Survey Research Methods (2007)
Vol.1 , No.2 , pp. 85-96
ISSN 1864-3361
http://www.surveymethods.org

Errors Depending on Costs in Sample Surveys

Daniela Marella
Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi di Roma

This paper presents a total survey error model that simultaneously treats sampling error, nonresponse error and measurement error. The main aim for developing the model is to determine the optimal allocation of the available resources for the total survey error reduction. More precisely, the paper is concerned with obtaining the best possible accuracy in survey estimate through an overall economic balance between sampling and nonsampling error.
Keywords: Sampling error, nonsampling error, nonresponse error, measurement error, total survey error model, cost model, total survey design.

Introduction

By "error in survey results" it is meant the difference between the survey estimate and the value to be estimated. Error in survey estimates is traditionally divided into two major categories: sampling error and nonsampling error. Sampling error occurs because only part of an entire population is studied: data are collected from a sample to draw conclusions about the population. Nonsampling error encompasses all other factors that contribute to the total error of a sample survey estimate, arising from deficiencies or mistakes in the survey process. This error source may also be present in censuses and may occur because of nonresponse, errors in sampling frame, mistakes in recording and coding of data, and other errors of collection, response and processing. Thus, sampling error measured by the sampling variance represents just the lower bound of the total error, achieved under the rather idealistic assumption that each sample unit gives the requested information without errors. As a consequence, a relative large sampling variance does not condemn all aspects of a given survey, just as a small sampling variance does not by itself assure a good data quality. Since in most surveys the sampling error may be small compared to nonsampling errors, their estimate is definitely important to asses the accuracy of the information being collected. Accuracy relates to the quality of survey results and it is distinguished from precision. Precision denotes only the inverse of the sampling variance, accuracy is "the inverse of the total error, including bias as well as the variance" (Kish 1965:25).

Understanding the causes and the prevention of nonsampling errors, through social science theories, is an essential step both to identify faulty operations that are in need of improvement and for effective error reduction. The next step should be to translate these human behavior theories into models of statistical error (Groves 1999). Total survey error models, whose objective is to measure the relative impact of each error source on survey estimates and to make probability statements about the total error, were developed in a series of important articles by Hansen et al. (1951, 1961, 1965). Kish (1965) proposed a very general model that decomposes the total error into fixed biases and variable errors. Such a model sufficiently general, must be made more specific in order to be useful in an actual survey. This model was used by Andersen et al. (1979) in their exploration of errors in a survey of health services use. More specifically, the model was focused on three components of nonsampling error: nonresponse bias, measurement bias and processing bias limited to imputation bias. Theoretical and empirical contributions on various types of nonsampling errors that can occur in surveys and on related costs could be found in Groves (1989) and Weisberg (2005).

As indicated by Lessler and Kalsbeek (1992), that proposed a model with sampling, nonresponse, measurement and frame errors, total error models are preliminary steps that must be more highly specified in an actual survey situation.

Forsman (1989, 1993) reviews the history of the survey error model theory, in particular its last 60 years. He notes that even though theory for specific sources of nonsampling error has had a positive development, we are far from realizing an integrated treatment of survey errors within one model, and above all connecting them to the budgetary conditions. The construction of total survey error models aims to translate the complex sequence of survey operations into a mathematical statement. As a consequence, the implementation of a total survey error approach involves a careful balance of results from mathematical statistics and empirical studies. More specifically, mathematical statistics provides the essential underlying framework while empirical studies allows us some preestimation of parameters appearing in the nonsampling errors models (Andersen et al. 1979).

This paper presents a total survey error model that simultaneously treats sampling error, nonresponse error and measurement errors. We ignore coverage error and data processing error. The former is due to the lack overlap between

Contact information: Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi di Roma, La Sapienza, P.le A.Moro 5, 00185 Roma, Italy (Daniela.Marella@uniroma1.it)

sampling frame and target population. The latter comprises errors arising from coding, imputation, keying, editing, and tabulating the survey data.

The main aim for developing the model comes from the desire to define a total survey design minimizing the total error, which can be implemented with costs that are consistent with the available budget. More precisely, the paper focuses on the task of obtaining the best possible accuracy in survey estimate through an overall economic balance between sampling and nonsampling error. The purposes of paper are twofold: (i) quantification of the total survey error using a model-based approach; (ii) study of the optimal allocation of the available budget for the total survey error reduction.

As far as the first point is concerned, we compute the Mean Square Error (MSE) for the model recognizing the existence of nonsampling error component in addition to sampling variance. In order to accomplish this, we introduce model assumptions to describe both how observations drop out due to nonresponse and measurement errors generating mechanism.

As far as the second point is concerned, we note that a survey involves a trade-off between survey costs and errors, increasing one reduces the other. The existence of this trade-off implies, conditionally on funds available, an inverse relation between sampling and nonsampling error. That is, given a budget to carry out the survey, the larger the observational process accuracy the larger the cost of measuring each of the sample cases. Hence, the lower the sample size with consequent increase of sampling variance. In other words, the larger the sample size the poorer the resources allocated for nonsampling errors reduction (Tranquilli 1995). Since we analyze a model with sampling error, nonresponse error and measurement error the question is how to allocate the available budget between reduction of sampling variance, maximization of response rate and minimization of measurement errors. As a matter of fact, in order to define a total survey design it is essential to investigate the relationship between survey errors and survey costs. For instance, the magnitude of error reduction achieved through the use of qualified interviewers, the interviewers training, the choice of data collection mode, the number of callbacks and so on, must be evaluated through empirical studies. Clearly, the impact of these actions depends heavily on combined factors such as: the characteristics of the underlying population, the survey topic.

The paper is organized as follows. In section 1 a hypothetical total survey error model that treats sampling error, nonresponse error and measurement errors is introduced, and the MSE is computed under a stratified sampling design. It is important to stress that the total survey error model does not come from an actual survey situation, but represents a preliminary step to answer to the following questions: (i) which parameters in total survey error model depend on survey costs (ii) how we can formalize the cost-error tradeoff for arriving at a total survey design (iii) how we can estimate the parameters appearing in the reparametrization model, connecting the nonsampling error to survey costs. Section 2 deals with the cost model depending on both the strata sample sizes $\underline{n} = (n_1, .., n_M)$ and strata per-unit costs $(\underline{c}^{(1)}, \underline{c}^{(2)}) = [(c_1^{(1)}, .., c_M^{(1)}), (c_1^{(2)}, .., c_M^{(2)})]$, that influence strata nonresponse rates and strata measurement errors respectively. Since the best possible accuracy in survey estimate is achieved by minimizing the total error subject to a fixed cost, we must solve a constrained optimization problem. In order to accomplish this, we introduce in section 3 the reparametrization model concept connecting the error components to the budgetary conditions. In more detail, while sampling error is related to cost model by the strata sample sizes, nonsampling error is not, since the strata cost variables $(\underline{c}^{(1)}, \underline{c}^{(2)})$ do not appear in the MSE expression. Hence, to formalize the cost-error tradeoff we introduce model assumptions associating the nonsampling error sources with the cost components.

Section 4 describes how cost and reparametrization models are used together with the total error model to find the optimal allocation of the available resources. In this section we distinguish two different approaches. The former deals with the optimal allocation problem assuming that $(\underline{c}^{(1)}, \underline{c}^{(2)})$ are fixed. The latter approach, called unconditional approach, achieves the best accuracy in survey estimate not conditional on the strata per-unit costs. More specifically, given $(\underline{c}^{(1)}, \underline{c}^{(2)})$ the response probability and the magnitude of measurement errors within the strata are given too. Hence, the optimal allocation problem can be formulated as the determination of the strata sample sizes minimizing the total error. In section 4.1 we discuss a possible application of the total error model developed in section 1 to business surveys. Finally, in section 5 we assess the limitations and the extensions of the unconditional approach.

