REPRESENTATIVENESS IN PANEL SURVEYS

Annamaria Bianchi[1], Silvia Biffignandi[2]

1 corresponding author:
Annamaria Bianchi
Department of Management, Economics and Quantitative Methods
University of Bergamo
via dei Caniana 2
24127 Bergamo, Italy
email: annamaria.bianchi@unibg.it

2 Silvia Biffignandi
Department of Management, Economics and Quantitative Methods
University of Bergamo, Italy

A panel is a set of units recruited and used in successive surveys. When the sample unit is the household, so-called R-indicators together with the comparison of distributions of certain variables to those known in the total population help to measure the representativeness of the panel. The method is applied to Understanding Society, a UK household longitudinal study. At each wave, under- and over-represented groups of individuals are identified. This allows the implementation of better survey designs and procedures to reduce the bias of non-response.

1

***Keywords**: R-indicator, household survey, longitudinal study, representativeness, targeted design, non response*

1. INTRODUCTION

A panel consists of a set of units (households, individuals, firms), recruited for participating in multi-wave surveys. The sample is drawn from the population using some probability scheme, and it is based on a sampling frame. Household panels, used for population-based studies (Hawkins et al., 2010; Ferro, 2011), are usually based on a convenient sample of households. The household is the sampling unit. Either all or some members of the selected households become panel members. Individuals are the observation units. The same panel members are interviewed at each wave.

Non-response endangers the representativeness of the panel. Indeed, people may fail to respond during recruitment or may drop out of the panel in the course of time: this is called panel attrition. When members are removed after a predetermined period of time, this is forced attrition. The decreasing participation in the panel, called unforced or normal attrition, is determined by non-response at several consecutive waves. The concern is that attrition can be systematic rather than random. A further problem is raised by decreasing response rates, which are reported for household panels (Leeuw and Heer, 2002). Scherpenzeel and Toepoel (2014) study different types of non response and attrition in panels and how efforts to obtain high response rates determine biases. Another non-sampling bias comes from item non-response (Särndal and Lundström, 2008; Mazza and Punzo, ????).

To mitigate the effects of non-response, Särndal and Lundström (2008) and Vandenplas et al. (????) use non-response adjustment techniques to reduce the non-response bias; Groves and Heeringa (2006) and Wagner (2008) focus on data collection monitoring and on targeting subpopulations in responsive and adaptive survey designs. These techniques refer to keeping data collection under control in order to identify under-represented subpopulations that could be treated differently. Responsive designs make reference to decisions taken during the fieldwork, while adaptive designs refer to interventions made according to rules specified before the data collection starts. These survey designs use available information from the sampling frame (frame data), administrative registers, and data related to the collection process (so-called *paradata*) in order to adapt collection procedures to the sample. For example, alternative modes (telephone, face-to-face, web) may be used for some households. Households in rural areas may be solicited more because they are less likely to respond. Such strategies are not much used in longitudinal surveys yet (Bianchi and Biffignandi, 2014, 2015; Lynn, 2015a, b). The term "panel maintenance" designates the set of means used to prevent attrition. It comprises incentives and inclusion of new individuals (Bethlehem and Biffignandi, 2012).

The practical implementation of these procedures requires to measure the representativeness of the collected data. The most widely used indicator of representativeness is the response rate. Groves (2006) and Groves and Peytcheva (2008) have criticized its poor link to the non-response bias. Using a meta-analysis technique, they find that the non-response rate of a survey is a poor predictor of the non-response bias. More reliable indicators are based on the use of auxiliary variables. Among those, subgroup response rates are also imperfect (Schouten et al., 2011), because small and large subgroups have equal importance, because they are not available at the variable level preventing the comparison of different variables for their effects, and because they do not easily allow conditioning on other variables. Schouten et al. (2011) propose partial R-indicators to determine the population subgroups with the strongest effect on the response. Särndal (2011) alternatively proposed balance indicators, which measure the similarity between the respondents and the sample by means of dissimilarity measures. Särndal and Lundström (2008) use goodness of fit statistics on propensity models which are helpful to choose auxiliary variables for non-response adjustment. Kreuter et al. (2010) use indicators involving not only frame data or paradata but also the survey data observed for respondents only, such as the correlations between auxiliary data and survey variables. Wagner (2012) reviews these indicators.

We review representativeness and its measurement for panel studies (section 2), propose to use comparisons between population distributions and R-indicators to evaluate the representativeness of household panels used to estimate individual characteristics (section 3). We apply our method to Understanding Society, a UK household longitudinal study (section 4). We identify groups of households and individuals which are under- or over-represented in the panel.

2. BASIC CONCEPTS FOR REPRESENTATIVENESS IN PANELS

Kruskal and Mosteller (1979a, b, c) show that the term "representativeness" may have different meanings, and that it is often used in a loose sense to convey a vague idea of good quality (Bethlehem et al., 2011).

