STATISTICA, anno LXVII, n. 3, 2007

COMPLEX SAMPLING DESIGNS FOR THE CUSTOMER SATISFACTION INDEX ESTIMATION

T. Di Battista, P. Valentini

1. INTRODUCTION

Customer satisfaction (CS) is a central concept in marketing and is adopted as an important outcome measure of service quality by service industries (Oliver, 1997). Customers are getting more demanding with the services they receive and the products they buy. This makes firms quickly adapt themselves to develop a customer-oriented management and deliver higher quality services.

For many firms, successfully managing customer dissatisfaction is crucial to stability and profitable growth. It requires a strategy that identifies the connection between the characteristics of customers and customer dissatisfaction responses.

This paper deals with the problem of identifying dissatisfied customers for the purpose of delineating quality improvement strategies. The hypothesis of our approach is that the phenomenon of dissatisfaction is relatively rare, that is not exceedingly diffused among the different customers groups, and clusterized, in the sense that dissatisfied customers share some specific characteristics.

Of critical importance to the general validity and reliability of customer satisfaction indices (CSI's) is the use of the best suited sampling design. In fact, frequently CS assessments are conducted with little regard to statistical problems. In particular, a non-probability sampling is usually used and the derivation of a suitable statistics relative to the phenomenon of study cannot be used to infer about the characteristics of the population from which the sample come. When probability sampling are used the estimates of the parameters of interest have estimation errors because only a subset of the population is observed, but inference is valid because samples are selected according to a sampling design that assigns a known probability to them.

Sometimes designed CS surveys detect relatively few dissatisfied customers that might share the same characteristics since they can be considered as a few rare clusters in the population. In this framework, estimates of population characteristics may have high uncertainty. For such populations, adaptive designs can produce gains in efficiency, relative to conventional designs, for estimating the population parameters. In addition, adaptive sampling designs can substantially increase the yield of interesting units in the sample.

In the case of CS, with adaptive sampling designs we may increase the number of dissatisfied customers for analyzing the profile and characteristics of these customers. Considering such sampling design, the Hansen-Hurwitz procedure gives unbiased estimators of population moments, but sometimes nonlinear estimators may be used to estimate customer satisfaction indices which are generally biased and the variance estimator may not be obtained in a closed-form solution. Accordingly, delta method, jackknife and bootstrap procedures may be introduced in order to reduce bias and estimating variance.

In order to evaluate the efficiency of our proposed method, the adaptive is compared with the conventional sampling design (Thompson and Seber, 1996; Di Battista and Di Spalatro, 1998, 1999; Di Battista, 2003). As the reader will see in the simulation study, for some dissatisfied populations, particularly those that are rare and clustered, adaptive sampling strategies produce remarkable increase in efficiency compared to conventional sampling designs of equivalent sample size. In particular, in Section 2, we introduce the basic ideas, formulas and implementation of different approaches to variance estimation, while Section 3 and 4 deal with the illustration of the adaptive sampling designs; in Section 5 we show procedures to estimating variance in the adaptive sampling designs. In Section 6 a simulation study is performed in order to estimate the variance of the estimator obtained from the sample design proposed. Finally, conclusions are given in Section 7.

2. GENERAL METHODS OF VARIANCE ESTIMATION FROM COMPLEX SURVEYS

In the usual setup for finite-population sampling the population U consists of N_T distinct units identified through the label $j = 1, 2,..., N_T$. Associated with the jth unit is a variable of interest Y_j and auxiliary variable X_j, each of which can be vector valued. The parameter of interest is a function of Y_j, $j = 1, 2,..., N_T$. A sample s is a subset of units from U selected according to a sampling plan that assigns a known probability $p(s)$. A statistical analysis involves: i) the choice of sampling design; ii) the choice of the estimate of parameter of interest and iii) the construction of variance estimates and confidence sets. Here we briefly describe sampling strategies which are widely used in practice. The simplest sampling design is the simple random sampling in which units are drawn with or without replacement. Usually, this sampling design is rarely used for practical and theoretical considerations. A stratified sampling consists of partitioning the units in the population into mutually exclusive and collectively exhaustive subgroups, called strata. Then a sample is drawn from each stratum and independently across the strata. The construction of strata should be done in order to ensure that the units are homogenous internally the strata. A primary purpose of stratification is to improve the precision of the survey estimates. For example, in a student satisfaction survey the stratification can be used to assure "representativeness" of student demographics (age, gender), enrolment status (full-time versus part-time, day versus evening), class size and academic department.