1 Total Error Model

Let U be a finite population partitioned in M nonoverlapping subpopulations $U = \{U_1, .., U_g, .., U_M\}$ called strata and denote by N_g the number of population elements in stratum g, for $g = 1, .., M$. Consider a single study variable θ, and suppose that an estimate is needed for the population total t_θ. Assume that a probability sample $\underline{s}_{n_g}$ of size n_g is selected from U_g according to a simple random sampling without replacement (for $g = 1, .., M$), and that the selection in one stratum is independent of the selections in all other strata. The resulting total sample $\underline{s}_n$ will thus be composed as

$$\underline{s}_n = (\underline{s}_{n_1} \cup \underline{s}_{n_2} \cup ... \cup \underline{s}_{n_M}) \quad \text{where} \quad n = \sum_{g=1}^{M} n_g$$

However, nonresponse occurs in the survey process and the response set

$$\underline{s}_r = (\underline{s}_{r_1} \cup \underline{s}_{r_2} \cup ... \cup \underline{s}_{r_M}) \quad \text{where} \quad r = \sum_{g=1}^{M} r_g$$

of size r is obtained. Since we do not ordinarily know how the strata response sets are generated we must make model assumptions about the true unknown response distribution.

In order to accomplish this we apply the theory for the two-phase sampling with estimated response probabilities as second phase inclusion probabilities. One such model is the *response homogeneity groups model* (RHGs). For each stratum g, the realized sample $\underline{s}_{n_g}$ can be partitioned into $H_{\underline{s}_{ng}}$ response groups $\underline{s}_{n_{gh}}$ indexed by $h = 1, 2, ..., H_{\underline{s}_{ng}}$. Given $\underline{s}_{n_g}$, all elements within one and the same group $\underline{s}_{n_{gh}}$ respond with the same probability and in an independent manner (Särndal et al. 1992). The auxiliary information required here is that we can uniquely classify every sampled element within a given stratum, respondent or nonrespondent, into one of the $H_{\underline{s}_{ng}}$ groups. Note that the grouping need not be the same for different samples, hence the subscript $\underline{s}_{n_g}$ in $H_{\underline{s}_{ng}}$.

Let $\underline{s}_{r_{gh}}$ be the response set of size r_{gh} in the *gh-th* RHG $\underline{s}_{n_{gh}}$, generated as the result of n_{gh} independent Bernoulli trials with constant probability α_{gh} of success that is, response. For a given $\underline{s}_{n_g}$, the response model we assume formalizes "the response mechanism" within each response group $\underline{s}_{n_{gh}}$ as a Bernoulli sampling design. Hence, the response set $\underline{s}_{r_g}$ of size r_g in the *g-th* stratum will be composed as

$$\underline{s}_{r_g} = (\underline{s}_{r_{g1}} \cup ... \cup \underline{s}_{r_{gH_{\underline{s}_{ng}}}}) \quad \text{where} \quad r_g = \sum_{h=1}^{H_{\underline{s}_{ng}}} r_{gh}$$

In practice to handle the problem of nonresponse a two-phase sampling was adopted: the first-phase sample $\underline{s}_n$ is selected by a simple stratified random sampling without replacement from population U; for each given $\underline{s}_{n_g}$, the response set $\underline{s}_{r_g}$ is distributed in accordance with the stratified Bernoulli sampling design.

Besides, suppose that the data collection operations generate errors in the individual data so that the observed value will differ from the true value $\theta_{i_{gh}}$, for each element i_{gh} in the response set $\underline{s}_{r_{gh}}$. It is assumed that, under the general conditions of the survey, the measurement for the i_{gh}*-th* individual varies over repeated trials of the survey measurement process. Under this perspective the survey at hand is only one of an infinite number of possible replications of the survey design, and the measurement for each element i_{gh} is a random variable given by

$$\tilde{y}_{i_{gh}} = \theta_{i_{gh}} + \tilde{\eta}_{i_{gh}} = \theta_{i_{gh}} + \delta_{i_{gh}} + \tilde{\epsilon}_{i_{gh}} \tag{1}$$

$\forall i_{gh} \in \underline{s}_{r_{gh}}, \forall g \in (1, .., M), \forall h \in (1, .., H_{\underline{s}_{ng}})$. The observed value $\tilde{y}_{i_{gh}}$ for the i_{gh}*-th* respondent is composed of the true value $\theta_{i_{gh}}$ and an error term $\tilde{\eta}_{i_{gh}} = \delta_{i_{gh}} + \tilde{\epsilon}_{i_{gh}}$ (Särndal et al. 1992) The measurement error $\tilde{\eta}_{i_{gh}}$ comes from the sum of a systematic error $\delta_{i_{gh}}$, constant over repeated trials of the survey, and a random error $\tilde{\epsilon}_{i_{gh}}$ such that

$$\begin{cases} E_\epsilon(\tilde{\epsilon}_{i_{gh}}|\underline{s}_{r_{gh}}) = 0 \\ Var_\epsilon(\tilde{\epsilon}_{i_{gh}}|\underline{s}_{r_{gh}}) = \sigma^2_{\epsilon_g} \\ Cov_\epsilon(\tilde{\epsilon}_{i_{gh}}, \tilde{\epsilon}_{j_{gh'}}|\underline{s}_{r_{gh}}, \underline{s}_{r_{gh'}}) = \rho_{\epsilon_g}\sigma^2_{\epsilon_g} \quad \forall g, \quad \forall h, h' \\ Cov_\epsilon(\tilde{\epsilon}_{i_{gh}}, \tilde{\epsilon}_{j_{g'h}}|\underline{s}_{r_{gh}}, \underline{s}_{r_{g'h}}) = 0 \quad \forall g \neq g' \end{cases} \tag{2}$$

where $E_\epsilon(.|\underline{s}_{r_{gh}})$ and $Var_\epsilon(.|\underline{s}_{r_{gh}})$ denote the conditional expectation and the conditional variance over all possible trials respectively, for any given response set $\underline{s}_{r_{gh}}$. Note that $\sigma^2_{\epsilon_g}$ is the variation between repeated measurements on any unit in the *g-th* stratum, and ρ_{ϵ_g} is the correlation between measurements on any two units within the same stratum g.

In the measurement error we distinguish the systematic measurement error from random measurement error. Systematic errors are biases that consistently affect the measurement process no matter what time the interview is conducted: we obtain the same results using measures in different occasions. Of the two types of measurement errors, systematic errors are the most serious but also the most controllable. In general, efforts are focused on reducing systematic errors concern with the choice of data collection mode (mail, telephone, face to face interview, administered questionnaire, etc), as well as a good questionnaire wording and field personnel training. Clearly all these factors have cost implications. We specify further the model (2)

$$\begin{cases} E_\epsilon(\tilde{y}_{i_{gh}}|\underline{s}_{r_{gh}}) = \theta_{i_{gh}} + \delta_{i_{gh}} = \mu_{i_{gh}} \\ Var_\epsilon(\tilde{y}_{i_{gh}}|\underline{s}_{r_{gh}}) = \sigma^2_{\epsilon_g} \\ Cov_\epsilon(\tilde{y}_{i_{gh}}, \tilde{y}_{j_{gh'}}|\underline{s}_{r_{gh}}, \underline{s}_{r_{gh'}}) = \rho_{\epsilon_g}\sigma^2_{\epsilon_g} \quad \forall g, \quad \forall h, h' \\ Cov_\epsilon(\tilde{y}_{i_{gh}}, \tilde{y}_{j_{g'h}}|\underline{s}_{r_{gh}}, \underline{s}_{r_{g'h}}) = 0 \quad \forall g \neq g' \end{cases} \tag{3}$$

asserting that, for any given $\underline{s}_{r_{gh}}$, the measurement $\tilde{y}_{i_{gh}}$ on element i_{gh} has mean $\mu_{i_{gh}}$, variance $\sigma^2_{\epsilon_g}$ and the covariance between elements within the same stratum g is given by $\rho_{\epsilon_g}\sigma^2_{\epsilon_g}$. Since we consider a partition of U into M strata, the following decomposition holds