Among the many interpretations of the term "representativeness," we first retain the one that a response is representative with respect to a variable if its percentage distribution is equal to its percentage distribution in the population. If a sample is representative with respect to enough auxiliary variables, it is also representative with respect to the variables of interest, also called target variables. To do so, auxiliary variables should be related both to target variables and to the response behavior. They should be measured in the survey, and their distribution in the population (or in the complete sample) should be known. This concept of representativeness forms the basis of many weighting adjustment techniques (Deville and Särndal, 1992; Särndal and Lundström, 2005; Bianchi and Biffignandi, 2013), which assign an adjustment weight to each survey respondent in order to reduce the effect of non-response.

The second interpretation of "representativeness," which we retain, is that the response is representative if each member of a population has the same probability of response when selected in a sample. This probability is called propensity. The response propensity for a unit *k* in the population is the probability of response, given a set of auxiliary variables *X*:

$$\rho_k\left(X\right)=P\left(r_k=1\,|\,X=x_k\right), \tag{1}$$

where the response indicator r_k equals 1 if the element *k* responds, and 0 otherwise (Rosenbaum and Rubin, 1983; Schonlau et al., 2009). A response is called "representative" with respect to covariates *X* when response propensities ρ_k are constant for all possible values of *X*. Denoting X^- the variables in *X* excluding a variable *Z*, the response to a survey is called "conditionally representative" for *Z* given X^- when the response propensities are equal for all choices of X^- **(**Schouten et al., 2009).

To measure representativeness (according to the second interpretation), Schouten et al. (2009, 2011) introduced representativity indicators, also called *R-indicators*. The overall measure of representative response is the R-indicator, which is defined as

$$R_\rho=1-2S_\rho, \tag{2}$$

where S_ρ denotes the standard deviation of the individual response propensities. It takes its values in [0, 1], with 1 indicating the most representative response and 0 the least representative response.

Besides the main R-indicator, Schouten et al. (2011) have introduced partial R-indicators to determine the population subgroups with the strongest effect on the response.

Unconditional and conditional partial indicators are defined at both the variable and the category level.

An unconditional partial R-indicator for a variable *Z* either is based on the between-variance, for a given stratification by categories of *Z*, or, equivalently, is defined as the standard deviation of response propensities based on the sole *Z*. It takes its values in [0.0, 0.5], with a high value reflecting an important contribution of the variable *Z* to the lack of representativeness.

The unconditional partial R-indicator is defined as the difference between the mean of the propensities for *Z*=*k* varying over the admissible range, $\bar{\rho}_k = N_k^{-1} \sum_{U_k} \rho_i$ (U_k denoting the set of population units in category *k*, and N_k the cardinality of U_k), and the overall mean, $\bar{\rho}_U = N^{-1} \sum_U \rho_i$, multiplied by the squared root of the proportion of units having *Z*=*k*, $\left(N_k / N\right)^{1/2}$. The unconditional partial R-indicator takes its values in [-0.5, 0.5], where zero indicates the absence of contribution. Positive values indicate over-representation; negative values indicate under-representation.

Conditional partial R-indicators measure the contribution of single variables to a lack of conditional representative response, given the other variables. They measure the remaining variance, due to the variable *Z* within sub-groups being formed by all other variables. The conditional partial R-indicator at the variable level is defined as the within-component of the total variance for a stratification based on X^{-}. This indicator takes its values in [0.0, 0.5], where zero means no conditional contribution of *Z*. Conditional indicators are useful in checking the collinearity of variables.

The conditional partial R-indicators for categories of variables are defined by distributing the within-component of the total variance over the classes of *Z*. They range from 0.0 to 0.5, where zero implies no conditional contribution of the category.

The indicators introduced above are population parameters. They are estimated from a sample s. We first estimate response propensities $\hat{\rho}_i$ through a logistic regression having the binary indicator for response (has responded or not) as the dependent variable and a set of covariates X as independent variables. Second, we weight the estimated propensities with sample design weights d_i, in order to estimate the population variances of the propensities. As an example, the overall R-indicator is estimated by $\hat{R}_{\hat{\rho}} = 1 - 2\hat{S}_{\hat{\rho}}$, where

$$\hat{S}^2_{\hat{\rho}} = \frac{1}{N-1}\sum_{i \in s} d_i \left(\hat{\rho}_i - \hat{\bar{\rho}}_U\right)^2 \tag{3}$$

and $\hat{\bar{\rho}}_U = \left(\sum_{i \in s} d_i \hat{\rho}_i\right) / N$. Shlomo et al. (2012) showed that the estimated R-indicator has a sample-dependent bias. They proposed an adjustment and derived an analytical variance estimator for $\hat{R}_{\hat{\rho}}$ under simple random sampling. Bianchi (2015) presented a bias adjustment and a variance estimator for the R-indicator in the case of cluster sampling. Similar other methods are used to estimate partial R-indicators (Schouten et al., 2011).