Another type of probability sampling is cluster sampling where sampling units consist of a group (cluster) of smaller units. If subunits within a cluster are selected then this technique is called two stage sampling. One can continue this process to have a multistage sampling. Multistage sampling is traditionally used in large-scale surveys and because of economic considerations. For example, in CS across organizational units (i.e. banks), we can first select subunits (bank branches) and then customers from each cluster are sampled.

Generally, we may consider complex sampling designs as a stratified multistage sampling designs since they include many commonly sampling designs. According to this sampling design, the population of interest has been subdivided into H strata with N_h clusters, $h = 1,2,...,H$. Within each stratum h a sample of n_h primary sampling units (PSU's) are selected, independently across the strata. The selection of PSU's can be done using several methods, like unequal probability sampling with replacement or equal probability sampling. For each (h,i)th first stage, n_{hi} ultimates units are sampled, $i = 1, 2, ... , n_h$, $h = 1, 2, ... , H$. The total number of final units is $n_T = \sum_{h=1}^{H} \sum_{i=1}^{n_h} n_{hi}$.

Once selected the sample from the general sampling design described above, we may be interested in the estimation of population parameter θ then a survey estimator of θ is $\hat{\theta} = g(\hat{Z})$ with a nonlinear known functional g, where

$$\hat{Z} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} z_{hij} \tag{1}$$

and z_{hij} is a vector of observed data relative to (hij)th final sampling unit and w_{hij} is the corresponding survey weight. After the construction of the survey estimator, a crucial part is the derivation of estimator of the variance of the estimator, which can be used in: i) measuring precision and quality of the estimation; ii) deciding the degree of detail with which the survey data may be meaningfully analyzed; iii) determining allocation and stratification under a specific design; iv) constructing confidence sets for unknown parameters.

Several methods are available for computing the sample estimates of the variances of nonlinear statistics and they can be classified in: i) approximation methods and ii) resampling methods. One of the most useful approximation technique is the delta method (or linearization) that is based on the Taylor series approximation. A Taylor series linearization of a statistic is formed and then substituted into the formula for calculating the variance of a linear estimate appropriate for the sample design. The delta method produces the following variance estimator (Rao, 1988):

$$v_l = \sum_{h=1}^{H} \frac{1 - \lambda_h f_h}{n_h} \nabla g(\hat{Z})' s_h^2 \nabla g(\hat{Z}) \tag{2}$$

where $f_h = n_h / N_h$, $s_h^2 = 1/(n_h - 1)\sum_{i=1}^{n_h}(z_{hi} - \bar{z}_h)(z_{hi} - \bar{z}_h)'$, $z_{hi} = \sum_{j=1}^{n_{hi}} n_h w_{hij} z_{hij}$, $\bar{z}_h = 1/n_h \sum_{j=1}^{n_h} z_{hi}$ and λ_h is equal to 1 if the first stage is without replacement and 0 if the first stage is with replacement.

This is the method which produces the usual large-sample formula for the variance of the ratio estimate given in literature (Cochran, 1977). Some underestimation of variance is to be expected at least for moderate-sized samples because higher order terms are neglected in delta-method. The underestimation of the variance of the ratio estimate by this method has been confirmed by Krewski and Rao (1981) and Efron (1982).

Resampling methods may be used to estimate standard errors. The most popular resampling methods in the complex sampling designs are: balanced repeated replication (BRR), jackknife and bootstrap. The basic idea behind them is to select subsamples repeatedly from the whole sample, to calculate the statistic of interest for each of these subsamples, and then use the variability among these subsamples or replicate statistics to estimate the variance of the full sample statistics. Here we present a detailed introduction of how the jackknife and bootstrap method are applied in survey problem.