$$t_\theta = \sum_{g=1}^{M} \overline{\theta}_g N_g \tag{4}$$

where $\overline{\theta}_g$ denotes the stratum mean. Consider as an estimator of $\overline{\theta}_g$

$$\tilde{\overline{y}}_{r_g} = \sum_{h=1}^{H_{\underline{s}_{ng}}} \left(\frac{1}{r_{gh}} \sum_{i_{gh} \in \underline{s}_{r_{gh}}} \tilde{y}_{i_{gh}} \right) \nu_{gh} = \sum_{h=1}^{H_{\underline{s}_{ng}}} \tilde{\overline{y}}_{r_{gh}} \nu_{gh} \tag{5}$$

where $\nu_{gh} = n_{gh}/n_g$ is the relative size of the response group $\underline{s}_{n_{gh}}$. The expression within parentheses is the weighting estimator of $\overline{y}_{n_{gh}}$, the mean of n_{gh} units in $\underline{s}_{n_{gh}}$. The nonresponse compensation adjustment weight is the inverse of the estimated response probability $\hat{\alpha}_{gh} = r_{gh}/n_{gh}$. Hence, the estimator of t_θ is given by

$$\tilde{\hat{t}}_\theta = \sum_{g=1}^{M} \left(\sum_{h=1}^{H_{\underline{s}_{ng}}} \tilde{\overline{y}}_{r_{gh}} \nu_{gh} \right) N_g \tag{6}$$

In order to analyze the inference not conditionally on $\underline{r} = (\underline{r}_1, .., \underline{r}_g, .., \underline{r}_M)$, with $\underline{r}_g = (r_{g1}, .., r_{gh}, ..r_{gH_{\underline{s}_{ng}}})$, we exclude the event

$$\{r_{gh} = 0 \quad \textit{for some} \quad g = (1, .., M), h = (1, .., H_{\underline{s}_{ng}})\}$$

since in that case the estimator $\tilde{\tilde{t}}_\theta$ is not defined. Denoted by $\underline{1}$ the unitary vector of length $\sum_{g=1}^{M} H_{\underline{s}_{ng}}$, the expected value for $\tilde{\tilde{t}}_\theta$ is given by (see Appendix A)

$$E(\tilde{\tilde{t}}_\theta) = E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon(\tilde{\tilde{t}}_\theta) = \sum_{g=1}^{M} \bar{\theta}_g N_g + \sum_{g=1}^{M} \bar{\delta}_g N_g \quad (7)$$

where $E_\epsilon(.|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n)$, $E_{\underline{s}_r}(.|\underline{r} \geq \underline{1}, \underline{s}_n)$, $E_{\underline{r}|\underline{r}\geq\underline{1}}(.|\underline{s}_n)$ denote conditional expectations with respect to the measurement model, the response model and the response set size $\underline{r}$ (for $\underline{r} \geq \underline{1}$) respectively, while $E_{\underline{s}_n}(.)$ denotes expectation over all possible samples. It follows that the approximate bias of $\tilde{\tilde{t}}_\theta$, given by

$$Bias(\tilde{\tilde{t}}_\theta) = E(\tilde{\tilde{t}}_\theta) - t_\theta = \sum_{g=1}^{M} \bar{\delta}_g N_g \quad (8)$$

is independent of the sample size n and is due to measurement error. Even in a census survey (i.e. when $n = N$), it would remain unchanged. The RHGs model is widely used in practice for modelling nonresponse. "No practitioner really believes that all elements in a group have exactly the same probability to respond, the point is that the assumption of constant probability within well-constructed groups removes most of the nonresponse bias" (Särndal et al. 1992:579).

The approximate variance of $\tilde{\tilde{t}}_\theta$ is given by (see Appendix B)

$$\begin{aligned} Var(\tilde{\tilde{t}}_\theta) \cong & \sum_{g=1}^{M} \left(\frac{1}{n_g} - \frac{1}{N_g}\right)\left[\sigma^2_{\theta_g} + \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g)\right] N_g^2 \\ & + E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} \left\{ \sum_{g=1}^{M} \left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{n_{gh}}{n_{gh}-1} \left(\frac{1}{r_{gh}} - \frac{1}{n_{gh}}\right) s^2_{\mu_{gh}} \right] N_g^2 | \underline{r} \geq \underline{1}, \underline{s}_n \right\} \\ & + E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} \left\{ \sum_{g=1}^{M} \left(\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{\sigma^2_{\epsilon_g}}{r_{gh}} \nu^2_{gh} \right) N_g^2 | \underline{r} \geq \underline{1}, \underline{s}_n \right\} \\ & + E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} \left\{ \sum_{g=1}^{M} \left(\sum_{h=1}^{H_{\underline{s}_{ng}}} (r_{gh}-1) \frac{\rho_{\epsilon_g} \sigma^2_{\epsilon_g}}{r_{gh}} \nu^2_{gh} \right) N_g^2 | \underline{r} \geq \underline{1}, \underline{s}_n \right\} \quad (9) \end{aligned}$$

where $\sigma^2_{\theta_g}$, $\sigma^2_{\delta_g}$ and $Cov(\theta_g, \delta_g)$ represent the variance of true values and systematic errors and their covariance in the g-th stratum respectively, while $s^2_{\mu_{gh}}$ is the sample variance of the expected measurement values in the response group $\underline{s}_{n_{gh}}$.

In the sequel, we assume that the strata used for the sample selection and the response homogeneity groups coincide. This means that we find it plausible that each sampled unit in a given stratum responds with the same probability: the response is random within strata. As a matter of fact, the response homogeneity groups and the strata are not the same. This assumption can be considered plausible for business surveys where the population is often stratified by three auxiliary variables: industry, size and geography. Under this assumption the nonresponse compensation weight is n_g/r_g, then the estimator of t_θ is given by

$$\tilde{\tilde{t}}_\theta = \sum_{g=1}^{M} \tilde{\tilde{y}}_{r_g} N_g \quad (10)$$

where $\tilde{\tilde{y}}_{r_g}$ is the sample mean of the r_g observed values in the g-th stratum. The approximate bias of estimator (10) is given by (8). With regard to the variance, we have (see Appendix C)

$$\begin{aligned} Var(\tilde{\tilde{t}}_\theta) \cong & \sum_{g=1}^{M} \left(E_{\underline{r}|\underline{r}\geq\underline{1}} \left(\frac{1}{r_g}\right) - \frac{1}{N_g} \right) \left[\sigma^2_{\theta_g} + \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g)\right] N_g^2 \\ & + \sum_{g=1}^{M} E_{\underline{r}|\underline{r}\geq\underline{1}} \left(\frac{1}{r_g}\right) \sigma^2_{\epsilon_g} N_g^2 + \sum_{g=1}^{M} E_{\underline{r}|\underline{r}\geq\underline{1}} \left(\frac{r_g - 1}{r_g}\right) \rho_{\epsilon_g} \sigma^2_{\epsilon_g} N_g^2 \quad (11) \end{aligned}$$

Through algebraic calculations the following decomposition holds

$$\begin{aligned} Var(\tilde{\tilde{t}}_\theta) \cong & \sum_{g=1}^{M} \left(\frac{1}{n_g} - \frac{1}{N_g}\right) \sigma^2_{\theta_g} N_g^2 + \sum_{g=1}^{M} \left[E_{\underline{r}|\underline{r}\geq\underline{1}} \left(\frac{1}{r_g}\right) - \frac{1}{n_g} \right] \sigma^2_{\theta_g} N_g^2 \\ & + \sum_{g=1}^{M} \left(\frac{1}{n_g} - \frac{1}{N_g}\right) \left[\sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) + \frac{\sigma^2_{\epsilon_g}}{n_g} \right] N_g^2 \\ & + \sum_{g=1}^{M} \left(\frac{1}{n_g} - \frac{1}{N_g}\right) \frac{(n_g - 1)}{n_g} \sigma^2_{\epsilon_g} \rho_{\epsilon_g} N_g^2 \\ & + \sum_{g=1}^{M} \left[E_{\underline{r}|\underline{r}\geq\underline{1}} \left(\frac{1}{r_g}\right) - \frac{1}{n_g} \right] \\ & \left(\sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) + \sigma^2_{\epsilon_g}(1 - \rho_{\epsilon_g}) \right) N_g^2 \end{aligned} \quad (12)$$

which reflects the contribution to the variance made by each error source. The first term is the sampling variance when each sampled unit gives the requested information without errors; the second term, conditionally on the strata sample sizes, is the increase in variance due to nonresponse; the third and fourth terms are related to the measurement variance of the observed values; and the last term is the interaction variance. In particular, the third component called the simple measurement variance arises from variability in measurements on individual elements. The fourth components called the correlated measurement variance depends on the covariances between measurements on different elements within the same stratum.

