.00369	0.00401	0.00018	8.78%	0.00432	0.00827	0.00019	91.57%
agegroup2	0.00494	0.00494	0.00005	0.02%	0.00533	0.00621	0.00018	16.49%
agegroup3	0.00532	0.00519	0.00011	2.27%	0.00562	0.00623	0.00022	10.99%
agegroup4	0.00387	0.00391	0.00038	1.02%	0.00498	0.00547	0.00010	9.91%
agegroup5	0.00316	0.00354	0.00067	12.03%	0.00482	0.00597	0.00011	23.93%
agegroup6	0.00245	0.00311	0.00064	26.89%	0.00493	0.00684	0.00013	38.59%
agegroup7	0.00345	0.00394	0.00056	14.35%	0.00491	0.00711	0.00017	45.02%
agegroup8	0.00473	0.00453	0.00008	4.34%	0.00486	0.00785	0.00017	61.65%
agegroup9	0.00444	0.00433	0.00014	2.51%	0.00457	0.00874	0.00021	91.38%
childa1	0.00317	0.00352	0.00153	11.29%	0.00256	0.00340	0.00004	33.02%
childa2	0.00348	0.00389	0.00173	11.67%	0.00446	0.00599	0.00009	34.37%

Table 14: Comparison of the empirical standard deviation (replication) and the average of the analytical standard deviations (with its standard deviation) for categorical level partial R-indicators under the 1:200 sample size

	Conditional R-Indicator				Unconditional R-Indicator			
	Replication	Analytical			Replication	Analytical		
	St. Dev	Average	St. Dev	Rel. Bias	St. Dev	Average	St. Dev	Rel. Bias
person1	0.00311	0.00618	0.00074	98.51%	0.00677	0.00989	0.00034	46.17%
person2	0.00489	0.00508	0.00011	3.81%	0.00563	0.00617	0.00009	9.69%
person3	0.00541	0.00549	0.00035	1.46%	0.00644	0.00920	0.00021	42.92%
person4	0.00572	0.00577	0.00010	0.81%	0.00597	0.00993	0.00021	66.19%
person5	0.00501	0.00553	0.00096	10.54%	0.00742	0.00956	0.00036	28.94%
person6	0.00652	0.00667	0.00020	2.26%	0.00770	0.00907	0.00051	17.70%
tza1	0.00656	0.00656	0.00015	0.12%	0.00694	0.00766	0.00018	10.46%
tza2	0.00211	0.00334	0.00061	57.95%	0.00501	0.00524	0.00008	4.58%
tza3	0.00412	0.00940	0.00256	128.36%	0.00739	0.01022	0.00041	38.41%
tza4	0.00617	0.00649	0.00022	5.10%	0.00672	0.01033	0.00032	53.69%
tza5	0.00503	0.01212	0.00873	141.02%	0.00713	0.01179	0.00150	65.36%
agegroup1	0.00492	0.00594	0.00088	20.88%	0.00617	0.01171	0.00035	89.71%
agegroup2	0.00658	0.00698	0.00009	5.94%	0.00717	0.00880	0.00034	22.82%
agegroup3	0.00722	0.00739	0.00016	2.32%	0.00777	0.00883	0.00044	13.54%
agegroup4	0.00503	0.00557	0.00072	10.82%	0.00680	0.00773	0.00021	13.66%
agegroup5	0.00406	0.00534	0.00083	31.41%	0.00689	0.00845	0.00022	22.54%
agegroup6	0.00377	0.00515	0.00076	36.62%	0.00728	0.00965	0.00026	32.64%
agegroup7	0.00451	0.00599	0.00065	32.69%	0.00669	0.01009	0.00033	50.75%
agegroup8	0.00649	0.00643	0.00017	1.01%	0.00658	0.01111	0.00033	68.96%
agegroup9	0.00602	0.00616	0.00028	2.26%	0.00625	0.01241	0.00042	98.57%
childa1	0.00406	0.00869	0.03517	114.03%	0.00348	0.00481	0.00009	38.38%
childa2	0.00448	0.00964	0.03921	115.06%	0.00604	0.00848	0.00017	40.48%