The jackknife method was originally introduced as a technique of bias reduction (Durbin, 1959). However, it has been widely used for variance estimation (Kish and Frankel, 1974). A detailed discussion of jackknife methodology can be found in Efron and Stein (1981) and Shao and Tu (1996). In the 'standard' version, each jackknife replication can be formed by eliminating one PSU from a particular stratum (h') at a time, and increasing the weights of the remaining PSU's in that stratum by using $g_{h'} = n_{h'}/(n_{h'} - 1)$. Each such replication provides an alternative, but an equally valid, estimate of the statistic concerned to that obtained from the full sample.
For a fixed $h' \le H$ e $i' \le n_{h'}$, let :

$$\hat{Z}_{h'i'} = \sum_{h \neq h'} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} z_{hij} + \frac{n_{h'}}{n_{h'} - 1} \sum_{i \neq i'} \sum_{j=1}^{n_{h'i}} w_{h'ij} z_{h'ij} \tag{3}$$

be the analog $\hat{Z}$ after the i'th cluster in the stratum h' is deleted,

$$\hat{\theta}_{h'i'} = T(\hat{Z}_{h'i'}) \text{ and } \hat{\theta}_{h'} = \frac{1}{n_{h'}} \sum_{l=1}^{n_{h'}} \hat{\theta}_{h'i'}$$

Then the jackknife variance estimator for $\hat{\theta}$ is given by:

$$v_{jack} = \sum_{h=1}^{H} (1 - \lambda_h f_h) \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{\theta}_{hi} - \hat{\theta}_h)^2 \tag{4}$$

Another resampling method widely used thanks to an increase in computing power is the bootstrap method. The bootstrap was first introduced by Efron

(1979) for samples of independent and identically distributed (i.i.d.) observations from some distribution F. An overview of the bootstrap theory and applications in the i.i.d. case can be found in Shao and Tu (1996).

A direct extension to surveys samples of the standard bootstrap method developed for i.i.d. samples is to apply the standard bootstrap independently in each stratum. This methodology is often referred to as the naïve bootstrap. Since the naïve bootstrap variance estimator is inconsistent in the case of bounded stratum sample sizes, several modified bootstrap methods were proposed.

Rao, Wu and Yue (1992) proposed a modification of the original bootstrap increasing the applicability of the method, from variance estimation for smooth statistics to the inclusion of non-smooth statistics as well.

Assuming $n_h \geq 2$, the bootstrap variance estimator for $\hat{\theta} = g(\hat{Z})$ is obtained calculating $\hat{\theta}^* = g(\hat{Z}^*)$ after the bootstrap sample is obtained with

$$\hat{Z}^* = \sum_{h=1}^{H} \left[\sqrt{\frac{(1-\lambda_h f_h) m_h}{n_h - 1}} \bar{z}_h^* + \left(1 - \sqrt{\frac{(1-\lambda_h f_h) m_h}{n_h - 1}} \right) \bar{z}_h \right] \tag{5}$$

where $\bar{z}_h^* = m_h^{-1} \sum_{i=1}^{m_h} z_{hi}^*$, $z_{hi}^* = \sum_{i=1}^{m_h} n_h w_{hij}^* z_{hij}^*$, $\bar{z}_h = n_h^{-1} \sum_{i=1}^{n_h} z_{hi}$, z_{hij}^* is the bootstrap analog of z_{hij} and w_{hij}^* is the bootstrap survey weight.

The bootstrap estimator variance of $\hat{\theta}$ is

$$v_{boot}(\hat{\theta}) = E_*(\hat{\theta}^* - E_* \hat{\theta}^*)^2 \tag{6}$$

where $\hat{\theta}^* = g(\hat{Z}^*)$ and E_* is the expectation respect to the bootstrap sampling. To estimate the variance of the estimator, the following steps (i) and (ii) are independently replicated B times, where B is quite large:

(i) Independently in each stratum h, select a bootstrap sample by drawing a simple random sample of m_h PSU's with replacement from the sample PSU's. Let m_{hi}^* be the number of times that PSU hi is selected $(\sum_i m_{hi}^* = m_h)$ in the bootstrap sample b (b=1, 2,..., B) and the initial bootstrap weight are rescaled as

$$w_{hij}^* = \left[1 - \left(\frac{(1-\lambda_h f_h) m_h}{n_h - 1} \right)^{\frac{1}{2}} + \left(\frac{(1-\lambda_h f_h) m_h}{n_h - 1} \right)^{\frac{1}{2}} \frac{n_h}{m_h} m_{hi}^* \right] w_{hij}$$

(ii) Calculate $\hat{\theta}_b^*$, the bootstrap replicate of estimator $\hat{\theta}$ by replacing the final survey weights w_{hij} with the final bootstrap weights w_{hij}^* in the formula for $\hat{\theta}$.