In order to simplify the constrained optimization problem of section 4, we introduce the following approximation

$$E_{\underline{r}/\underline{r}\geq\underline{1}} \left(\frac{1}{r_g}\right) \geq \frac{1}{E_{\underline{r}/\underline{r}\geq\underline{1}}(r_g)} = \frac{[1 - (1 - \alpha_g)^{n_g}]}{n_g \alpha_g} \simeq \frac{1}{n_g \alpha_g} \quad (13)$$

where the response set size r_g is a binomially distributed random variable with parameters (n_g, α_g), for $g = 1, .., M$. On the accuracy of approximation (13), note that the mean difference $E[(E_{\underline{r}|r\geq 1}(1/r_g) - 1/(n_g\alpha_g))/\alpha_g]$ is less then 0.01 for $n_g > 30$. If the terms $1/N_g$ are negligible then the approximate mean square error is given by

$$MSE(\tilde{\hat{t}}_\theta) \cong \sum_{g=1}^{M} \frac{1}{n_g\alpha_g}\left(\sigma^2_{\theta_g} + \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) + \sigma^2_{\epsilon_g}\right)N_g^2 + + \sum_{g=1}^{M} \frac{(n_g\alpha_g - 1)}{n_g\alpha_g}\sigma^2_{\epsilon_g}\rho_{\epsilon_g}N_g^2 + \left(\sum_{g=1}^{M}\bar{\delta}_g N_g\right)^2 \quad (14)$$

where the last term is the squared bias. The total error depends on both the strata sample sizes n_g and the unknown parameters

$$(\alpha_g, \sigma^2_{\delta_g}, 2Cov(\theta_g, \delta_g), \bar{\delta}_g, \sigma^2_{\epsilon_g}, \rho_{\epsilon_g})$$

coming from the nonsampling errors models. Clearly, these parameters must be estimated. For details on this topic, see Särndal et al. (1992).

2 Cost Model

In order to determine the optimal allocation of the available resources minimizing the total error, in this section we introduce a cost model describing the survey costs structure on the basis of the error sources taken into account.

It is known that an efficient sample design must provide reasonably precise estimate under the constraint of a fixed budget. Then the most efficient design is achieved by minimizing the sampling variance subject to a fixed cost B. The solution is limited to one source of error: sampling error. In presence of nonsampling errors the question is to determine the optimal allocation of the available resources B to obtain the best accuracy in survey estimate. Formally, we must solve the following constrained optimization problem

$$\min_{(\underline{c}^{(1)}, \underline{c}^{(2)}, \underline{n})} MSE(\tilde{\hat{t}}_\theta) \quad s.t. \quad B = f(\underline{c}^{(1)}, \underline{c}^{(2)}, \underline{n}) \quad (15)$$

where the cost model $B = f(\underline{c}^{(1)}, \underline{c}^{(2)}, \underline{n})$ depends on both the strata sample sizes $\underline{n} = (n_1, .., n_M)$ affecting the sampling error, and the strata per-unit costs $[\underline{c}^{(1)}, \underline{c}^{(2)}] = [(c_1^{(1)}, .., c_M^{(1)}), (c_1^{(2)}, .., c_M^{(2)})]$ affecting the strata nonresponse rates and the strata measurement errors respectively. That is, executing the survey operations related to the data collection phase more carefully leads to a decrease in nonsampling error through an increase in survey costs $(\underline{c}^{(1)}, \underline{c}^{(2)})$.

Suppose that the overall cost B of the survey is decomposed into more detailed components, associated with various aspects of its design and implementation. More specifically, we postulate that the overall cost is a linear function expressed as

$$\begin{aligned} B &= C_0 + \sum_{g=1}^{M} c_g^{(1)} n_g + \sum_{g=1}^{M} c_g^{(2)} E(r_g | r_g \geq 1) \\ &\cong C_0 + \sum_{g=1}^{M} (c_g^{(1)} + c_g^{(2)}\alpha_g) n_g \underline{c}^{(1)} > \underline{c}^{(0)}, \underline{c}^{(2)} > \underline{c}^{(0)} \quad (16) \end{aligned}$$

where $\underline{c}^{(0)}$ is a vector of length M with all components equal to $c^{(0)}$, the questionnaire sending cost by mail. In other words $c^{(0)}$ is the minimum cost that is needed to pay for interviewing the sample units. Note that C_0 is a fixed cost, to be incurred regardless of what sample size is chosen. This component includes the costs related to preparatory activities, that do not depend on the total sample size n such as coordination of survey planning, frame development, sample design and so on. In more detail, in the cost model (16)

1. $c_g^{(1)}$ represents the per-unit cost to keep nonresponse low or alternatively to increase the response probability α_g in the *g-th* stratum. Reducing nonresponse means to employ strategies that reduce the reluctance of the sample units to cooperate with an interview request. For instance, if the first contact with the potential respondent is an unsuccess it is common to instruct the interviewers to return, conducting the interview at a time more convenient for the potential respondent. Then, the larger will be the per-unit cost $c_g^{(1)}$ the larger will be the callbacks to obtain a response, or alternatively more accurate will be the data collection mode in the follow-up procedure (mail questionnaire, telephone call, face to face interview). Hence, the component $c_g^{(1)} n_g$ in the model (16) is the cost related to the effort of contacting each sampled unit within the *g-th* stratum.

2. $c_g^{(2)}$ represents the per-unit cost to improve the responses quality in the *g-th* stratum. This cost includes the use of qualified interviewers, the increase of budget reserved to supervision phase, an accurate questionnaire wording and so on. For instance, since some items in the questionnaire could be ambiguous and lead to misunderstanding on the part of the respondent, bias from question wording is minimized through careful design of the survey questions, pilot testing, analysis of pilot-test results and interviewer feedback that can reveal problem with understanding of the question. Analogously the use of qualified interviewers or the interviewers training allows to reduce the errors coming from the interviewer misunderstanding, the incorrectly recording and the interviewer influence on the responses. Finally, efforts to instruct respondents in appropriate respondent behavior (e.g., thinking carefully, seeking clarification from the interviewer), increasing the length of the questionnaire and the interviewing time, also have cost implications (Groves

1989). Hence, the component $c_g^{(2)} n_g \alpha_g$ in the model (16) represents the cost to improve the data quality provided by the expected number of respondents $n_g \alpha_g$ in the *g-th* stratum.

Note that the cost model (16) is continuous. In practice a more realistic cost function is frequently a stepwise function rather than a linear function of the principal cost factors. For instance, if 10 interviews can be conducted in a single day, then the addition of one interview requires an extra day of work and thus a substantial cost increase, whereas the addition of two interviews may add little cost. Clearly, the discontinuities in the cost models imply that partial derivatives do not exist, hence no single optimum total survey design can be found through standard calculus methods. It should be pointed out that the cost model is an approximation of the reality affecting the total survey design features. Hence, some attention needs to be paid to the specification of the cost model to determine whether its form is sufficiently appropriate to the survey. Groves has a relatively large discussion on cost models, including various complex and realistic forms, e.g., non-linear, discontinuous, step-function cost expression (Groves 1989). However, the scarcity of detailed information on costs associated with various aspects of survey implementation often makes to set up a cost model a hard task.