The representativeness of a panel can be understood either as the representativeness of the set of those who agreed to become members of the panel (the recruited panel), or as the representativeness of the panel at different waves, because members may drop out of the panel in the course of time. In the latter case, a “marginal” analysis devised to assess the representativeness of the survey with respect to the panel clarifies the functioning of the panel and the behavior of panel members. A “cumulative” analysis devised to assess the representativeness with respect to the population (or to the original sample) clarifies the joint effect of non-response due to recruitment and attrition.

Both types of representativeness may be studied at each wave of the panel. In order to study representativeness with respect to a variable, the variable should be measured in the survey and its distribution in the population should be known. To use the definition based on response propensities, data should be available for both respondents and non-respondents. Bianchi et al. (2015) introduce R-indicators when only population totals are available, and data on non-respondents are not available.

3. DATA

The UK Household Longitudinal Study is a longitudinal survey of individuals residing in the United Kingdom. Households recruited in the first round are visited each year to update their information on health, work, education, income, family, and social life.

Most data are collected through computer-assisted personal interviewing. Each wave spans over a 24-month period. The periods of waves overlap, and individual respondents are interviewed around the same time each year. Four waves are available, for the years 2009, 2010, 2011, 2012, 2013.

The UK Household Longitudinal Study has four sample components: the so-called "General Population Sample," the Ethnic Minority Boost Sample, the sample made of the participants from the British Household Panel Survey, and the Innovation Panel. In the empirical analysis, we consider the General Population Sample. It is based on an initial overall sample of 49,920 addresses, selected by a clustered stratified sample design for the British sub-sample, and a simple random sample for Northern Ireland, where the probability of being drawn is twice that in the British sub-sample. The overall response rate at the first wave of the General Population Sample at the household level is 57.2%.

The UK Household Longitudinal Study uses several instruments for members of selected households. Only one member of the household fills the household enumeration grid and the household interview. This task takes about 15 minutes. Then, each member of the household aged 16 or older answers to an individual adult interview (32 minutes) and fills a self-completion questionnaire (8 minutes). Young people aged from 10 to 15 are asked to fill a paper questionnaire with a pencil. There is also a brief proxy interview, where basic information is collected about unavailable other adult household members.

Rules specify which household members are to be surveyed (for example, new born and those who leave the household) in subsequent waves. They are used to take into account births and deaths, partnership formation and dissolution, and emigration, which may have occurred in the household. They however ignore immigrants into the UK (Buck and McFall, 2012). Apart from immigration, and if there were no attrition, the sample would remain representative of the UK population in its changes over time.

We investigate the representativeness of the General Population Sample with reference to adult interviews. The focus is on the longitudinal sample.

4. RESULTS

We examine the representativeness of the UK Household Longitudinal Study panel at each wave. As the panel is a clustered sample of households, no information on individuals in households who do not respond at the first wave is available. The only available information is the distribution of a few variables measured in the questionnaire, which is published by the Office for National Statistics, in 2009. As for households, only geographical variables are available from the frame. We add interviewers' notes and data related to the collection process (paradata).

On the first wave, we use R-indicators to assess the representativeness of households and compare the sample to the population. Variables are urban or rural residence, country of residence, year of interview, and interviewers' notes on locked common entrance of the residence (with categories mentioned or not mentioned), children's presence in the household, and the presence of intense traffic passing near the residence (with categories mentioned or not mentioned). Some variables comprise the category “missing”. Rather than treating these data as missing values (and excluding the household from the analysis), we consider them. We use design weights.

We compute R-indicators using the SAS code provided on the webpage of the Representativity Indicators for Survey Quality (RISQ) project. We adapt the code to take the complex survey design into account (Bianchi, 2015).

The overall R-indicator is 0.85, with a 95% confidence interval [0.84, 0.86]. It is significantly different from 1. Unconditional partial indicators at the variable level show that all variables contribute significantly to the lack of representativeness (Table 1). They remain significant even when conditioning on the other variables, even if the effect is widely reduced. Households from Scotland and Northern Ireland are under-represented, as well as households in rural areas. Households interviewed in the second year of operation are under-represented. This may be related to the fact that interviewers were also involved in interviews at the second wave. As regards interviewers’ notes, households without a locked common entrance to the residence are over-represented unconditionally, but not conditionally on the other variables. Households with at least one child are over-represented, both conditionally and unconditionally. This is due to the poor correlation of this variable with the other variables of the model, so that, after conditioning on other variables, the effect of the presence of at least one child remains.

Table 1. Partial R-indicators at the household level for respondents at the first wave compared to the sample. Values in italics refer to the variable-level indicators, the other variables to the corresponding category-level partial indicators.