The bootstrap variance estimator of $\hat{\theta}$ is given by:

$$v_{boot}(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B}(\hat{\theta}_b^* - \bar{\hat{\theta}}^*)^2 \tag{7}$$

where $\bar{\hat{\theta}}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^*$.

3. ADAPTIVE SAMPLING DESIGNS IN CS SURVEYS

In this section, we illustrate adaptive resampling designs that may be useful in customer satisfaction survey when the phenomenon of dissatisfaction is rare and is concentrated in some cluster of the population. For example, in a job satisfaction survey we may suppose that people sharing the same office have the same degree of dissatisfaction because of environmental condition. Then it might be convenient to adopt adaptive sampling designs in order to increase the number of dissatisfied people in the sample since the management might be interested to analyze the characteristics of the dissatisfied workers.

The use of information gathered during the survey to inform sampling procedures is a key feature that distinguishes adaptive sampling and conventional sampling. In conventional sampling, the sampling design is based entirely on a priori information, and is fixed before the study begins.

Adaptive sampling designs are those in which the selection procedure may depend sequentially on observed values of the variable of interest (Thompson and Seber, 1996). They are sampling designs that may redirect sampling efforts using information gathered during the survey.

A researcher using a conventional sampling design would identify the universe of individuals eligible for sampling before any sampling was actually done, and would not add any eligible individuals discovered during the course of the study. For example, using a conventional sampling design, a researcher interested in dissatisfied customers might administer an interview to a random sample of customers. Suppose during the interview procedure a customer mentions that his social network includes a group of people that are extremely dissatisfied. Using a conventional sampling design, these newly discovered customers could not be added to the group that was fixed before the study began. By contrast, in adaptive sampling, the selection of people to include in the sample adapts based on observations made during the survey. In an adaptive sampling design, the sampling procedure might call for an initial random sample of customers to be interviewed. Then customers who report to be extremely dissatisfied might be asked for the names of several friends that use the same service or buy the same product. These people would then be given the interview, and if they themselves are dissatisfied, asked for the names of several of their friends that use the same service or buy the same product.

When interesting values are observed, sampling intensity may be adaptively increased for neighbouring or linked units. In the CS surveys the neighbourhood is difficult to establish since it is not clear how to define criteria that let us add interesting units. However, in this setting, investigators can decide a protocol that makes the decision to add interesting people dependent on behavioural or characteristics of the person already in the sample.

In the next section, sampling strategies for graph-structured populations will be briefly reviewed (Tompson and Collins, 2002), and a design-based strategy from adaptive cluster sampling will be described and illustrated with a simulation example.

4. TYPES OF ADAPTIVE DESIGNS

As mentioned above, any sampling design that adapts to observations made in the course of the study is adaptive.

Human population with social structure can be conceptualized as graphs, where the nodes of the graph represent people and the edges or arcs linking some nodes to others represent social relationships between people.

Sampling methods such as network sampling, snowball sampling, chain referral sampling, adaptive cluster sampling, and other link-tracing designs in which investigators use links between people to find other people to include in the sample are examples of survey sampling in graphs.

However, a graph sampling design is adaptive if decisions on whether to follow links depend on the observed values in the sample. For example, suppose the variable of interest is an indicator of whether or not an individual is satisfied about a specific service that he receives. If an individual in the sample is asked to name friends that they are supposed to be dissatisfied only if the individual reports to be dissatisfied, the survey is adaptive, whereas it is not adaptive if every person sampled is asked to name dissatisfied friends.

Snowball sampling, as described by Goodman (1961), has been applied to a variety of graph sampling procedures. In one type (Kalton and Anderson, 1986) an initial sample of individuals were asked to name different individuals of the population, who in turn were asked to identify further members and so on, for the purpose of constructing a nonprobability sample or obtaining a frame from which to sample. For example, workers who are dissatisfied might be asked to identify any colleagues who are dissatisfied, the colleagues might be asked to identify any of their colleagues who are dissatisfied, and so on. In this type of design the sampling procedure continues until no new individuals are identified, or the limits of the study's resources (time or financial) are reached. In another type of snowball sampling (Goodman, 1961) individuals in the sample are asked to identify a fixed number of other individuals, who in turn are asked to identify the same number of individuals.