3 Reparametrization Model

This section deals with the reparametrization model allowing us to associate the nonsampling error with the cost components $(\underline{c}^{(1)}, \underline{c}^{(2)})$, affecting nonresponse and measurement error respectively. More specifically, in the constrained optimization problem (15) while sampling error is related to the cost model by the strata sample sizes, nonsampling error is not, since $(\underline{c}^{(1)}, \underline{c}^{(2)})$ do not appear in the $MSE(\hat{\tilde{t}}_\theta)$. As a consequence, we formalize the cost-error tradeoff through model assumptions expressing the strata MSE parameters

$$(\alpha_g, \sigma^2_{\delta_g}, 2Cov(\theta_g, \delta_g), \overline{\delta}_g, \sigma^2_{\epsilon_g}, \rho_{\epsilon_g}) \tag{17}$$

coming from the nonsampling errors models, by strata per-unit costs $(c_g^{(1)}, c_g^{(2)})$. In order to accomplish this, one must (i) give the mathematical formulation of the model, and (ii) estimate the parameters appearing in it. Looking for a compromise between intuitive validity and simplicity of the reparametrization model, we have assumed functional hypotheses regarding a hyperbolic model. We suppose that the stratum response probability α_g depends on $c_g^{(1)}$ according to the function

$$\alpha_g = \alpha(c_g^{(1)}) = \frac{\alpha^* \gamma_g c_g^{(1)}}{1 + \gamma_g c_g^{(1)}} \tag{18}$$

where the parameter $\gamma_g > 0$ measures the effect of additional financial resources on the response probability. The larger the cost $c_g^{(1)}$ the larger is the effort of contacting each sampled unit within the *g-th* stratum, then the larger will be the stratum response rate. Note that as $c_g^{(1)}$ increases, the function (18) increases to the asymptote α^*. We assume $\alpha^* < 1$ since in most surveys there is a fraction of hard core nonresponse, composed of elements that do not under any circumstances respond. Besides, since the per-unit cost $c_g^{(2)}$ affects the measurement error, we assume

$$\begin{cases} \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) + \sigma^2_{\epsilon_g} = \frac{1}{a_g c_g^{(2)}} = \frac{\beta_g}{c_g^{(2)}} \\ \sigma^2_{\epsilon_g} \rho_{\epsilon_g} = \frac{1}{f_g c_g^{(2)}} = \frac{\varphi_g}{c_g^{(2)}} \\ \overline{\delta}_g = \frac{1}{k_g c_g^{(2)}} = \frac{\tau_g}{c_g^{(2)}} \end{cases} \tag{19}$$

where $\beta_g = 1/a_g, \varphi_g = 1/f_g, \tau_g = 1/k_g > 0$. The functions (19) decrease as $c_g^{(2)}$ increases. For instance, the larger is the per-unit cost $c_g^{(2)}$ more accurate will be the measurement process then the lower will be the measurement error magnitude. For instance, face to face interview is the most expensive data collection mode since fieldwork and its organization requires more resources and generates more costs, while one of the advantages of a postal survey is low cost. At the same time, the presence of an interviewer makes the interview situation more controllable and makes it possible to clarify both questions and answers reducing the measurement error. The parameters (a_g, f_g, k_g) measure the effect of additional financial resources on the measurement error components (simple response variance, correlated response variance and response bias).

As we show in section 4.1, in order to obtain estimates of unknown parameters appearing in (18)-(19) we could refer to previous surveys or carry out a pilot survey. The effects evaluation on both response rate and measurement error components of additional financial resources could require an experiment too. For instance, to determine how much additional callbacks will increase the response probability, or how much additional interviewers training will reduce measurement bias and variance could require experimenting with different callbacks number or different levels of interviewers training.

It is important to stress that the functions appearing in (18)-(19) do not come from quantitative studies or actual survey data, but represent just an attempt to formalize the cost-error tradeoff for arriving at a total survey design. As a consequence, the reparametrization model can be seen as a key tool in the search for an overall optimization of a survey. In practice, more realistic functions are probably not smooth and may not even be continuous.

4 Optimal allocation of the available resources

This section essentially deals with the optimal allocation problem of the available resources minimizing the total error. The goal is to obtain the best possible accuracy in survey estimate through an overall economic balance between sampling and nonsampling error. The problem can be stated in this way: given the cost structure represented in the cost model, how should the budget be allocated to minimize the total survey error. Formally, we must solve the following constrained optimization problem

$$\min_{(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n})} MSE(\tilde{\hat{t}}_\theta) \quad s.t. \quad D = \sum_{g=1}^{M} (c_g^{(1)} + c_g^{(2)}\alpha_g) n_g \tag{20}$$

where $D = B - C_0$. As a guideline to the analysis, it could be useful to consider two different approaches: the former is conditional on the strata per-unit costs $(\underline{c}^{(1)}, \underline{c}^{(2)})$, the latter is unconditional. As a matter of fact, assuming as given the strata per-unit costs is equivalent to fix the survey conditions under which the data will be collected. In addition to the sampling method, any data collection effort involves making decisions on a number of survey operations: the data collection mode, the questionnaire wording, the interviewers training, the use of supervisors, the maximum number of callbacks, the data collection mode in follow-up procedure and so on. All these decision have cost implications. Hence, given the strata per-unit costs $(\underline{c}^{(1)}, \underline{c}^{(2)})$ the strata response probabilities and the strata measurement errors magnitude are given too, since the features of the total survey design have been defined. As a consequence, prior of the sample selection and given an estimate of parameters (17), the constrained optimization problem (20) can be formulated as the determination of the n_g minimizing the MSE under the cost model (16). Formally, we have

$$\min_{\underline{n}} MSE(\tilde{\hat{t}}_\theta) \quad s.t. \quad D = \sum_{g=1}^{M} (c_g^{(1)} + c_g^{(2)}\alpha_g) n_g \tag{21}$$

A method for carrying out such optimization is the Lagrange's Multiplier Method. It consists of introducing a new function which incorporates the $MSE(\tilde{\hat{t}}_\theta)$ together with the constraint. Let $\widehat{A}_g$ be the estimate of the stratum parameter

$$A_g = \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) + \sigma^2_{\epsilon_g}(1 - \rho_{\epsilon_g}) \tag{22}$$

and denote by $\widehat{\alpha}_g$ the stratum response probability estimate. Let $L(\underline{n}, \lambda)$ be the Lagrangian function, where λ is the Lagrange multiplier, the M+1 simultaneous conditions

$$\frac{\partial L(\underline{n}, \lambda)}{\partial n_g} = 0 \quad \text{for} \quad g=1,..,M; \quad \frac{\partial L(\underline{n}, \lambda)}{\partial \lambda} = 0 \tag{23}$$

give the solution

$$n_g^* = D \frac{\sqrt{(\sigma^2_{\theta_g} + \widehat{A}_g) N_g^2} / \sqrt{\widehat{\alpha}_g (c_g^{(1)} + c_g^{(2)}\widehat{\alpha}_g)}}{\sum_{g=1}^{M} \sqrt{(\sigma^2_{\theta_g} + \widehat{A}_g)(c_g^{(1)} + c_g^{(2)}\widehat{\alpha}_g) N_g^2} / \sqrt{\widehat{\alpha}_g}} \tag{24}$$

for $g=1,..,M$

representing the strata sample sizes minimizing the total error subject to a fixed cost, given the survey conditions. Note that, in absence of nonresponse and measurement errors we obtain the standard result of optimum sample allocation (Cochran 1977).