	Partial R-indicators	
Variable	Unconditional partial (×1000)	Conditional partial (×1000)
Country	*34.8**	*32.4**
England	13.8*	13.5*

Wales	-3.9	2.7*
Scotland	-30.5*	27.3*
Northern Ireland	-8.6*	10.7*
Urban or rural area	*27.2**	*23.3**
urban	13.9*	12.4*
rural	-23.4*	19.8*
Year of Interview	*18.9**	*23.0**
first year	13.1*	16.3*
second year	-13.6*	16.3*
Locked common entrance to the residence	*39.5**	*6.3**
not mentioned	3.2*	2.9
mentioned	4.5	5.6
missing	-39.1*	0.9
Presence of at least one child	*39.3**	*35.0**
definitely or likely	33.2*	30.7*
definitely not or unlikely	1.8	3.2
unknown or missing	-20.8*	16.5*
Presence of intense traffic	*44.1**	*17.5**
not mentioned	4.7*	3.9
mentioned	2.2	1.3
missing	-43.8*	17.0*

*: significant at the 5% level.

For individuals, Table 2 shows the design-weighted panel distribution computed with design weights (inverse of selection probabilities), the population distribution, and p-values for Rao-Scott chi-square tests (computed taking the sample design into account) for sex, age, and country of residence. Each variable shows significant differences between the panel composition and the population. In comparison to the population, the panel contains older individuals and more women. Wales is slightly over-represented; Northern Ireland is slightly under-represented. However, although significant, the differences in percentage remain small. Our results are consistent with Burton, Laurie, and Lynn (2011), who compare the UK Household Longitudinal Study first wave with the 2009 Labor Force Survey. They find a higher proportion of women than in the Labor Force Survey, and a lower proportion of 65 or older. They find that the two surveys have similar sample distributions for other variables.

Table 2. Design-weighted panel composition and population distribution, in percent

Variable	Category	Panel members	Population	p*
Sex	women	56.2	51.5	<.0001
	men	43.8	48.5	
Age	16-24	12.5	14.6	<.0001
	25-39	24.7	24.9	
	40-54	27.1	25.9	
	55-64	15.8	14.5	
	65+	19.8	19.9	
Country	England	83.7	83.7	<.0001
	Wales	5.4	4.9	
	Scotland	8.4	8.5	
	Northern Ireland	2.4	2.8	

Source: Office for National Statistics 2009 population statistics, * p-values of the Rao-Scott chi-square test.

Moving to the representativeness of subsequent waves, the interest is in understanding the mechanisms of panel members' responses. We thus evaluate the representativeness of the response, in a marginal analysis. We consider the individuals responding to a wave, modelling the propensity to respond to the subsequent wave. The 0-1 response indicator identifies respondents to the current wave among the respondents to the previous wave. While some variables have the category "missing" in the first wave due to item non-response, this does not happen for subsequent waves, because most of the time it was possible to retrieve the missing information from previous responses. In the cases, where still missing values remain, the observations were deleted.

Variables considered for the estimation of response propensities are not only socio-demographics, but also "psychographic" variables known to be related to survey participation, and interviewers' notes (paradata). The socio-demographic variables included in the model are: sex, age (in classes), country of origin (UK, other), education (degree, other higher, A level, secondary, other qualification, no qualification), marital status (single, in a relationship, separated or divorced, widow), total number of children (0, 1, 1+), in paid employment (yes, no), urban or rural area of residence, country of residence, household size

(1, 2, 2+), and house tenure (owned, rented). The psychographic variables included in the analysis are: frequency of using the internet (never, sometimes, often), health (excellent or good, fair or poor), long-standing illness (yes, no), prefers to move house (stays here, prefers to move), expects to move next year (yes, no), supports a political party (yes, no), and level of interest in politics (very much, fairly, not very much, not at all). Among interviewers' notes, the variables locked common entrance to the residence (mentioned, not mentioned), unkempt garden (yes, no, no obvious garden), presence of intense traffic (mentioned, not mentioned), condition of residential properties in the area (good, fair, bad or very bad) are included. The variable locked common entrance to the residence was measured at the household level in the first wave only.

Table 3 shows the R-indicator for the response to each wave compared with the previous wave. The R-indicator for respondents to the second wave is 0.80, which is significantly different from 1. The R-indicator of each subsequent wave is higher than the previous one. Even though the confidence intervals are not overlapping, it is not possible to use them for testing hypotheses of equality of the R-indicators, as the samples are not independent from one another. To our knowledge, no procedure is available in this case yet. The composition of the panel in subsequent waves tends to be increasingly similar, likely because panel members are loyal.

Table 3. Size of the sample considered, total number of respondents, R-indicators, 95% confidence intervals for respondents to subsequent waves, and response rate with respect to respondents to previous wave (individual level).

Wave	Size of the sample	Total number of respondents	R-indicator	Confidence intervals	Response rates
1	-	41047	-	-	-
2	41047	31288	0.80	[0.79, 0.81]	75.99

3	31288	25543	0.85	[0.84, 0.86]	81.50
4	25543	22311	0.89	[0.88, 0.90]	88.81

Figures 1, 2, and 3 show the partial R-indicators. Precise values with identification of significant indicators at the 5% level are shown in Tables A1, A2, and A3. Partial R-indicators allow to identify which variables are contributing the most to the lack of representativeness, and which groups are over- or under-represented.

