TREATMENT OF TOTAL NONRESPONSE VIA SEQUENTIAL WEIGHT ADJUSTMENT IN THE ITALIAN DISABILITY SURVEY

Claudia De Vitiis[1], Francesca Inglese, Marco D. Terribili

ISTAT - Italian National Institute of Statistics, Italy

Daniela Cocchi

Department of Statistical Sciences "Paolo Fortunati", University of Bologna, Italy

Received: 15 May 2013 / Accepted: 20 March 2014

Abstract. *Weighting adjustment techniques, adopted in the estimation phase to compensate nonresponse effect, are based on the use of auxiliary information known both for respondent and not respondent units, usually without distinguishing among different types of nonresponse. This paper proposes to treat separately the components of nonresponse, i.e. noncontact and refusal, with the aim of reducing nonresponse bias. Our weight adjustment method takes into account the sequential nature of the response process. Nested models are used to construct nonresponse adjustment weights in a two-stage response process, i.e. contact and participation conditional on contact. The assumption underlying this approach is that, conditional on auxiliary information, the different types of response are linearly independent.*

Keywords: *Nonresponse components, Propensity score, Sequential weight adjustment, Logit nested model, CART nested model.*

1. INTRODUCTION

This work addresses the reduction of the negative effects of nonresponse in surveys. As is well known, not observing the variable under study in some statistical units in a survey contributes to non-sampling error. Nonresponse comes from multiple causes: failure to contact the selected unit, inability to answer, refusal to cooperate with the survey. As a consequence, nonresponse reduces the accuracy of the final estimates, due both to the increase of the sample variance and the introduction of bias in estimates.

---

[1] Claudia De Vitiis, email: devitiis@istat.it

To reduce the effects of nonresponse on estimate quality, in the estimation phase the sampling weights associated to respondent units are adjusted, in order to take account of not respondent units. The methods of correction of the sampling weights of the units actually included in the sample are known as *weighting adjustment techniques* (Rizzo et al., 1996; Kalton and Flores-Cervantes, 2003).

In this paper, the correction for nonresponse directly models the response probability, under the assumption of a functional link between the probability response and a set of auxiliary variables known for all units belonging to the sample, either respondents and not respondents. The procedure is presented in two ways: under the traditional adjustment approach (Rizzo et al., 1996; Kalton and Flores-Cervantes, 2003), in which the components of nonresponse are considered without distinction, and under an alternative approach that considers the different types of nonresponse, *i.e.* noncontact and refusal, and the sequential nature of the response process, known as *sequential weight adjustment* (Groves and Couper, 1998; Bethlehem et al., 2011).

The ISTAT survey "Social integration of people with disabilities", named Disability survey in the following, was conducted in 2010 (ISTAT, 2012) as a specific follow-up of the 2004-2005 Multipurpose survey "Health conditions and use of health services" (ISTAT, 2008), named Health survey in the following. The Disability survey suffered from a high nonresponse rate, due more to the lack of contact of the disabled individuals highlighted by the first survey rather than to the refusal to cooperate when contacted. In this survey the response process is, according to the two-stage process above sketched, ascribed first to the contact phase of the disabled individuals and secondly to participation or refusal of disabled contacted individuals. However, a nice peculiarity of the Disability survey is that all the survey variables collected in the first occurrence (2004-2005 Health survey) were available in the 2010 Disability survey for the whole list of individuals, irrespective of their effective retrieval, helping to define nonresponse adjustment factors; this information corresponds to the set of auxiliary variables known for respondent and non respondent units, necessary for the construction of the weights. The adjustment procedure is developed in two successive steps: in the first the initial sampling weights are adjusted to account for noncontact, while in the second they are further adjusted to correct for nonparticipation. The sequentially obtained adjustment factors come from nested models, the first estimating the probability of being contacted and the second estimating the probability of contacted individuals to participate in the survey. To evaluate these probabilities we apply not only parametric models as in Groves and Couper (1998) and Iannacchione (2003), but we also propose the use of nested not parametric models, namely CART model

(Classification And Regression Tree, Breiman et al., 1984).

Sequential weight adjustment is a promising alternative to one shot traditional weight adjustments for nonresponse in the surveys of the Italian Statistical Institute. The validity of the method in this specific survey is assessed in two ways: first, an appropriate analysis of the survey nonresponse components is conducted, then a comparative analysis between the results obtained with the sequential and one shot approaches is performed. Each of the two approaches for constructing weight adjustments is implemented using the *propensity score* method (Rosenbaum and Rubin, 1984) and the techniques based on CART algorithm (Breiman et al., 1984; Rizzo et al., 1996) respectively; the final assessment relies on indicators providing both a measure of the goodness of fit, and a measure of the impact of the final weights on the variance of the estimates.

This paper is structured as follows: Section 2 outlines the issue of total nonresponse, considering causes and effects of its different components on estimation and introducing the theory of nonresponse treatment; Section 3 illustrates the nonresponse adjustment methods according to the two approaches (traditional and sequential), each of them seem under the perspective of parametric and not parametric models; Section 4 focuses on the Disability survey and on the application of the proposed methods and illustrate the studies on the nonresponse components, followed by a comparative assessment of the results obtained; finally, in Section 5, some concluding remarks are given.

2. TOTAL NONRESPONSE IN SURVEYS

2.1 OUTLINE

The term *total nonresponse* mirrors the situation where the expected response of a statistical unit called to participate in the survey is not obtained by the institute conducting the survey, for whatever reason (Särndal and Lundström, 2005).

In a sample survey, the lack of response produces two effects: i) the unplanned reduction of the sample size and, consequently, of the degree of precision of the estimates obtained, ii) the introduction of bias effects that reduce the accuracy of the final estimates. The occurrence of both effects leads to a loss in the reliability of the estimates, which increases the sampling variance, and introduces the risk of bias, the greater the more non respondent units differ systematically from respondent units with respect to the survey variables of interest.

In the estimation phase the compensation of nonresponse effects consists of adjusting the sampling weights associated to respondent units, so that they represent not respondent units. These techniques, called *weighting adjustment techniques*, are based on the use of auxiliary information known both for respondent and not respondent units and use known totals referred to the complete sample or to the whole population of interest. A particular class of weighting adjustment methods consist of forming cells of homogeneous sample units with respect to characteristics observed both on respondents and not respondents. The inverse of the response rate in each cell is assigned as a correction factor to all respondent units in the cell to compensate for the not respondent units. A thorough overview of adjustment methods appears in Kalton and Flores-Cervantes (2003).

2.2 CAUSES AND EFFECTS OF NONRESPONSE COMPONENTS

Nonresponse can be determined by multiple causes: the statistical unit does not receive the questionnaire or is not contacted by the interviewer (*noncontact*), or the contacted unit does not respond because unable (*inability*) or because he/she explicitly does not cooperate (*refusal*).

The distinction among the various components of nonresponse is important both in the prevention of nonresponse and in the estimation of the parameters of the population under investigation. The actions to be taken to prevent the phenomenon are different according to the type of nonresponse; indeed, refusal may be, unlike the noncontact, determined by a mental attitude of the statistical unit and thus can be strongly conditioned by the collection technique or by the particular phenomenon investigated.

The identification of different types of nonresponse (Groves and Couper, 1998) is important in the application of adjustment methods, as the effects on estimates bias of the different types of nonresponse may vary.