On the other hand, to adopt the unconditional approach means that in the planning phase of a survey the researcher can influence the magnitude of nonsampling errors through the choice of strata per-unit costs $(c_g^{(1)}, c_g^{(2)})$, defining the survey conditions under which the data will be collected. Then, the question is how to allocate a budget available between reduction of sampling variance, maximization of response rate and minimization of measurement errors. It then becomes clear how the cost-error tradeoff knowledge, formalized by (18)-(19), is fundamental for an optimal allocation of the resources when an unconditional approach is employed. Formally, substituting the reparametrization model in the objective function of the constrained optimization problem (20) we obtain

$$\begin{aligned} \min_{(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n})} & \sum_{g=1}^{M} \frac{1+\gamma_g c_g^{(1)}}{n_g \alpha^* \gamma_g c_g^{(1)}} \left(\sigma^2_{\theta_g} + \frac{\beta_g}{c_g^{(2)}} \right) N_g^2 \\ & + \sum_{g=1}^{M} \left(1 - \frac{1+\gamma_g c_g^{(1)}}{n_g \alpha^* \gamma_g c_g^{(1)}} \right) \frac{\varphi_g}{c_g^{(2)}} N_g^2 \\ & + \left(\sum_{g=1}^{M} \frac{\tau_g}{c_g^{(2)}} N_g \right)^2 \\ & \text{s.t.} \\ D = & \sum_{g=1}^{M} \left[c_g^{(1)} + c_g^{(2)} \left(\frac{\alpha^* \gamma_g c_g^{(1)}}{1+\gamma_g c_g^{(1)}} \right) \right] n_g \end{aligned} \tag{25}$$

Denoted by $L(\underline{c}^{(1)}, \underline{c}^{(2)}, \underline{n}, \lambda)$ the Lagrangian function, the solution is determined by solving the following system of $3M + 1$ equations

$$\begin{cases} \frac{\partial L(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n},\lambda)}{\partial n_g} = 0 \quad \frac{\partial L(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n},\lambda)}{\partial c_g^{(1)}} = 0 \quad \frac{\partial L(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n},\lambda)}{\partial c_g^{(2)}} = 0 \\ \frac{\partial L(\underline{c}^{(1)},\underline{c}^{(2)},\underline{n},\lambda)}{\partial \lambda} = 0 \qquad \text{for} \quad g=(1,..,M); \end{cases} \tag{26}$$

where λ is the Lagrange multiplier. In next section we discuss a possible application of the model developed in section 1, in order to show how the unconditional approach could be implemented.

4.1 *A possible application to business surveys*

Nevertheless a total error model must be highly specified in an actual survey situation, in this section we suppose to apply the model of section 1 to business surveys. Consider a survey that collects data for only one item and estimates a single population-level total (e.g., total employment, total sales).

Business surveys are usually designed as one stage stratified simple random samples selected without replacement. A more common method for stratifying a skewed population is to create a certainty stratum of large businesses in which all units are sampled. This work can be easily extended including a take all or completely enumerated stratum. In business surveys the stratification usually has three dimensions: size, type of activity and region. Nevertheless the efforts to obtain full response to a survey a certain amount of nonresponse is inevitable. As indicated by Cox et al. (1995:490), in business surveys the list frame design strata tend to form good weighting classes for nonresponse adjustment. As a consequence, we assume that the strata used for the sample selection and the response homogeneity groups coincide: businesses of similar size, in the same region and in the same industrial sector have the same chance of response. Furthermore, we postulate that the measurement errors generating mechanism is described by (1)-(2). As a consequence, the stratification variables contribute to explain both the stratum response propensity and the measurement errors generating mechanism.

In the unconditional approach the available resources must be allocated to provide the highest overall accuracy, shifting them to the area where they are more effective. Then, the survey designer needs to understand potential sources of error and attempt to model them as function of the resources available.

With regard to nonresponse error, it is known that certain techniques (follow-up contacts, financial incentives to induce compliance with the survey request, prior notice) have emerged as important in achieving high response in business surveys. Then, given a set of techniques to keep nonresponse low and assumed as known the costs and the response rates associated to them, we can estimate through the least squares method the parameter γ_g appearing in the reparametrization model (18). The response rates can be obtained from previous surveys or carry out empirical studies.

The same consideration holds for the measurement error components. One important source of measurement error in business surveys is the survey instrument, that is the questionnaire and the instructions to the respondent for supplying the requested information. The economic concepts traditionally used in business surveys are not easily understood by respondents.

In this case, standard techniques that keep the measurement error low regard the questionnaire wording, the use of qualified interviewers or the interviewers training and the use of supervisors. For instance, suppose we want to evaluate through an experiment the effect on the measurement error components (measurement bias, simple response variance, correlated response variance) of the number of days devoted to interviewers training (e.g., one day, five days, ten days) in conjunction with the use of supervisors. Through an empirical study we can estimate the magnitude of measurement error components associated to each option: the magnitude of error reduction. These data together with the costs associated to each survey method, allows us to estimate the parameters in (19).

After estimating the parameters appearing in the reparametrization model we can solve the constrained optimization problem (25), obtaining the combination of strata sample sizes, nonresponse and measurement procedures which minimize the total error within the resources available for the survey.

Note that we assumed a continuous reparametrization model. For instance, the response probability α_g is a continuous increasing function of per-unit cost $c_g^{(1)}$. As a consequence, each increase in $c_g^{(1)}$ implies an increase in the *g-th* stratum response probability. A considerable problem concerns with the stepwise nature of reparametrization model. As stressed in section 3, more realistic functions are probably not smooth and may not even be continuous. For instance, to increase the response rate could be necessary numerous callbacks with a substantial investment on $c_g^{(1)}$. Note that the discontinuities offer the researcher large benefits if they can be identified. Besides, as we will stress in the next section, a set of methods to keep nonresponse low could be eliminated as feasible alternatives in an actual survey situation because of time constraints.

As a matter of fact, evaluation studies can be very useful in defining the reparametrization model and getting estimates of parameters appearing in it. This implies that extra data must be collected and experimental methods must be used to estimate the cost-error tradeoff. Resources need to be provided to support these studies.

5 Conclusions

Estimating total survey error assumes that a mathematical model exists that includes all the key error sources of the survey. The first step in the development of such a model consists in deciding what nonsampling errors should be considered. Since the important sources of error vary from survey to survey, this choice is strictly depending on the survey characteristics. For instance, coverage and interviewer errors might be the most important error sources in some surveys, while in other surveys the most considerable sources of error might be nonresponse or recall errors. Besides, most surveys produce many estimates that may be affected differently by the error sources. The total survey error approach should be flexible enough both to allow different errors to be incorporated and to take into account their covariance structure (Federal Committee on Statistical Methodology 2001)

Since inference in presence of nonsampling errors must rely in part on modelling, we must make model assumptions about their generating mechanism in order to evaluate their magnitude. Note that the selection of a model needs a careful analysis of the social environment where the survey is to be conducted, study variable nature as well as all possible interactions between errors coming from different phases of survey. Since the formulation of a given model is the same as introducing some not fully testable hypotheses, any inference will depend on the validity of such assumptions. Model misspecifications will lead to invalid conclusions. As we stressed in the introduction, the implementation of a total error model represents a preliminary step for arriving at a total survey design.

In practice, an integrated unified total survey error model is as improbable as an unified theory for predicting human behavior across a myriad of situations. The "best" model for a given survey will not be the "best" for another. In other words, the nonsampling errors models are not transferable from one survey to another, hence a general model valid for all surveys is a non-sense. For instance, the nonsampling errors models for an income survey will differ from the models for other survey topics. More generally, since the rich tend to underestimate their income while the poor tend to overestimate it the measurement model will have probably a multiplicative form.