- Figure 1 about here -

Figure 1 shows the variable-level unconditional and conditional partial R-indicators for waves 2, 3, and 4. The unconditional partial R-indicators show that the variables with the largest deviations from representativeness are age, “expects to move next year,” marital status, housing tenure, and country of origin. The deviation is highest at the second wave, and decreases at subsequent waves. Figure 2 shows that the 16-25 age class is under-represented at waves 2 and 3 with respect to the previous wave, and the 56-65 age class is over-represented at all waves. People who expect to move next year are under-represented, and those who do not expect to move are over-represented with respect to the previous wave. In all waves, singles are under-represented with respect to the previous wave, and those living in a relationship are over-represented. Individuals living in owned houses are over-represented, while individuals living in rented houses are under-represented. Individuals of UK origin are over-represented and individuals of foreign origin are under-represented.

Conditioning on the other variables (using conditional indicators) removes part of the influence of the variables that are not representative enough. The variable “country of origin” remains significant for waves 2 and 3, as do “expects to move next year” and age for wave 2 (Figure 1). The lack of representativeness for the variable “age” is much reduced, although still significant for every class (Figure 3). Unconditionally health-related variables do contribute to the lack of representativeness, but, conditioning on other variables, their influence almost disappears entirely. The same holds true for political variables, and for interviewers' notes. This is due to the correlation of these variables with variables present in the analysis. For example, health is correlated to age and education. In sum, conditioning on other variables removes part of the influence of a given variable.

This allows us to identify subpopulations needing better representativeness: persons aged 16-25, those of foreign origin, persons with no qualification, those who expect to move. These results could be used for future data collection. The analysis of representativeness could be carried out at different times during data collection, in order to provide information, which could be used for adequate adjustment.

- Figures 2 and 3 about here –

5. CONCLUSION

Our method based on R-indicators and comparison of distributions is devised to assess the representativeness of household panels. It allows us to monitor the representativeness of the panel, and to identify subgroups requiring additional coverage. It should contribute to reduce non-response error by means of responsive and adaptive survey designs and panel maintenance.

We applied our method to the first four waves of the UK Household Longitudinal Study panel. R-indicators computed at the household level with respect to the sample allowed us to evaluate the representativeness of responding households at the first wave with respect to the sample. To assess the representativeness of individuals responding at the first wave, we compared the distribution of sex, age, and country of origin to the population distribution.

We computed R-indicators to evaluate non-response at successive waves. Representativeness varies mostly in the first waves. Representativeness can be improved for some groups. Young people and those who expect to move next year need better follow-up. More efforts could be directed towards such groups, in an adaptive design framework. For example, we could contact young people through their favorite communication devices. Those who expect to move next year are likely to move actually. Specific contact must be developed for them.

ACKNOWLEDGEMENTS

This article was supported by Biffignandi's research grants: "Ex 60% University of Bergamo founding grants" and PAADEL project (Lombardy region – University of Bergamo-CASI Center project). The authors acknowledge the networking support of the COST Action IS1004. The first author is thankful to Natalie Shlomo (University of Manchester) and Barry Schouten (Statistics Netherlands) for their helpful discussions.

We also thank guest-editor Vera Toepoel and two anonymous reviewers.

APPENDIX A

Table A1. Variable-level and category-level partial R-indicators for waves 2, 3, and 4 (marginal analysis). Values in italics refer to variable-level indicators, values in roman to the corresponding category-level partial indicators.