Bias is the most important effect of nonresponse. To investigate the effects of nonresponse on population parameter estimators, treatment of nonresponse should be integrated in sampling theory. This occurs through two approaches (Särndal et al., 1992): the first is the *fixed response model*, which assumes that the population consists of two subpopulations or strata, exclusive and exhaustive, the stratum of the responding units that, with certainty, provide an answer and the stratum of the not responding units that, still with certainty, do not respond. The second approach is the *random response model*, which assumes that each unit of the population has an unknown response probability. In both models the bias of an estimator depends on two factors: the response rate and the relationship between the response variables and the target variables of the survey.

Following the first approach, bias depends on differences in the characteristics under investigation between the two strata of respondent and not respondent units in the population. In the second approach, bias depends on the fact that the response probabilities associated to each unit of the population are not known. Furthermore, while the increase in variance can usually be assessed, the impact of bias on estimates is difficult to measure in both approaches, because in the first case a sample extracted from the population of not respondent units ought to be available for comparison, while in the second the knowledge of the process generating nonresponse would be required.

In more formal terms, let us denote by U the target population of interest, and assume that a sample s of size n (i = 1, ..., n) has been selected through a sample design, with a probability measure $p(s)$, s_R ($s_R \subseteq s$) is the subsample of respondent units of s. Considering the population mean $\overline{Y}$ of the surveyed variable y as the parameter of interest; in the fixed response model the nonresponse bias of a survey estimate $\hat{\overline{Y}}_{s_R}$, obtained via the responding units, is:

$$B\left(\hat{\overline{Y}}_{s_R}\right) = \frac{N_{NR}}{N}\left(\overline{Y}_R - \overline{Y}_{NR}\right), \tag{1}$$

where N is the total number of units in the population, N_{NR} is the number of not respondent units in the population, $\overline{Y}_R$ and $\overline{Y}_{NR}$ are the means of the target variable y calculated respectively on respondent and not respondent units in the population.

If nonresponse is attributed to two causes, contact failure and refusal, and the sequential nature of the response process is assumed, expression (1) becomes:

$$B\left(\hat{\overline{Y}}_{s_R}\right) = \frac{N_{NC}}{N}\left(\frac{N_{R|NC}}{N}\left(\overline{Y}_R - \overline{Y}_{R|NC}\right) + \frac{N_{RF|NC}}{N}\left(\overline{Y}_R - \overline{Y}_{RF|NC}\right)\right) + \frac{N_C}{N}\frac{N_{RF|C}}{N_C}\left(\overline{Y}_R - \overline{Y}_{RF|C}\right), \tag{2}$$

where N_{NC} is the number of missing contacts in the population, which can be divided into two parts, $N_{R|NC}$ and $N_{RF|NC}$, representing respectively the number of units that would have responded when contacted and the number of units that would have refused participation once contacted; $\overline{Y}_{R|NC}$ and $\overline{Y}_{RF|NC}$ are respectively the means of the two groups in the population. The last term in the equation takes into account the fact that participation in the survey and the refusal of the units are a consequence of the success in being contacted, in fact N_C is the number of contacted units, $N_{RF|C}$ the number of units that have expressed a refusal when contacted, $\overline{Y}_{RF|C}$ is the mean of the target variable *y* of units refusing to participate in the survey even if contacted (Bethlehem et al., 2011).

2.3 ESTIMATION IN THE PRESENCE OF NONRESPONSE

The theoretical basis on which the proposal of a correction factor relies is two-phase sampling (Särndal et al., 1992): the first phase sample is constituted by the initially selected sample, while the second phase sample coincides with the set of respondent units. The inclusion probabilities of the units can be defined for both sampling phases: in the first phase, the units of the sample *s* are selected according to the chosen sampling design that assigns the first order inclusion probability to the generic *i*-th unit of the population *U*, $\pi_i = \mathrm{P}(i \in s),\ (i \in U)$; in the second phase, the units of sample *s* of size *n* ($i = 1,\ldots, n$) are divided, on the basis of a random unknown mechanism, into two subsets, the respondents and the not respondents; this mechanism is summarised by assigning to each unit of the sample s_R of size n_R (i=1,…,n_R) the response probability $\theta_i = \mathrm{P}(i \in s_R \mid s)$, that represents the first order inclusion probability of the *i*-th unit in the second phase of the sampling.

When the parameter of interest is the population mean $\bar{Y}$ of the variable *y*

$$\bar{Y} = \frac{1}{N}\sum_{i\in U} y_i \tag{3}$$

an estimator for $\bar{Y}$ can be expressed as a linear homogenous function of the sampled units

$$\hat{\bar{Y}}_{HT} = \frac{1}{N}\sum_{i\in s_R} w_i y_i \quad (i=1,\ldots,n_R) \tag{4}$$

with final weights

$$w_i = d_i \gamma_i \qquad i \in s_R \ \text{ and } \ s_R \subseteq s\,, \tag{5}$$

where $d_i = 1/\pi_i$ is the initial direct weight in sample *s*, obtained as the reciprocal of the inclusion probability, γ_i is the multiplication factor in the sample of respondents s_R to correct for total nonresponse.

Assuming a response probability θ_i known for each respondent unit, an unbiased estimator for $\bar{Y}$ can be obtained. This estimator can be expressed in the following form

$$\hat{\bar{Y}}_{HT} = \frac{1}{N}\sum_{i \in s_R} d_i \gamma_i y_i = \frac{1}{N}\sum_{i \in s_R} \frac{y_i}{\pi_i \theta_i} \qquad (i=1,\ldots,n_R). \tag{6}$$

Briefly, the knowledge of the response probability for each respondent unit allows to obtain an unbiased estimate of $\bar{Y}$. However, since the subsample s_R is only one of the possible samples selected from the overall sample s, and the random mechanism that generates the set of respondents samples is unknown, the probability distribution $p(s_R \mid s)$ is not known and the inclusion probabilities of units in s_R cannot be computed.

So, the propensity of units of the first phase sample s to be respondent is to be estimated according to a model specification. Response models are discussed in Särndal and Lundström (2005). The estimator obtained by substituting θ_i by its estimate $\hat{\theta}_i$ is, however, biased.

By adding a random response model into design-based sampling theory, a further type of randomness is introduced: every unit selected in sample s carries also an unknown response probability. Consequently, the design-based theory becomes a quasi-randomisation theory (Oh and Sheuren, 1983).

3. NONRESPONSE ADJUSTMENT METHODS

3.1 AN OVERVIEW OF TWO METHODS

This section describes the nonresponse adjustment methods, respectively based on parametric and non-parametric models, assessed in the application which follows.

In the first method considered, known as *propensity score* (Rosenbaum and Rubin, 1984), response probabilities can be estimated through parametric models such as the logit (or probit), that link the expected value of the response indicator variable to the values assumed by a set of auxiliary variables. This method is based on two important assumptions: the *missing at random* (MAR) assumption and the *matching* assumption. According to the MAR assumption, response depends on a set of auxiliary variables, while the relationship between the variables of interest and the response variable is established in an indirect way. The stronger the relationship between the target and auxiliary variables, the more the use of correction techniques based on auxiliary information can lead to estimators reducing nonresponse bias. Moreover, the propensity score method assumes that the response probability is defined within the interval (0,1) (*matching assumption*). This assumption ensures that for every value of the observed characteristics both respondent and not respondent sample units can occur (Rosenbaum and Rubin, 1984; Bethlehem et al., 2011).