As we stressed previously, the main problem is how we can formalize the cost-error tradeoff for arriving at a total survey design. In this paper, we proposed the reparametrization model as a key tool in the search for an overall optimization of a given survey. Note that in the reparametrization model, the MSE parameters $(\alpha_g, \sigma^2_{\delta_g}, 2Cov(\theta_g, \delta_g), \overline{\delta}_g, \sigma^2_{\epsilon_g}, \rho_{\epsilon_g})$ have been expressed just as functions of the corresponding strata per-unit costs. For instance, the response probability α_g depends on $c_g^{(1)}$ according to (18). In practice, it is known that many actions reducing measurement errors (for example the use of qualified interviewers, an accurate questionnaire construction) also reduce nonresponse and *vice-versa*. Hence, further developments could include more complex reparametrization models. For instance, the stratum response probability could assume the following form

$$\alpha_g = \alpha(c_g^{(1)}, c_g^{(2)}) \tag{27}$$

depending on both costs $(c_g^{(1)}, c_g^{(2)})$.

In conclusion, the paper is concerned with obtaining the best possible accuracy in survey estimate through an overall economic balance between sampling and nonsampling error. As Groves emphasizes it is important to think to the survey from a cost stand point because of its importance as limiting factor (Groves 1989) The best practices in different fields like sampling, avoiding nonresponse, decreasing measurement errors, are often not compatible in the same survey because they are generally also the most expensive (Desrosieres et al. 2001). Hence, we need to identify the best practices in each field conditionally on the available money.

It is important to stress that the optimal allocation of the available budget loses its optimality in an empirical context both because the MSE is the error resulting from the error sources that the model takes into account and for misspecifications regarding the nonsampling errors, cost and reparametrization models. Besides, the decision problem is so complex that an optimum, in the sense of a mathematical solution to a closed-form problem, is inconceivable. There are too many interrelated decisions and too many variables to take into account (Särndal et al. 1992; Särndal and Platek 2001).

The decision, therefore, would likely be made on the basis of quantitative studies or assumptions, budgetary and priority considerations, time constraints but not on the basis of a mathematical model alone (Fellegi and Sunter 1974). In actual situations, the choice criteria may conflict and require a compromise. For instance, the time required to conduct the follow-up of nonrespondents may eliminate it as a feasible alternative to reduce nonresponse. Hence, the time constraint introduces in the balance between error components a source of limitation to the alternatives that need to be considered. Despite of its limitations we believe that this optimization approach stimulating the research on both survey costs and cost-error tradeoff, could improve the information about the survey estimates accuracy.

Lack of data on costs of survey methods and their associated errors seriously impedes the use of the total survey design approach at the planning stages. As a consequence, there is a great need for accumulating a systematic body of knowledge regarding the nonsampling errors, the survey costs and their relationship. In order to accomplish this, experimental procedures must be introduced in the survey process to evaluate both the effect of a given error source on the total error, and the effect that alternative survey methods have on reducing its magnitude.

Since is expensive to develop a total error model and carry out an overall optimization of the survey, such expense is justified for a large-scale survey that will continue for a number of years: the knowledge accumulates during the years can be used in the planning phase of surveys that follow.

Much work still remains to be done in this context both in theoretical and applied field. As we showed, the attempt to formalize the tradeoff between cost and error for arriving at a total survey design generates a set of parallel research fields which development is needed.

References

Andersen, R., Kasper, J., & Frankel, M. (1979). *Total survey error : Applications to improve health surveys*. San Francisco: Jossey-Bass.

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

Cox, B. G., Binder, D. A., Chinnappa, B. N., Christianson, A., Colledge, M. J., & Kott, P. S. (1995). *Business survey methods*. New York: Wiley.

Desrosières, A., Deville, J., & Sautory, O. (2001). Discussion (C. E. Särndal and R. Platek, 2001). *Journal of Official Statistics*, *17*, 33-37.

Federal Committee on Statistical Methodology. (2001). *Measuring and reporting sources of error in surveys.* Statistical Policy Working Paper, Vol. 31, online available at www.fcsm.gov.

Fellegi, I. P., & Sunter, A. B. (1974). Balance between different sources of survey errors - some canadian experiences. *Sankhyā: The Indian Journal of Statistics*, *36*(C), 119-142.

Forsman, G. (1989). Early survey models and their use in survey quality work. *Journal of Official Statistics*, *5*(1), 41-55.

Forsman, G. (1993). Recent advances in survey error modelling. *Jahrbücher für Nationalökonomie und Statistik*, *211*, 331-350.

Groves, R. M. (1989). *Survey errors and survey costs*. New York: Wiley.

Groves, R. M. (1999). Survey error models and cognitive theories of response behavior. In M. Sirken, D. Herrmann, S. Schechter, N. Schwarz, J. Tanur, & R. Tourangeau (Eds.), *Cognition and survey research* (p. 235-250.). New York: Wiley.

Hansen, M. H., Hurwitz, W. N., & Bershad, M. A. (1961). Measurement errors in censuses and surveys. *Bulletin of the International Statistical Institute*, *38*(2), 359-374.

Hansen, M. H., Hurwitz, W. N., Marks, E. S., & Mauldin, W. P. (1951). Response errors in surveys. *Journal of the American Statistical Association*, *46*, 147-190.

Hansen, M. H., Hurwitz, W. N., & Pritzker, L. (1965). *The estimation and interpretation of gross errors and the simple response variance.* Oxford: Pergamon.

Kish, L. (1965). *Survey sampling*. New York: Wiley.

Lessler, J. T., & Kalsbeek, W. D. (1992). *Nonsampling error in surveys*. New York: Wiley.

Särndal, C. E., & Platek, R. (2001). Can a statistician deliver? *Journal of Official Statistics*, *17*(1), 1-20.

Särndal, C. E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer.

Tranquilli, G. B. (1995). *Teoria attuale: incompiuta o disattesa?* Atti del Convegno SIS (in Italian).

Weisberg, H. F. (2005). *The total survey error approach: A guide to the new science of survey research*. Chicago: University of Chicago Press.

Appendix A

The expected value for $\tilde{\hat{t}}_\theta$ is given by

$$
\begin{aligned}
E(\tilde{\hat{t}}_\theta) &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta) \\
&= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon \\
&\quad \left\{ \sum_{g=1}^{M} \left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \nu_{gh} \left(\frac{1}{r_{gh}} \sum_{i_{gh}\in\underline{s}_{r_{gh}}} \tilde{y}_{i_{gh}} \right) \right] N_g | \underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n \right\} \\
&= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} \\
&\quad \left\{ \sum_{g=1}^{M} \left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \nu_{gh} \left(\frac{1}{r_{gh}} \sum_{i_{gh}\in\underline{s}_{r_{gh}}} \mu_{i_{gh}} \right) \right] N_g | \underline{r} \geq \underline{1}, \underline{s}_n \right\} \\
&= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} \\
&\quad \left\{ \sum_{g=1}^{M} \left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \nu_{gh} \left(\frac{1}{n_{gh}} \sum_{i_{gh}\in\underline{s}_{n_{gh}}} \mu_{i_{gh}} \right) \right] N_g | \underline{s}_n \right\} \\
&= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} \left[\sum_{g=1}^{M} \overline{\mu}_{n_g} N_g | \underline{s}_n \right] \\
&= E_{\underline{s}_n} \left[\sum_{g=1}^{M} \overline{\mu}_{n_g} N_g \right] \\
&= \sum_{g=1}^{M} \overline{\mu}_g N_g \\
&= \sum_{g=1}^{M} \overline{\theta}_g N_g + \sum_{g=1}^{M} \overline{\delta}_g N_g \qquad (28)
\end{aligned}
$$

where $E_\epsilon(.|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n)$, $E_{\underline{s}_r}(.|\underline{r} \geq \underline{1}, \underline{s}_n)$, $E_{\underline{r}|\underline{r}\geq\underline{1}}(.|\underline{s}_n)$ denote conditional expectations with respect to the measurement model, the response model and the response set size $\underline{r}$ (for $\underline{r} \geq \underline{1}$) respectively, while $E_{\underline{s}_n}(.)$ denotes expectation over all possible samples. Besides, $\overline{\mu}_{n_g}$ and $\overline{\mu}_g$ represent the sample mean and the population mean of the expected measurement values in the *g-th* stratum.