	Unconditional R-indicators (×1000)			Conditional R-indicators (×1000)		
Variable	Wave 2	Wave 3	Wave 4	Wave 2	Wave 3	Wave 4
Socio-demographics						
Sex	*14.7**	*6.6**	*4.8**	*4.8*	*2.4*	*2.4*
men	-11.0*	-5.0	-3.7	3.5*	1.7*	1.7*
women	9.7*	4.3	3.1	3.4*	1.7*	1.6*
Age	*62.1**	*44.3**	*31.2**	*6.2**	*4.9*	*4.1*
16-25	-50.7*	-38.9*	-26.0	2.8*	2.2*	1.8*
26-40	-10.6*	-5.6	-3.0	3.0*	2.5*	1.9*
41-55	13.8*	10.1*	9.0	2.8*	2.4*	1.6*
56-65	27.9*	16.9*	13.8*	3.3*	2.4*	2.2*
66+	14.1*	6.2	-4.4	1.5*	1.1*	1.5*
Country of origin	*46.2**	*34.0**	*14.3**	*7.5**	*7.3**	*2.3*
UK	17.1*	11.5	4.6*	4.8*	4.7*	1.5*
other	-42.5*	-32.0*	-13.5*	5.8*	5.5*	1.7*
missing	-6.0	-	-	0.1	-	-
Education	*17.2**	*25.3**	*23.6**	*4.2*	*6.9**	*3.7*
degree	6.9	16.7*	15.2	1.5*	3.9*	2.1*
other higher	10.2*	8.4*	6.5	1.5*	2.1*	1.2*
A level	-7.1	-7.6	-4.2	1.4*	2.2*	1.2*
secondary	-0.9	-3.4	-2.1	1.4 *	2.4*	1.2*
other qualification	0.4	0.4	-2.2	2.0*	2.1*	1.1*
no qualification	-7.5	-14.8*	-16.0*	2.3*	3.7*	2.0*
missing	-5.9	-	-	0.3	-	-
Marital status	*49.9**	*32.9**	*17.6**	*1.7*	*1.8*	*0.7*
single	-44.1*	-28.5*	-15.3*	0.9*	0.7*	0.2
relationship	20.4*	8.4*	7.9*	0.8*	0.8*	0.4
Separated or divorced	9.7*	13.6*	2.0	1.0*	1.3*	0.3
widow	4.9	4.1	-3.2	0.6	0.6*	0.5
missing	-1.6	-	-	0.0	-	-
Total number of children	*9.2**	*8.1**	*1.7*	*3.6*	*3.9*	*1.4*
0	-4.3	-3.0	-0.8	2.5*	2.5*	1.0*
1	1.6	-1.1	1.4	1.5*	1.7*	0.6
1+	8.0	7.4	0.6	2.0*	2.5*	0.7
In paid employment	*4.4*	*4.9*	*12.7**	*1.0*	*0.7*	*0.4*
yes	2.9	3.2	8.5	0.7	0.5	0.3
no	-3.2	-3.6	-9.4	0.7	0.5	0.3
missing	-0.6	-	-	0.0	-	-
Urban or rural area of residence	*20.5**	*6.4**	*10.3**	*2.0*	*0.5*	*1.9*
urban	-9.9*	-3.2	-5.2	1.3*	0.4	1.3*
rural	18.0*	5.6	9.0	1.5*	0.4	1.4*
Country of residence	*14.7**	*8.0**	*10.3**	*5.4*	*3.4*	*2.7*
England	1.2	1.0	2.8	2.8*	1.9*	1.5*
Wales	2.2	2.7	0.1	0.7	0.7	0.6
Scotland	-10.8*	-7.0*	-4.1	3.3*	2.2*	1.3*

Northern Ireland	9.6*	2.5	-9.0*	3.1*	1.7*	1.8

*: significant the 5% level.

Table A2. Variable-level and category-level partial R-indicators for waves 2, 3, and 4 (marginal analysis). Values in italics refer to variable-level indicators, values in roman to the corresponding category-level partial indicators.

	Unconditional R-indicators (×1000)			Conditional R-indicators (×1000)		
Variable	Wave 2	Wave 3	Wave 4	Wave 2	Wave 3	Wave 4
			Psychographics			
Frequency of using internet	*18.2**	*10.3**	*18.8**	*3.2*	*2.7*	*2.9*
never	-2.4	-9.0	-15.7*	2.2*	1.8*	1.7*
sometimes	4.9	1.1	-2.5	1.1*	0.9*	1.2*
often	0.4	4.9	10.1*	2.1*	1.8*	2.0*
missing	-17.3*	-	-	0.6	-	-
Health	*11.1**	*10.9**	*17.1**	*1.8*	*2.9*	*2.7*
excellent or good	1.0	5.2*	8.1*	1.2*	2.0*	1.9*
fair or poor	-1.1	-9.6*	-15.1*	1.3*	2.1*	2.0*
missing	-11.0*	-	-	0.1	-	-
Long-standing illness	*26.5**	*10.2**	*1.0*	*3.3*	*2.1*	*1.6*
yes	17.9*	8.1	0.8	2.4*	1.6*	1.2*
no	-13.0*	-6.2	-0.6	2.2*	1.5*	1.1*
missing	-14.7*	-	-	0.3	-	-
Prefers to move house	*22.4**	*14.0**	*12.2**	*1.0*	*0.5*	*1.3*
stays here	13.9*	8.4*	7.2	0.7	0.3	0.8*
prefers to move	-17.4*	-11.2*	-9.9	0.7	0.4	0.9*
missing	-2.4	-	-	0.4	-	-
Expects to move next year	*57.7**	*37.3**	*22.7**	*8.2**	*5.3*	*2.8*
yes	-51.7*	-34.8*	-21.3*	6.1*	4.0*	2.1*
no	23.2*	13.6*	7.7*	5.4*	3.5*	1.9*
missing	-11.1*	-	-	1.1*	-	-
Supports a political party	*24.2**	*15.2**	*6.9**	*0.5*	*1.0*	*0.1*
yes	16.9*	12.3*	5.7	0.3	0.7*	0.1
no	-10.4*	-9.0	-3.9	0.3	0.7	0.1
missing	-13.8*	-	-	0.3	-	-
Interested in politics	*33.2**	*25.1**	*20.6**	*4.3*	*2.8*	*2.5*
very	6.2	6.8	6.6	0.9*	0.6	0.6
fairly	16.8*	12.9*	9.7	2.5*	1.6*	1.2*
not very	2.3	0.6	1.6	1.7*	1.1*	1.1*
not at all	-23.9*	-20.4*	-16.9*	2.9*	2.0*	1.8*
missing	-14.3*	-	-	0.0	-	-
			Household			
Household size	*19.9**	*16.2**	*8.1**	*2.5*	*2.7*	*1.9*
1	-1.4	7.5	1.5	0.6*	0.9	0.9
2	15.2*	7.9*	5.3	1.6*	1.7	1.1*
2+	-12.7*	-12.0*	-6.0	1.8*	2.0*	1.3*
Housing tenure	*50.0**	*28.0**	*27.4**	*5.0*	*1.7*	*2.7*
Owned	28.2*	15.1*	14.4*	3.0*	1.1*	1.8*
Rented	-40.7*	-23.5*	-23.3*	3.3*	1.4*	2.0*
Missing	-7.0	-	-	1.1	-	-