Response propensity may be used for nonresponse adjustment in different ways: the first approach, known as *response propensity weighting*, makes use of response propensities directly in estimating the survey variable, the second approach indirectly uses response propensities that intervene in the construction of strata or weighting cells (*response propensity stratification*). This approach is based on the assumption that within each stratum the response behaviour is the same for respondents and not respondents.

The response propensity weighting approach is strongly influenced by the correct specification of the estimation model and for this reason may lead to unreliable estimates; in addition, this strategy can considerably increase the variability of the weights and, therefore, of the estimates (Bethlehem et al., 2011). The response propensity stratification approach produces final estimates which are less dependent on the accuracy of the model because it uses response propensities only to build cells; this strategy can reduce the bias due to nonresponse if units, with regard to response propensity, are highly homogeneous within cells and present the highest heterogeneity in response propensity between cells. When response propensity is used indirectly for the construction of strata or weighting cells, a usual method for weighting cells construction is related to optimal stratification (Cochran, 1977) and consists of splitting the sample s on the basis of equal quantiles of the distribution of predicted probabilities for the sample units. According to this method each weighting cell contains an equal number of units in the sample. Thus, an adjustment factor for the propensity stratification method is calculated as the inverse of the response rate in the various cells. Little (1986) provides an overview of the weighting cell method.

Alternatively, nonresponse adjustment can be accomplished using a tree-based approach, such as the CART algorithm (Breiman et al., 1984). Following this non-parametric approach, used only to construct adjustment cells, assumptions about the missing data mechanism are not necessary. Decision trees algorithms provide a partition of the sample into groups, based, for each individual, on the relationship between the response indicator variables and the vector of the auxiliary variables. The algorithm searches, through the auxiliary variables, to find the variable to use for splitting the full sample in two or more groups in a way that best explains the response variable. This partition is based on a rule that, at each step, maximises the homogeneity of the response variable within the groups corresponding to terminal nodes of the classification tree. The final groups (cells) of the resulting optimal tree are treated as weighting adjustment cells; the adjustment factors are obtained, also in this case, as the inverse of the response rate in the obtained cells.

In recent years, methods based on parametric response models were applied in the *sequential weight adjustment* (Bethlehem et al., 2011), that takes into account the sequential nature of the response process and, therefore, the different types of nonresponse. Under this perspective, the participation of an individual in the survey constitutes the last stage of a process characterised by a sequence of events, namely the different stages of the response process, each of which is nested in the previous. The components of nonresponse are thus hierarchically distinct and the final probability of response of an individual is structured as a conditional probability of the outcome depending on the different stages of the process in which the response is nested. The fundamental assumption of the sequential adjustment method is that the stages of the response process are independent conditionally on a set of auxiliary variables. If nonresponse is characterised by the two “noncontact” and “refusal” components, the response process is modelled as a two-stage process and the correction procedure is also developed in two phases: in the first phase the base weight is adjusted to take noncontact into account, in the second phase the corrected weights are further adjusted to take the nonparticipation of contacted units into account.

For the construction of the adjustment factors, logit nested models can be used (Groves and Couper, 1998; Iannacchione, 2003) as two separate logistic regression models, one for each stage of the response process, first estimating the sample units probability to be contacted and then the contacted units probability to participate in the survey. In general, when the estimated response probabilities are used directly to correct the sample weights, the variance of the estimator (6) can became large. This effect increases with the number of nested stages (Little and Vartivarian, 2005).

In this paper, the two adjustment factors constructed for the sequential adjustment are defined both through the propensity score and the CART models.

3.2 PROPENSITY MODEL ADJUSTMENTS IN THE TRADITIONAL APPROACH

The *response propensity score* method (Rosenbaum and Rubin, 1984) assumes (*random response model*) that each *i*-th unit of the population has an unknown response probability, and that the probability that the *i*-th unit is respondent is $P\left(R_i = 1\right) = \theta_i$, while the probability of being not respondent is $P\left(R_i = 0\right) = 1 - \theta_i$.

Each *i*-th unit is still associated with a $p \times 1$ vector of auxiliary variables, known for all units in the sample *s*, $\mathbf{X}_i^R = (X_{i1}, X_{i2}, ..., X_{ip})'$.

The individual response probability is therefore a latent, *i.e.* not observed, variable of the model. Rather, the outcome of a survey is a random indicator variable R_i that assumes value 1 if the *i*-th unit included in the sample *s* is respondent and value 0 otherwise.

The response probability for the *i*-th sample unit conditionally on the values of the characteristics $\mathbf{X}_i^R$

$$\theta_i = \theta\left(\mathbf{X}_i^R\right) = P\left(R_i = 1 \middle| \mathbf{X}_i^R\right) \tag{7}$$

can be estimated through logit (or probit) models that associate the expected value of the indicator variable R_i with the values assumed by the vector of *p* auxiliary variables, $\mathbf{X}_i^R$ (Bethlehem et al., 2011). For each unit belonging to sample *s*, the logit model is

$$log\left(\frac{R_i}{1-R_i}\right) = logit\left(\theta\left(\mathbf{X}_i^R\right)\right) = \mathbf{X}_i^R{}'\boldsymbol{\beta}^R \quad (i=1,\ldots,n), \tag{8}$$

where: $\boldsymbol{\beta}^R$ is the $p \times 1$ vector of regression coefficients corresponding to the auxiliary variables in $\mathbf{X}_i^R$; *n* is the number of sample units for which the random indicator variable $\mathbf{R} = (R_1, R_2, \ldots, R_n)'$ is defined. The model has no quadratic terms and interactions and the auxiliary variables are all categorical. The parameters of the model are estimated by the maximum likelihood method (MLE).

The response propensities $\hat{\theta}\left(\mathbf{X}_i^R\right)$ are obtained as

$$\hat{\theta}_i = \hat{\theta}\left(\mathbf{X}_i^R\right) = \frac{exp\left(\mathbf{X}_i^R{}'\hat{\boldsymbol{\beta}}^R\right)}{1 + exp\left(\mathbf{X}_i^R{}'\hat{\boldsymbol{\beta}}^R\right)}. \tag{9}$$

The response propensity can be inserted in the Horvitz-Thompson estimator expressed by (6)

$$\hat{\hat{\bar{Y}}}_{HT} = \frac{1}{N}\sum_{i \in s_R} \frac{y_i}{\pi_i \hat{\theta}\left(\mathbf{X}_i^R\right)} \quad (i=1,\ldots,n_R). \tag{10}$$

This estimator is referred to as the *response propensity weighting*; the adjustment factor for *i*-th unit is specified as the reciprocal of the response propensity $\hat{\theta}\left(\mathbf{X}_i^R\right)$

$$\gamma_i = \frac{1}{\hat{\theta}\left(\mathbf{X}_i^R\right)} \quad i \in s_R . \tag{11}$$

If response propensities $\hat{\theta}\left(\mathbf{X}_i^R\right)$ are used to form *H* strata (cells), the *response propensity stratification* estimator is obtained using the correction factors calculated in each stratum. For the cell *h* (*h*=1,…, *H*) defined on sample *s* (*i*=1,…,*n*), the adjustment factor can be expressed as

$$\gamma_h^{'} = \left(\hat{\theta}_h\right)^{-1} = \left(\frac{n_{R,h}}{n_h}\right)^{-1} \tag{12}$$

where: the estimated response probability in cell *h*, $\hat{\theta}_h$, is the *h*-th scalar element of $\hat{\mathbf{\Theta}}^R = \left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_H,\right)$; n_h is the total number of sample units (respondents and not respondents) in the cell *h* and $n_{R,h}$ is the number of responding units in the same cell. This post-stratification assigns the same adjustment factors to all respondent units in the same stratum. The estimator defined using the correction factor (12) calibrates to the sample level through the use of response propensities (response propensity stratification estimator).