Appendix B

The variance of $\tilde{\hat{t}}_\theta$ is shown to be composed of four components. This can be seen by writing the variance expression in the following form

$$\begin{aligned} Var(\tilde{\hat{t}}_\theta) &= Var_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n) \\ &+ E_{\underline{s}_n} Var_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n) \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} Var_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n) \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} Var_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n) \\ &= A + B + C + D \end{aligned} \tag{29}$$

Consider each of these separately, we have

$$\begin{aligned} A &= Var_{\underline{s}_n}\left[\sum_{g=1}^{M} \overline{\mu}_{n_g} N_g\right] \\ &= \sum_{g=1}^{M} \frac{N_g}{N_g - 1}\left(\frac{1}{n_g} - \frac{1}{N_g}\right)\sigma^2_{\mu_g} N_g^2 \\ &\cong \sum_{g=1}^{M}\left(\frac{1}{n_g} - \frac{1}{N_g}\right)\sigma^2_{\mu_g} N_g^2 \end{aligned} \tag{30}$$

where

$$\begin{aligned} \sigma^2_{\mu_g} &= \frac{1}{N_g}\sum_{i\in U_g}(\mu_{i_g} - \overline{\mu}_g)^2 \\ &= \frac{1}{N_g}\sum_{i\in U_g}[(\theta_{i_g} - \overline{\theta}_g) + (\delta_{i_g} - \overline{\delta}_g)]^2 \\ &= \sigma^2_{\theta_g} + \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g) \end{aligned} \tag{31}$$

and $Cov(\theta_g, \delta_g)$ represents the covariance between systematic errors and true values in the g-th stratum.

Considering the term $B = E_{\underline{s}_n} Var_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n)$, we have

$$\begin{aligned} B &= E_{\underline{s}_n} Var_{\underline{r}|\underline{r}\geq\underline{1}}\left[\sum_{g=1}^{M} \overline{\mu}_{n_g} N_g|\underline{s}_n\right] \\ &= 0 \end{aligned} \tag{32}$$

Considering the term $C = E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} Var_{\underline{s}_r} E_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n)$, we have

$$\begin{aligned} C &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} Var_{\underline{s}_r}\left\{\sum_{g=1}^{M}\left(\sum_{h=1}^{H_{\underline{s}_{ng}}} \nu_{gh}\overline{\mu}_{r_{gh}}\right) N_g|\underline{r} \geq \underline{1}, \underline{s}_n\right\} \\ &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{n_{gh}}{n_{gh} - 1}\left(\frac{1}{r_{gh}} - \frac{1}{n_{gh}}\right) s^2_{\mu_{gh}}\nu^2_{gh}\right] N_g^2|\underline{s}_n\right\} \end{aligned} \tag{33}$$

where $s^2_{\mu_{gh}}$ is the sample variance of the expected measurement values in the gh-th response group $\underline{s}_{n_{gh}}$, while $\overline{\mu}_{r_{gh}}$ is the sample mean of the r_{gh} respondents in the same group. Finally, to complete the variance we need $D = E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r} Var_\epsilon(\tilde{\hat{t}}_\theta|\underline{s}_r, \underline{r} \geq \underline{1}, \underline{s}_n)$. It follows that

$$\begin{aligned} D &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left(\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{\sigma^2_{\epsilon_g}}{r_{gh}}\nu^2_{gh}\right) N_g^2|\underline{s}_n\right\} \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left(\sum_{h=1}^{H_{\underline{s}_{ng}}}(r_{gh} - 1)\frac{\rho_{\epsilon_g}\sigma^2_{\epsilon_g}}{r_{gh}}\nu^2_{gh}\right) N_g^2|\underline{s}_n\right\} \end{aligned} \tag{34}$$

To summarize, we have

$$\begin{aligned} Var(\tilde{\overline{\theta}}) \cong& \sum_{g=1}^{M}\left(\frac{1}{n_g} - \frac{1}{N_g}\right)\left[\sigma^2_{\theta_g} + \sigma^2_{\delta_g} + 2Cov(\theta_g, \delta_g)\right] N_g^2 \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left[\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{n_{gh}}{n_{gh} - 1}\left(\frac{1}{r_{gh}} - \frac{1}{n_{gh}}\right) s^2_{\mu_{gh}}\nu^2_{gh}\right] N_g^2|\underline{r} \geq \underline{1}, \underline{s}_n\right\} \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left(\sum_{h=1}^{H_{\underline{s}_{ng}}} \frac{\sigma^2_{\epsilon_g}}{r_{gh}}\nu^2_{gh}\right) N_g^2|\underline{r} \geq \underline{1}, \underline{s}_n\right\} \\ &+ E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left\{\sum_{g=1}^{M}\left(\sum_{h=1}^{H_{\underline{s}_{ng}}}(r_{gh} - 1)\frac{\rho_{\epsilon_g}\sigma^2_{\epsilon_g}}{r_{gh}}\nu^2_{gh}\right) N_g^2|\underline{r} \geq \underline{1}, \underline{s}_n\right\} \end{aligned} \tag{35}$$

Appendix C

If we assume that the strata used for the sample selection and the response homogeneity groups coincide, the terms A and B in the variance (29) remain unchanged. The term (33) becomes

$$\begin{aligned} C &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} Var_{\underline{s}_r}\left[\sum_{g=1}^{M}\left(\frac{1}{r_g}\sum_{i_g\in\underline{s}_{r_g}}\mu_{i_g}\right)N_g|\underline{r}\geq\underline{1},\underline{s}_n\right] \\ &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}}\left[\sum_{g=1}^{M}\frac{n_g}{n_g-1}\left(\frac{1}{r_g}-\frac{1}{n_g}\right)s^2_{\mu_g}N_g^2|\underline{r}\geq\underline{1},\underline{s}_n\right] \\ &= E_{\underline{s}_n}\left[\sum_{g=1}^{M}\frac{n_g}{n_g-1}\left(E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{1}{r_g}\right)-\frac{1}{n_g}\right)s^2_{\mu_g}N_g^2|\underline{s}_n\right] \\ &= \sum_{g=1}^{M}\left[E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{1}{r_g}\right)-\frac{1}{n_g}\right]\sigma^2_{\mu_g}N_g^2 \end{aligned} \tag{36}$$

The term (34) becomes

$$\begin{aligned} D &= E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r}\left[\sum_{g=1}^{M}\left(\frac{\sigma^2_{\epsilon_g}}{r_g}\right)N_g^2|\underline{r}\geq\underline{1},\underline{s}_n\right] \\ &\quad + E_{\underline{s}_n} E_{\underline{r}|\underline{r}\geq\underline{1}} E_{\underline{s}_r}\left[\sum_{g=1}^{M}\left(\frac{r_g-1}{r_g}\right)\rho_{\epsilon_g}\sigma^2_{\epsilon_g}N_g^2|\underline{r}\geq\underline{1},\underline{s}_n\right] \\ &= \sum_{g=1}^{M}E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{1}{r_g}\right)\sigma^2_{\epsilon_g}N_g^2 + \\ &\quad + \sum_{g=1}^{M}E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{r_g-1}{r_g}\right)\rho_{\epsilon_g}\sigma^2_{\epsilon_g}N_g^2 \end{aligned} \tag{37}$$

To summarize, we have

$$\begin{aligned} Var(\tilde{\hat{t}}_\theta) \cong& \sum_{g=1}^{M}\left(E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{1}{r_g}\right)-\frac{1}{N_g}\right)\left[\sigma^2_{\theta_g}+\sigma^2_{\delta_g}+2Cov(\theta_g,\delta_g)\right]N_g^2 \\ &+ \sum_{g=1}^{M}E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{1}{r_g}\right)\sigma^2_{\epsilon_g}N_g^2 + \sum_{g=1}^{M}E_{\underline{r}|\underline{r}\geq\underline{1}}\left(\frac{r_g-1}{r_g}\right)\rho_{\epsilon_g}\sigma^2_{\epsilon_g}N_g^2 \end{aligned} \tag{38}$$