*: significant at the 5% level.

Table A3. Variable-level and category-level partial R-indicators for waves 2, 3, and 4 (marginal analysis). Values in italics refer to variable-level indicators, values in roman to the corresponding category-level partial indicators.

	Unconditional R-indicators (×1000)			Conditional R-indicators (×1000)		
Variable	Wave 2	Wave 3	Wave 4	Wave 2	Wave 3	Wave 4
			Interviewer notes			
Locked common entrance to the residence	*7.0**	*4.9*	*6.9**	*2.2*	*1.5*	*2.0*
not mentioned	-2.7*	2.0	3.1	1.3*	0.9*	1.3*
mentioned	6.4*	0.1	-0.7	1.7*	0.3	0.4
missing	0.7	-4.4	-6.1	0.6	1.1*	1.5*
Unkempt garden	*34.3**	*15.9**	*11.0**	*2.1*	*0.2*	*0.5*
Yes	-10.6	-6.4	-7.6	0.4	0.1	0.2
No	17.3*	7.8	5.2	1.3*	0.2	0.3
no obvious garden	-26.1*	-12.2	-6.0	1.6*	0.2	0.4
Missing	-9.3	-	-	0.1	-	-
Presence of intense traffic	*13.9**	*9.6**	*1.4*	*0.3*	*0.8*	*0.3*
not mentioned	3.9	3.2	0.5	0.2	0.5	0.2
Mentioned	-8.2	-9.0	-1.3	0.2	0.7	0.2*
Missing	-10.5	-	-	0.0	-	-
Condition of residential properties in the area	*20.6**	*13.6**	*16.6**	*1.8*	*1.3*	*1.1*
Good	13.4*	8.9*	8.7	1.2*	0.8	0.5*
Fair	-11.3*	-8.6	-7.4	1.2	0.9	0.6
bad or very bad	-8.7	-4.6	-11.8	0.2	0.3	0.8
unable to obtain information or missing	-6.6	-3.1	-	0.6*	0.3	0.3

*: significant at the 5% level.

REFERENCES

Bethlehem, J., Cobben, F., and Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys*. Hoboken, NJ: Wiley.

Bethlehem, J. and Biffignandi, S. (2012). *Handbook of Web Surveys*. Hoboken, NJ: Wiley.

Bianchi, A., Shlomo, N., Schouten, B., Da Silva, D., and Skinner, C. (2015). Estimation of response propensities using population-level information with an application to Indicators of Representative Response, submitted.

Bianchi, A. (2015). R-indicators for Complex Survey Designs, in Preparation.

Bianchi, A. and Biffignandi, S. (2015). Targeted Design in Surveys and Longitudinal Studies, submitted to *Statistical Journal of the International Association for Official Statistics*.

Bianchi, A. and Biffignandi, S. (2013). Web panel representativeness, in *Statistical Models for Data Analysis*, P. Giudici, S. Ingrassia, M. Vichi (eds). Berlin: Springer, 37-44.

Bianchi, A. and Biffignandi, S. (2014). Responsive design for economic data in mixed-mode panels, in *Contribution to Sampling Statistics*, P. L. Conti, F. Mecatti, M. G. Ranalli (eds). Springer Series *Contributions to Statistics*, Springer, 85-102.

Buck, N. and McFall, S. (2012). Understanding Society: Design Overview. *Longitudinal and Life Course Studies*, *31*(1): 5-17.

Burton, J., Laurie, H., and Lynn, P. (2011). Appendix: Understanding society design overview, in *Early findings from the first wave of the UK's household longitudinal study*, S. L. McFall, C. Garrington (eds). Colchester: Institute for Social and Economic Research, University of Essex, 129-140.