3.3 SEQUENTIAL TWO-STAGE PROPENSITY MODEL ADJUSTMENTS

Different functional forms can be used for the two components of nonresponse in specifying good response propensity models. The theory for constructing adjustment factors exposed in the former subsection, based on response propensity models, has to be adapted to reflect the two phases of units participation process to the survey.

Denote by s_c $(s_c \subseteq s)$ the subsample of contacted units and with $s_p (s_P \subseteq s_C)$ the sample of units participating in the survey when contacted.

In the first phase, the contact probability that the *i*-th unit belonging to the sample *s* is contacted conditionally on the characteristics $\mathbf{X}_i^C = (X_{i1}, X_{i2}, ..., X_{iq})'$ can be expressed as

$$\theta_{1i} = \theta_1\left(\mathbf{X}_i^C\right) = P\left(C_i = 1 \middle| \mathbf{X}_i^C\right), \tag{13}$$

where C_i equals 1 if the *i*-th unit of the sample *s* is contacted and equals 0 otherwise. At the second stage, conditional on the contact, i.e. when C_i=1, the participation probability that the *i*-th unit participates conditionally on the characteristics $\mathbf{X}_i^C$ is

$$\theta_{2i} = \theta_2\left(\mathbf{X}_i^P\right) = P\left(P_i = 1 \middle| \mathbf{X}_i^C, C_i = 1\right), \tag{14}$$

where P_i equals 1 if *i*-th unit included in the contacted sample s_c participates and equals 0 otherwise.

In the sequential approach nested models can be used for the construction of the adjustment factors, considering two different logit models, one for each stage of the response process.

The logit model for contact probability is

$$log\left(\frac{C_i}{1-C_i}\right) = logit\left(\theta_1\left(\mathbf{X}_i^C\right)\right) = \mathbf{X}_i^{C}{}'\boldsymbol{\beta}^C \quad (i=1,\ldots,n) \tag{15}$$

for each *i*-th unit belonging to the initial sample *s of size n*, that is the number of sample units for which the random variable indicator $\mathbf{C} = (C_1, C_2, ..., C_n)'$ is defined.

The contact probability is estimated as

$$\hat{\theta}_{1i} = \hat{\theta}_1\left(\mathbf{X}_i^C\right) = \frac{exp\left(\mathbf{X}_i^C \hat{\boldsymbol{\beta}}^C\right)}{1 + exp\left(\mathbf{X}_i^C \hat{\boldsymbol{\beta}}^C\right)}. \tag{16}$$

At the second stage of the response process, the logit model for the participation probability is

$$log\left(\frac{P_i}{1-P_i}\right) = log\,it\left(\theta_2\left(\mathbf{X}_i^{(P|C=1)}\right)\right) = \mathbf{X}_i^{(P|C=1)}{}'\boldsymbol{\beta}^{(P|C=1)} \quad (i=1,\ldots,n_c), \tag{17}$$

for each *i*-th unit belonging to the sample s_c of size n_c that is the number of contacted units for which the random variable indicator $\mathbf{P} = (P_1, P_2, \ldots, P_n)'$, is defined, while the vector of auxiliary variables is $\mathbf{X}_i^{P|C=1} = \left(X_{i1}, X_{i2}, ..., X_{iv}\right)'$.

The participation probability is estimated as

$$\hat{\theta}_{2i} = \hat{\theta}_2\left(\mathbf{X}_i^P\right) = \frac{exp\left(\mathbf{X}_i^{P|C=1}\,\hat{\boldsymbol{\beta}}^{P|C=1}\right)}{1 + exp\left(\mathbf{X}_i^{P|C=1}\,\hat{\boldsymbol{\beta}}^{P|C=1}\right)}. \tag{18}$$

The estimator of $\bar{Y}$ is computed substituting the contact probability θ_{1i} and the participation probability θ_{2i} respectively with the contact propensity $\hat{\theta}_1\left(\mathbf{X}_i^C\right)$ and the participation propensity $\hat{\theta}_2\left(\mathbf{X}_i^P\right)$, obtaining a *response propensity weighting* estimator as

$$\hat{\bar{Y}}_{HT} = \frac{1}{N}\sum_{i\in s_R}\frac{y_i}{\pi_i\hat{\theta}_1\left(\mathbf{X}_i^C\right)\hat{\theta}_2\left(\mathbf{X}_i^P\right)} \qquad (i=1,\dots,n_R) \qquad (19)$$

The adjustment factor for the ***i***-th unit is specified as the reciprocal of the contact propensity $\hat{\theta}_1\left(\mathbf{X}_i^C\right)$ and as the reciprocal of the participation propensity $\hat{\theta}_2\left(\mathbf{X}_i^P\right)$

$$\gamma_{1i} = \frac{1}{\hat{\theta}_1\left(\mathbf{X}_i^C\right)} \quad \text{and} \quad \gamma_{2i} = \frac{1}{\hat{\theta}_2\left(\mathbf{X}_i^P\right)}. \qquad (20)$$

If response propensity is used to form cells according to the sequential two stage approach, the final weight of each respondent unit is the product of the direct weight and two correction factors (*response propensity stratification*): the first is the inverse of the estimated contact probability in the cell *f* (*f*=1,…, *F*); the second is the inverse of the estimated participation probability in the cell *g* (*g*=1,…, *G*). For cell *f* defined on sample *s* (*i*=1,8,*n*) and for cell *g* defined on sample $s_c\,(i=1,\dots,n_c)$ the adjustment factors are respectively

$$\gamma'_{1f} = \left(\hat{\theta}_f\right)^{-1} = \left(\frac{n_{C,f}}{n_f}\right)^{-1} \text{ and } \gamma'_{2g} = \left(\hat{\theta}_g\right)^{-1} = \left(\frac{n_{P,g}}{n_{C,g}}\right)^{-1}, \qquad (21)$$

where: the estimated contact probability in cell f, $\hat{\theta}_f$, is the *f*-th scalar element of $\hat{\mathbf{\Theta}}^C = \left(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_F,\right)$ and $n_{C,f}$ is the number of contacted sample units in that cell (first stage); the estimated participation probability in cell *g*, $\hat{\theta}_g$, is the *g*-th scalar element of $\hat{\mathbf{\Theta}}^P = \left(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_G,\right)$ and $n_{p,g}$ is the number of respondent sample units, being contacted, in cell *g* itself (second stage). In this case two subsequent poststratifications occur: the first assigns the same adjustment factor to all contacted units within each stratum defined on sample *s*; the second assigns the same adjustment factor to all respondent units within each stratum defined on sample s_c. In this way, the response propensity stratification estimator calibrates both to the sample *s* and to the sample s_c using contact propensities and participation propensities.