In Tables 12 to 14, we see that there is overestimation of the variance for the category level unconditional partial R-indicator. This result needs further investigation. For example, in Section 3.2, there is an assumption that the variable Z is not included in the response model to estimate the response propensities although this is generally not the case. It is possible that this assumption is inflating the analytical expressions of the standard deviations. For the category level conditional partial R-indicator, the analytical expressions of the standard deviations seem to be more similar to their empirical standard deviations but there are large discrepancies with some categories better estimated than others. For both category level partial R-indicators, there does not seem to be an improvement in the variance approximations as the sample size increases.

In spite of the overestimation, we recommend including the variance estimates of the category level partial R-indicators into the new version of the RISQ software given that the estimates of the variance are conservative. Further work will be carried out to see if the theory can be improved.

5. Conclusions and Future Work

The following is a summary of the recommendations in this document based on the results of the evaluation study in Section 4.

1. For the variable level partial R-indicators, no bias corrections should be applied if sample sizes are over 15,000. For smaller sample sizes, we propose the method of prorating, i.e. decomposing the overall variance of the response propensities into the between and within variance components and carrying out a prorating of the overall bias term as calculated in (1).
2. For the categorical level partial R-indicators, no bias corrections should be applied.
3. Include the analytical expressions for the variance of the categorical level partial R-indicators given their conservative estimation. Further work will be carried out to see if the analytical expressions of the variance can be improved.

This report can be viewed as a first exploration of the theoretical properties of the partial R-indicators. We have highlighted the problems and pitfalls in the theory developed so far for the bias corrections and analytical expressions of the variance. We are continuing to research and develop the theory to enable interpretation, statistical testing and comparison of the partial R-indicators in practical settings.

Further work on the R-indicators and partial R-indicators include:

- A new version of the software for the website with an accompanying manual,
- Research and develop R-indicators when only population level auxiliary information is known including bias corrections and standard errors,
- Continue developing and improving the theoretical properties of the partial R-indicators, particularly the analytical expressions of the variance for the categorical level partial R-indicators,
- Investigate the use of R-indicators and partial R-indicators in practical settings, including their use in driving follow-up in longitudinal surveys, data collection for web surveys and other selective designs, assessing representativity under informative response models.

Acknowledgements: We thank Xiaoling Ou from the University of Southampton for the computations.

References

Schouten, B., Cobben, F. and Bethlehem, J. (2009), Indicators for the Representativeness of Survey Response, *Survey Methodology* 35, pp. 101-113.

Schouten, B., Shlomo, N. and Skinner, C.J. (2011) Indicators for Monitoring and Improving Representativeness of Response, *Journal of Official Statistics,* Vol. 27, No. 2, pp. 231-253.

Schouten, B., Morren, M., Bethlehem, J., Shlomo, N. and Skinner, C. (2009) How to use R-indicators?. Work Package 4, Deliverable 3, RISQ Project, 7th Framework Programme (FP7) of the European Union, available at http://www.risq-project.eu/publications.html.

Shlomo, N., Skinner, C.J., Schouten, B., Bethlehem, J., Zhang, L. (2008) Statistical Properties of R-indicators, Work Package 3, Deliverable 2.1, RISQ Project, 7th Framework Programme (FP7) of the European Union , available at http://www.risq-project.eu/publications.html.

Shlomo, N., Skinner, C., Schouten, B., Carolina, T. & Morren, M. (2009), Partial Indicators for Representative Response, Version 2, Work Package 5, Deliverable 4 , RISQ Project, 7th Framework Programme (FP7) of the European Union , available at http://www.risq-project.eu/publications.html.

Shlomo, N., Skinner, C.J. and Schouten, B. (2012) Estimation of an Indicator of the Representativeness of Survey Response, *Journal of Statistical Planning and Inference* 142, pp. 201-211.