Comley, P. (2007). Online market research, in *Market Research Handbook*, ESOMAR (eds). Hoboken, NJ: John Wiley & Sons, 401-420.

Deville, J. C. and Särndal, C. E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, *87*(418): 376-382.

Ferro, M. A. (2011). A Population-based Study of the Prevalence and Sociodemographic Risk Factors of Self-reported Epilepsy among Adults in the United Kingdom. *Seizure*, *20*(10): 784-788.

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5): 646-675.

Groves, R. M. and Heeringa, S. G. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. *Journal of the Royal Statistical Society: Series A*, *169*(3): 439-457.

Groves, R. M. and Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly*, *72*(2): 167-189.

Hawkins, J., Hollingworth, W., and Campbell, R. (2010). Long-term Smoking Relapse: a Study Using the British Household Panel Survey. *Nicotine & Tobacco Research*, 12(12): 1228-1235.

Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., and Raghunathan, T.E. (2010). Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Nonresponse: Examples from Multiple Surveys. *Journal of the Royal Statistical Society: Series A*, *173*(2), 389-407.

Kruskal, W. and Mosteller, F. (1979a). Representative Sampling, I: Non-scientific Literature. *International Statistical Review*, *47*(1): 13-24.

Kruskal, W. and Mosteller, F. (1979b). Representative Sampling, II: Scientific Literature, Excluding Statistics. *International Statistical Review*, *47*(2): 111-123.

Kruskal, W. and Mosteller, F. (1979c). Representative Sampling, III: The Current Statistical Literature. *International Statistical Review*, *47*(3): 245-265.

Leeuw (de), E. D. and Heer (de), W. (2002). Trends in household survey nonresponse: a longitudinal and international comparison, in *Survey Nonresponse*, R. M. Groves, D. A. Dillman, J. L. Eltinge, R. J. A. Little (eds). New York: Wiley, 41-54.

Lohr, S. (2010). *Sampling: Design and Analysis*. Boston, MA: Cengage Learning.

Lynn, P. (2015a). Targeted appeals for participation in letters to panel survey members, to appear in *Public Opinion Quarterly*.

Lynn, P. (2015b). Targeted response inducement strategies on longitudinal surveys, in *Improving Survey Methods: Lessons from Recent Research*, U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, P. Sturgis (eds). New York: Routledge, 322-338.

Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrica*, *70*(1): 41-55.

Särndal, C. E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. Hoboken, NJ: Wiley.

Särndal, C. E. (2011). The 2010 Morris Hansen Lecture: Dealing with Survey Nonresponse in Data Collection, in Estimation. *Journal of Official Statistics, 27*(1): 1-21.

Särndal, C. E. and Lundström, S. (2008). Assessing Auxiliary Vectors for Control of Nonresponse Bias in the Calibration Estimator. *Journal of Official Statistics*, *24*(2): 167-191.

Scherpenzeel, A. and Toepoel, V. (2014). Informing panel members about panel results: Effects on traditional and innovative forms of feedback on participation, in *Online Panel Research: A Data Quality Perspective*, M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, P. J. Lavrajkas (eds). Chichester: Wiley, 192-214.

Schonlau, M., van Soest, A., Kapteyn, A., and Couper, M. (2009). Selection Variables in Web Surveys and Propensity Scores. *Sociological Methods and Research*, *37*(3): 291-318.

Schouten, B., Cobben, F., and Bethlehem, J. (2009). Indicators for the Representativeness of Survey Response. *Survey Methodology*, *35*(1): 101-113.

Schouten, B., Shlomo, N., and Skinner, C. (2011). Indicators for Monitoring and Improving Representativeness of Response. *Journal of Official Statistics*, *27*(2): 231-253.

Schouten, B., Bethlehem, J., Beullens, K., Kleven, O., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N., and Skinner, C. (2012). Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-indicators and Partial R-indicators. *International Statistical Review*, *80*(3): 382-399.

Shlomo, N., Skinner, C., and Schouten, B. (2012). Estimation of an Indicator of the Representativeness of Survey Response. *Journal of Statistical Planning and Inference*, *142*(1): 201-211.

University of Essex. Institute for Social and Economic Research and National Centre for Social Research (2014). Understanding Society: Waves 1-4, 2009-2013 [data collection]. 6th Edition. Colchester, Essex: UK Data Archive. SN: 6614.

Wagner, J. (2008). Adaptive Survey Design to Reduce Non-response Bias. *PhD Thesis*, University of Michigan, USA.

Wagner, J. (2012). A comparison of alternative indicators for the risk of nonresponse bias. *Public Opinion Quarterly*, *76*(3): 555-575.

Figures' captions

Figure 1. Variable-level unconditional and conditional partial R-indicators (×1000).

Figure 2. Category-level unconditional partial R-indicators (×1000).

Figure 3. Category-level conditional partial R-indicators (×1000).