3.4 ONE STAGE AND SEQUENTIAL CART MODELS

A CART model describes the conditional distribution of the target variable **Z** given a set of *p* predictors defined by matrix **X** of size $n \times p$. This model has two main components: a tree *T* with *M* terminal nodes, and a parameter $\hat{\boldsymbol{\Theta}} = (\theta_1, \theta_2, ..., \theta_M)$ which associates the parameter values θ_m with the *m*-th terminal node (*m=1,…, M*).

A decision tree model is fully specified by a couple $(T, \boldsymbol{\Theta})$. If **X** ies in the region corresponding to the *m*-th terminal node, then **Z**|**X** has the distribution $f = (\mathbf{Z}|\theta_m)$, were *f* is used to represent a conditional distribution indexed by . The model is called a regression tree if **Z** is quantitative and classification tree if **Z** is qualitative.

The main idea behind tree methods is to recursively partition the data into smaller and smaller strata in order to improve the fit at best. They partition the sample space into a set of rectangles and fit a model in each one. Generally, the preferred strategy to find the optimal partitions is to consider a full tree *T* and to prune a large tree using a "cost-complexity" function.

The optimal classification tree is usually based on a compromise between the complexity of the tree and the misclassification of statistical units distributed among the groups and corresponds to the tree with minimum value of the function of cost-complexity $\Phi_\alpha(T)$:

$$\Phi_\alpha(T) = \Phi(T) + \alpha \cdot Q(T), \tag{22}$$

in which $\Phi(T)$ is the misclassification error associated with a given tree *T*, equal to the fraction of misclassified observations, *Q(T)=M* is the number of terminal nodes and α is a coefficient that penalises the complexity of the tree. The misclassification error $\Phi(T)$ takes the value zero for the maximum partition, and tends to infinity when the tree coincides with the root node. The parameter $\alpha \geq 0$ controls the tradeoff between the tree size and its goodness of fit to the data. Large values of α result in smaller T_a, and conversely for smaller values of α. With = 0 the solution is the full tree *T*. To obtain the final tree $T_{\hat{\alpha}}$ it is necessary to estimate the coefficient α, that minimises the cost-complexity $\Phi_\alpha(T)$ function. The estimation of α (i.e. $\hat{\alpha}$) is achieved through repeated cross-validation analysis (Breiman et al., 1984).

In the present study the tree-based approach is used to construct nonresponse adjustment cells (as mentioned in paragraph 3.1), therefore the classification tree models are suitably defined both in the traditional and in the sequential approach.

In the traditional approach, for the estimation of the response probability $\mathbf{\Theta}^R$, the conditional distribution of $\mathbf{R}|\mathbf{X}^R$ for the m-th terminal node is defined as $f = (\mathbf{R}|\theta_m)$, where: $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ is the random response indicator defined on sample s of size n (i=1,…,n) that assumes value 1 for respondent units and value 0 for not respondent units; $\mathbf{X}^R$ is the $n \times p$ matrix of predictors (CART response model).

The estimated parameters of the traditional response model $(T^R, \mathbf{\Theta}^R)$ are: the final tree $T_{\hat{\alpha}}^R$ with K (K<M) terminal nodes (k=1,…,K) and the estimated response rate in each terminal node $\hat{\mathbf{\Theta}}^R = (\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_K)$.

Nonresponse adjustment factors are calculated in each terminal node as the inverse of the response rates, obtaining the correct weight in the same way as expressed in formula (12).

As regards the sequential weight adjustment, two nested CART models need to be defined, one for each stage of the response process: the first, with parameters $(T^C, \mathbf{\Theta}^C)$, for estimating the probability of the units of sample s to be contacted and the second, with parameters $(T^P, \mathbf{\Theta}^P)$, for estimating the probability to participate in the survey of contacted units. In the first stage the conditional distribution of $\mathbf{C}|\mathbf{X}^C$ for the m-th terminal node is defined as $f_1 = (\mathbf{C}|\theta_m)$, where: $\mathbf{C} = (C_1, C_2, \ldots, C_n)$ is the random contact indicator defined on sample s of size n (i=1,…,n) with values 1 and 0 respectively for contacted and not contacted units; $\mathbf{X}^C$ is a $n \times q$ matrix of predictors (CART contact model); in the second stage the conditional distribution of $\mathbf{P}|\mathbf{X}^P$ for the m-th terminal node is defined as $f_2 = (\mathbf{P}|\theta_m)$, where: $\mathbf{P} = (P_1, P_2, \ldots, P_{n_C})'$ is the participation indicator defined on sample s_C of size n_c (i=1,…,n_c) with value 1 if the contacted units are respondents and 0 otherwise; $\mathbf{X}^P$ is a $n_C \times v$ matrix of predictors (CART participation model).

The estimated parameters of the sequential contact and participation models are: the final trees $T_{\hat{\alpha}}^C$ with L (L<M) terminal nodes (l=1,…,L) and the estimated contact rate $\hat{\mathbf{\Theta}}^C = (\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_L)$ in the first model; the final tree $T_{\hat{\alpha}}^P$ with J (J<M) terminal nodes (j=1,…,J) and the estimated participation rate $\hat{\mathbf{\Theta}}^P = (\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_J)$ in the second model. Nonresponse adjustment factors are calculated as the inverse of the estimated contact rates and the inverse of the estimated participation rates obtaining the adjustment factors in the same way as expressed in formula (21).

4. THE NONRESPONSE ADJUSTMENT IN THE ITALIAN DISABILITY SURVEY

This section describes the application of the nonresponse adjustment methods sketched in Section 3 to the Disability survey (ISTAT, 2012). Firstly the main characteristics of the survey are illustrated, together with a discussion of the nonresponse issues, highlighting the peculiarity of the nonresponse components. Our study has been carried out by comparing the traditional and sequential approaches to nonresponse, since the peculiarities of this survey are suitable for a sequential formulation of the nonresponse adjustment treatment, exploiting the rich auxiliary information, available from the Health survey (ISTAT, 2008).

Nonresponse is analysed through: i) a descriptive analysis of the different sub-sets of sample units individuated by the response outcome of the dichotomous variables R, C and P (respondent/not respondent, contacted/not contacted and, conditional to contact, participating/not participating); ii) the assessment of bivariate relationships between the three response variables and the auxiliary variables, using the Cramer’s V statistic; iii) multivariate analyses through logit and classification models. Furthermore, for the sequential approach, the relationship between contact probability and participation probability, estimated on the basis of the logit models, is examined, with the aim of checking whether contact probabilities and participation probabilities can be considered linearly independent, since independence assumptions are the basis of the sequential approach to nonresponse treatment. Finally, the outcomes of models have been utilised for constructing the adjustment factors. The performance of the various methods has been evaluated in terms of goodness of fit and of the impact of the variability of final weights on the variance of the estimates.

It is important to underline that, despite the fact that the Disability survey investigates a large number of phenomena, the final weight based on the nonresponse adjustment factors are obtained exploiting the relation between response and the available auxiliary variables, not taking into account the survey target variables.

4.1 THE SURVEY AND THE NONRESPONSE CONTEXT

The 2010 survey on “Social integration of people with disabilities” (ISTAT, 2012) is part of the project “System of Statistical Information on Disability” based on an agreement between the Italian National Institute of Statistics and the Ministry of

Labour and Social Policy. The project aims to develop a system of indicators that, through several sources of institutional data, is able to monitor disability in Italy and to provide support for related social policies planning.

Collected information refer to the social integration of people with disabilities (networks of relationships, school, work, leisure, etc.), focusing on the factors preventing full participation (limitations in mobility, lack of adequate support, etc.). The survey concerned the sub-set of individuals belonging to the sample of the former survey "Health conditions and use of health services" (ISTAT, 2008) carried out in 2004-2005, who reported at that time to have difficulties in functions (motor, sensory or related to activities of daily living) or to be suffering from a disability or from reduced independence.

The 2004-2005 Health survey was a two-stage sample survey: the primary sampling units (PSUs) were the Italian municipalities, while the final sample was constituted by households, randomly selected from the population registers of sampled PSUs; each member of the sample households was interviewed by PAPI (paper and pencil interview). The selected sample consisted in approximately 60,000 households in nearly 1,400 municipalities.

The 2010 Disability survey was carried out on the set of units corresponding to the subclass of disabled people in the 6-80 age group of the former Health survey. A contact with all this people was tried in 2010 and respondents were interviewed by CATI (computer assisted telephone interview). The interview was administered to a family member or any other person who took care of the disabled person (proxy) in all cases in which the disabled person was unable to respond to the interview and for disabled children under 14 years.

The initial sample of the Disability survey (2010) is constituted by 3,502 individuals with serious functional limitations and 7,482 individuals with mild functional limitations. The size of the original sample from the former Health survey has been reduced because a part of the individuals formerly identified as disabled, were no longer classified as such, becoming ineligible units (758). The actual sample was therefore composed of 2,744 individuals with severe functional limitations and 6,293 individuals with mild functional limitations. The study of this paper refers to the first group: the strictly disabled.

The Disability survey suffered from a high nonresponse rate, due to the failure of contact rather than to the refusal to participate expressed by contacted individuals. The high proportion of not contacted units came from the combination of two main factors: first, the Health survey did not collect the telephone number for all units, or the information supplied was wrong; and secondly, at the time of the survey, some disabled individuals were no longer available at the formerly provided number because it changed or was discontinued.

The latter fact is due to the lag between the reference survey and the call back on disabled people but also to a critical feature of CATI surveys. Surveys based on telephone interviews suffer by a fall of responses due to the decreasing coverage of the fixed telephone network, which especially occurs for particular population groups, which affect the representativeness of the sample with respect to the target population.

Table 1 shows the response results, highlighting the partition of the effective sample of disabled people in respondent and not respondent units, on one hand, as a premise to the traditional approach, and (Table 2) in contacted and not contacted, being the latter subdivided in participant units and refusals, on the other hand, useful for the sequential approach. The number of not contacted units is very high (1,290 units), while the number of refusals is limited (340 units). Overall nonresponse rate is very high (59.4%) and is due to noncontact in 47.0% of cases, according to the traditional approach perspective.

Table 1: Nonresponse in a traditional approach perspective

Outcome	Number of cases	Rate
Respondent	1114	40.6%
Not respondent units (not contacted + refusals)	1630	59.4%
Effective sample	2744	100%

Table 2: Type of nonresponse in a sequential approach perspective

Sequential stage	Outcome	Number of cases	Rate
First stage	Not contacted	1290	47.0%
	Contacted units	1454	53.0%
	Effective sample	2744	100%
Second stage	Participant units	1114	76.6%
	Refusals	340	23.4%
	Contacted units	1454	100%

4.2 RESPONSE ANALYSIS AND MODELS

Nonresponse in the Disability survey has been deepened by considering the distributions by age and geographical area of the sample units first in two groups according to the traditional approach (respondent and not respondent) and then into the nested groups (not contacted, contacted, participating and not participating after contact) according the sequential approach.

Tables 3 and 4 show the distributions by age class and geographic area and point out remarkable differences when comparing not respondents with refusals highlighting substantial dissimilarities between the two groups (for example, refusals are mostly concentrated in North areas and in the 61-70 ages).

Table 3: Distribution of disabled individuals by response group and age class (row percentages)

Group		Age class								
		06-10	**11-20**	**21-30**	**31-40**	**41-50**	**51-60**	**61-70**	**71-80**	**Total**
Traditional approach perspective										
Respondent		1.08	1.35	2.96	5.12	6.82	10.50	23.25	48.92	100%
Not respondent units (not contacted + refusals)		4.97	2.70	2.64	5.71	4.66	8.90	20.80	49.63	100%
Sequential approach perspective										
First stage	Not contacted units	5.66	2.48	2.79	6.05	4.88	9.53	18.99	49.61	100%
	Contacted units	1.38	1.86	2.75	4.95	6.12	9.56	24.28	49.11	100%
Second stage	Participant units	1.08	1.35	2.96	5.12	6.82	10.50	23.25	48.92	100%
	Refusals	2.35	3.53	2.06	4.41	3.82	6.47	27.65	49.71	100%

Table 4: Distribution of disabled individuals by response group and geographic area (row percentages)

Group		Geographic Area					
		North-West	North-East	Centre	South	Islands	Total
Traditional approach perspective							
Respondent		18.58	16.97	17.50	33.93	13.02	100%
Not respondent units (not contacted + refusals)		17.42	13.37	16.93	37.98	14.29	100%
Sequential approach perspective							
First stage	Not contacted units	15.89	12.64	17.29	39.22	14.96	100%
	Contacted units	19.67	16.78	17.06	33.77	12.72	100%
Second stage	*Participant units*	18.58	16.97	17.50	33.93	13.02	100%
	Refusals	23.24	16.18	15.59	33.24	11.76	100%

The relationship between the outcome of the indicator variables defined in par. 3.2 and 3.3, respectively denoting, for unit *i-th*, response **R**, contact **C** and participation **P**, and the auxiliary variables coming from the information collected in the Health survey is further investigated by means of the Cramer's *V* statistic, logistic regression models and CART models.

The auxiliary variables considered in these analyses are: telephone number provided in the first survey (yes/no), household size (between 1 and 8), age, occurrence of motor difficulty (yes/no), level of education (11 levels), number of invalidities (5 levels), number of disabilities (3 levels), disability level (3 categorical level), sex, occurrence of difficulty of sight, hearing or speech, marital status (6 levels) and occurrence of difficulties in daily life functions (yes/no).

The Cramer's *V* statistic is a bivariate test of independence based on χ^2 which takes into account the different degrees of freedom of the categorical variables and the sample size (Bethlehem et al., 2011). Here the variables *age* (≤12, 13-21, 22-75, >75) and the values of the *education level* (below and above school leaving certificate) and *marital status* (married/not married) were aggregated using the CART classification algorithm.

This preliminary analysis leads to important results highlighting that the contact itself **C** is more closely linked to the characteristics of units than participation **P**.

Table 4 shows the results of the test that associates indicator values with the auxiliary variables above: as far as response **R** is concerned, the strongest bivariate relationship is found for the telephone variable (*V*=0.296), followed by the household size. For the contact variable, **C**, the bivariate relationship with the telephone variable is even stronger (*V*=0.367). Regarding to participation, **P**, the values of Cramer's *V* statistics are overall low.

Table 5: Tests of bivariate dependence between auxiliary variables and response, contact and participation indicators (Cramer's V statistic)

Auxiliary variables	R	C	P
Telephone	0.296	0.367	0.026
Household size	0.137	0.143	0.075
Age class	0.137	0.125	0.135
Motor difficulty	0.078	0.088	0.021
Level of education	0.091	0.075	0.093
Number of invalidities	0.090	0.071	0.092
Number of disabilities	0.066	0.067	0.059
Disability level	0.068	0.057	0.060
Sex	0.016	0.027	0
Difficulty of sight, hearing, speech	0	0	0.016
Difficulties in daily life functions	0	0	0.010
Marital status	0	0	0

In the following analysis, the auxiliary variables taken into account were those listed above plus the geographical area. All these auxiliary variables were differently recoded in order to obtain the best fitting of the models.

Model assessment was performed, in the parametric case, via the Akaike Information Criterion (AIC), i.e. a goodness of fit criterion, and, for classification trees, via the "cost-complexity" function (22), i.e. $\Phi_\alpha(T)$.

Table 6 summarises the best resulting models with the most influential covariates and the values of the AIC for the logit model and of $\Phi_\alpha(T)$ for the CART model, respectively for the three outcome variables, response **R**, contact **C** and participation **P**.

Table 6: Response, contact and participation: logit and CART models assessment

	Traditional approach		**Sequential approach**			
	Response		**Contact**		**Participation**	
Model	**Covariates**	**Index**	**Covariates**	**Index**	**Covariates**	**Index**
Logit Model AIC	*telephone* *age in 4 classes* *marital status* *disability level* *motor difficulty* *number of invalidities*	3.388	*Telephone* *age in 2 classes* *marital status* *motor disability* *number of invalidity* *number of disability*	3.347	*age in 5 classes*	1.564
CART $\Phi_\alpha(T)$	*telephone* *age in 4 classes* *difficulties in daily life functions*	0.406	*Telephone*	0.325	*age in 3 classes*	0.249

Response models. In response models (traditional approach), the variable denoting the availability of a telephone number is the most influential for both models. The other relevant variables are the four age classes (≤12, 13-21, 22>75, >75), the marital status (married or not), the disability level, the presence or motor difficulty and the number of invalidities. The AIC of the model assumes the value 3,388. In the model based on the CART algorithm (in which the variable age is classified as in logit model), the most influential predictors on the dependent variable **R** are broadly similar and the cost function $\Phi_\alpha(T)$ is 0.406.

Contact models. The contact models regard the first stage of the sequential approach: in the logit models the significant covariates are broadly the same of the response model, but the age has been recoded into two classes ((≤12, >12). For the logit, the AIC assumes the value 3.347. For the CART model, where the only split variable is the telephone; the cost function is 0.325.

Participation models. In the logit model for estimating the participation probability (second stage in the sequential approach), the only explanatory variable is age expressed in 5 classes (≤21; 22-55; 56-59; 60-77; >77); the value of AIC is 1,564. The only predictor of classification model is age in three classes (≤21; 22-59; >59); the cost function $\Phi_\alpha(T)$ assumes the value 0.249.

In order to continue in assessing the appropriateness of the sequential approach, the estimates of the logit models for contact and participation probability are jointly considered, following Olson (2006). Table 7 reports the quartiles of the joint distribution of the estimated propensities.

Table 7: Joint percentage distribution of the quartiles of contact and participation propensities

		Participation propensity				
	quartiles	1°	2°	3°	4°	Total
Contact propensity	1°	11	5.03	5.57	3.5	25
	2°	5.39	7.72	7.09	4.85	25
	3°	4.94	5.57	6.1	8.44	25
	4°	3.77	6.64	6.19	8.26	25
	Total	25	25	25	25	100

We do not perform a formal test, but, since the total percentage falling in the main diagonal is slightly higher than 33%, meaning that only for this fraction a strong linear relationship exists between contact and participation probabilities, while most units do not fall in the main diagonal.

The conjecture of a non linear relationship between the two probabilities is strengthened by the checks of Table 8.

Table 8: Results of independence assessment

Statistics	Value	p-value	Conclusion
χ^2chi-square	0.000	1	No independence
ρ(Pearson linear correlation)	0.012	0.41	No linear correlation
ρ(Spearman correlation)	0.395	1	Nonlinear relationship

4.4 ADJUSTMENT FACTORS

The nonresponse correction factors for obtaining the final weights for the Disability survey participants were calculated following the methods described in Section 3.

According to the logit response model, following the traditional approach, adjustment cells were defined on the basis of quartiles, quintiles and deciles of the distribution of the predicted response propensities. Analogously, in the first stage of the sequential approach, for the logit contact model the cells were defined through the quartiles, quintiles and deciles of the distribution of the predicted contact propensities, while for the second stage the cells were defined only on the basis of quartiles of the distribution of the predicted participation propensities. A further type of correction factor was determined, for each logit model, as the inverse of the predicted probabilities (response propensity weighting).

According to the response CART model, always following the traditional approach, the final tree contains four terminal nodes (cells) as follows (Figure 1); the first cell contains the individuals who provided the telephone number, the second the individuals aged between 22 and 75 who have not provided the telephone number, the third the individuals younger than 22 or older than 75, which have no difficulty in performing functions of daily life, and the fourth cell the individuals of the same age classes but with difficulties.

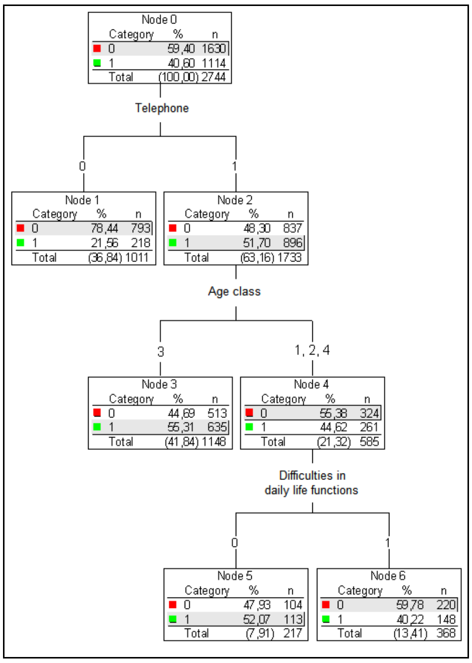

Figure 1: Response classification tree (traditional approach)

In the contact CART model (first stage of sequential approach), the terminal nodes (cells) of the final tree depend on having provided or not a telephone number in the Health survey (Figure 2). Finally, for the participation CART model (second stage of sequential approach), the classification procedure leads to three terminal nodes (Figure 3); the first cell focuses on individuals aged up to 21 years, the second on those aged between 22 and 59 and the third on those over 60 units.

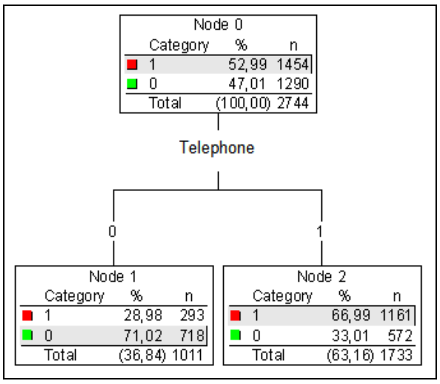

Figure 2: Contact classification tree (sequential approach – first stage)

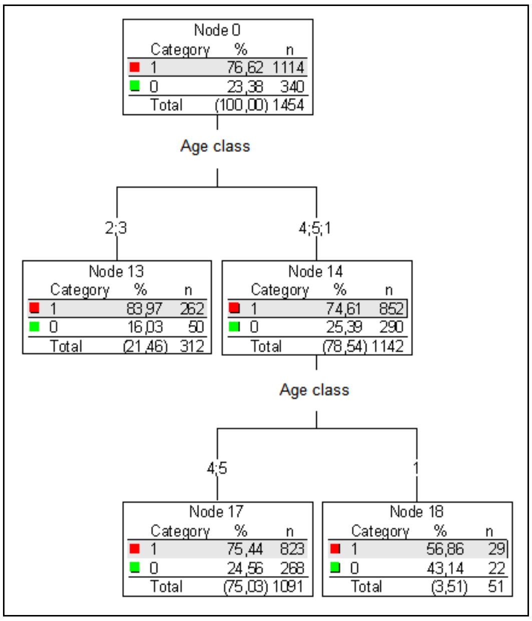

Figure 3: Partecipation classification tree (sequential approach – second stage)

4.5 RESULTS ASSESSMENT

All results were compared in two different ways to evaluate, on one hand, the goodness of fit obtained according to the estimation methods and, on the other hand, to measure the impact of the variability of final weights on the variance of the estimates. For goodness of fit, a concordance index (*CI*) was calculated as

$$CI = 1 - \frac{\sum_{i=1}^{n} \left| obs_i - \hat{\theta}_i \right|}{n} . \tag{23}$$

Index (23) can be expressed in various forms, which are summarised in Table 9, where $\hat{\theta}_i$ indicates the estimated probability of response and n is the number of observations in sample s, and, alternatively, $obs_i = R_i$ indicates the response variable observed on the i-th unit in the sample s, $obs_i = C_i$ the observed value of the contact variable and $obs_i = P_i$ the observed participation value (in this case n is replaced by the number of contacted persons n_C). Index (21) always shows higher values for the contact model than for the response model, while the goodness of fit of the participation model is always significantly better.

The evaluation of the impact of variability of final weights is based on the statistics $1+CV^2(w)$ (where $CV(w)$ is the coefficient of variation of the generic weight w) proposed by Kish (1992), which keeps into account an approximated increase of the variance of the estimates due to the application of correction factors to the sampling weights. This synthetic measure permits to appreciate the extent to which the proposal of adjustment factors produces an increase of the variability in the weights and thereby a reduction of the precision of the survey estimates, with respect to the situation of equal weights.

Table 9: Concordance index

Model	Approach	Technique	Concordance Index (*CI*)		
			Response	Contact	Participation
LOGIT	Response propensity stratification	Quartiles	0.569	0.574	0.645
		Quintiles	0.569	0.581	
		Deciles	0.573	0.584	
	Response propensity weighting	Individual propensity	0.565	0.569	0.647
CART		Terminal	0.574	0.583	0.648

The following Tables 10 and 11 summarise the distributions of the final weights obtained through the different techniques listed in Table 9, together with the statistics $1+CV^2(w)$, respectively according to the traditional and sequential approach.

Table 10: Summaries of the distributions of final weights (Traditional approach)

Model	Approach	Technique	Average	Max	Min	$CV(w)$	$1+CV^2(w)$
LOGIT	Response propensity stratification	Quartiles	1046.72	7692.57	98.83	0.825	1.680
		Quintiles	1037.98	8861.92	99.02	0.821	1.673
		Deciles	1037.62	9781.18	89.22	0.855	1.731
	Response propensity weighting	Individual propensity	1022.55	7235.38	94.09	0.784	1.615
CART		Terminal nodes	1035.76	6796.77	94.09	0.753	**1.567**

Table 11: Summaries of the distributions of the final weights (Sequential approach)

Model	Approach	Technique	Average	Max	Min	$CV(w)$	$1+CV^2(w)$
LOGIT	Response propensity stratification	Quintiles	1028.87	7081.31	104.13	0.732	1.555
	Response propensity weighting	Individual propensity	1027.73	7350.38	101.51	0.732	1.555
CART		Terminal nodes	1026.71	7003.45	102.98	0.729	**1.531**

From both tables, the CART algorithm produces the final set of weights with the lowest value of the statistic $1+CV^2(w)$. Moreover, whatever the methods proposed for the adjustment factors, the sequential approach performs much better than the traditional one.

7. CONCLUSIONS

A new proposal for sequential nonresponse treatment produced final weights with better performances when compared with the traditional one shot method; this performance was evaluated both for the impact on the weight variability and in terms of bias reduction. Actually, in the sequential approach two different models, parametric and not parametric, are considered, both exhibiting a goodness of fit better than the corresponding model for response (traditional approach), providing an indication about a stronger bias reduction. Furthermore, the procedure based on two sequential adjustment factors produces an additional gain due to a decrease in the variability of the sample weights. Finally, what emerges is also the good performance of the CART model also when proposed in nested mode in the sequential perspective.

REFERENCES

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

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth International Group, Belmont.

Cochran, W.G. (1977). *Sampling Techniques*. Wiley, New York.

Groves, R.M., and Couper, M.P. (1998). *Nonresponse in Household Interview Surveys*. Wiley, New York.

Iannacchione, V.G. (2003). Sequential weight adjustments for location and cooperation propensity for 1995 national survey of family growth. In *Journal of Official Statistics*, 19: 31-43.

ISTAT (2008). *Indagine Multiscopo Condizioni di salute e Ricorso ai Servizi Sanitari* http://www.istat.it/it/archivio/10836.

ISTAT (2012). *Inclusione Sociale delle Persone con Limitazioni dell'Autonomia Personale*. Statistiche report. http://www.istat.it/it/archivio/77546.

Kalton, G., and Flores-Cervantes, I. (2003). Weighting methods. In *Journal of Official Statistics*, 19: 81-97.

Kish, L. (1992). Weighting for unequal Pi. In *Journal of Official Statistics* 8: 183-200.

Little, R.J.A. (1986). Survey nonresponse adjustments for estimates of means. In *International Statistical Review*, 54: 139-157.

Little, R.J.A., and Vartivarian S. (2005). Does weighting for nonresponse increase the variance of survey means? In *Survey Methodology*, 31: 161-168.

Oh, H.L., and Scheuren, F. (1983). Weighting adjustments for unit nonresponse. In W.G. Madow, I. Olkin, and D.B. Rubin, eds., *Incomplete Data in Sample Surveys, Vol. 2, Theory and Bibliographies*, 143-184. Academic Press, New York.

Olson, K.M. (2006). *Survey Participation, Non Response Bias, Measurement Error Bias, and Total Bias*. Sociology Department, Faculty Publications, University of Nebraska, Lincoln.

Rizzo, L., Kalton, G., and Brick, J.M. (1996). A comparison of some weighting adjustment methods for panel nonresponse. In *Survey Methodology*, 22: 43-53.

Rosenbaum, P.R., and Rubin, D.B. (1984). Reducing the bias in observational studies using subclassification on the propensity score. In *Journal of the American Statistical Association*, 79: 516-524.

Särndal, C.E., and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. Wiley, New York.

Särndal, C.E., Swensson, B., and Wretman, J.H. (1992). *Model Assisted Survey Sampling*, Cap XV, Springer. New York.