Klovdahl (1989) used the term random walk to describe a variation of the link-tracing sampling procedure in which an initial respondent is asked to identify other members of the population of interest, and one from this list is selected at random to be the next respondent. The pattern continues for a number of waves.

The motivation for using designs like this in practice is that initial respondents may be atypical in their characteristics. For example, in a study of job satisfaction where the satisfaction is a rare phenomenon, the initial dissatisfied respondents may be spread in several office. After a few waves researchers may find research participants who are dissatisfied only in a particular office.

With adaptive allocation designs, the starting point is a sample obtained using a conventional design such as simple or stratified random sampling. Based on the observed values in key variables for the initially selected units, an additional sampling is then concentrated in areas or strata based on the initial observations. For example, an initial stratified random sampling is taken and the dissatisfaction is measured. In strata or group of people where dissatisfaction is highly concentrated, a larger sample is allocated.

In adaptive cluster sampling, an initial sample is selected with a conventional sampling design such as simple random sampling, cluster sampling two-stage sampling or stratified sampling. Whenever a particular variable of interest satisfies a specified condition for an individual in the sample, units in the neighbourhood of that unit are added to the sample. If in turn any of the added units satisfies the condition, still more units are added, and so on. For example, a study of customer satisfaction where dissatisfaction is a phenomenon rare might begin by taking a random sample of customers. Whenever a customer is found to be dissatisfied, the "neighbouring" customers would be sampled. In this example, neighbourhoods may be defined by social or institutional connections as well as geographically.

5. ESTIMATION IN ADAPTIVE DESIGNS

For the development of estimators, it will be convenient to define network as a collection of observation units that share the same linkage pattern. We consider the situation in which if a unit of a network is selected then every unit in the network will be included. More complex sampling procedures might be applied to adaptive sampling, but for illustrative purposes we limit ourselves to adaptive sampling designs with an initial sample selected with replacement.

When a clustered distribution of units is suspected then the population may be suitably sampled by using an adaptive sampling. In this case, the units may be aggregated in network, in such a way that the reference population is constituted by the set of networks $\{A_1, A_2,..., A_k\}$ ($k \leq N_T$).

Since most of CSI estimators may be obtained as a (linear/nonlinear) function of population means, we follow the approach proposed by Di Battista (2003) that allows us to obtain a reduced bias estimators of CSI and consistent estimators of the sampling variances in the case of adaptive sampling designs. Let $\tilde{Z}_l = M_l^{-1}\sum_{j\in A_l} Z_l$ be the mean for the lth network, where M_l represents the number of units in the network l (l = 1, 2,..., k). Now it is well known (Thompson and Seber, 1996) that an adaptive sampling starting from a random sample with replacement of n units actually represents a sampling with replacement of n network, in which at each time the lth network inclusion probability is M_l / N_T (l = 1, 2,...,k). Accordingly, if G denotes a sample of n networks selected from the population by the adaptive design it is that $\bar{Z} = n^{-1} \sum_{l \in G} \tilde{Z}_l$ provides an unbiased estimator of the population mean.

In particular, for measuring customer satisfaction an ample literature exists. The traditional indicators of CS, such as those based on SERVQUAL and SERVPERF instruments, are constructed as a linear combination of manifest variables and the estimate of the variance is quite easy to calculate. In particular, if we are interested to measure how much customers are satisfied in some specific dimension then we can define the following index:

$$CSI_d = \frac{\sum_{j=1}^{N_T} \frac{1}{K} \sum_{k=1}^{K} Y_j^k \delta_j}{\sum_{j=1}^{N_T} \delta_j} \tag{8}$$

where Y_j^k is the kth variable (k=*1,2,..., K*) associated with unit j and δ_j assumes value 1 if the individual j satisfies some condition (i.e. if individual j is globally dissatisfied) and 0 otherwise.
In the adaptive sampling design setup, an estimator of (8) is

$$C\hat{S}I_d = \frac{\sum_{l=1}^{n} \tilde{Z}_l}{\sum_{l=1}^{n} \delta_l} \tag{9}$$

where $\tilde{Z}_l = M_l^{-1} \sum_{j \in A_l} \frac{1}{K} \sum_{k=1}^{K} Y_j^k \delta_j$.

The estimator (9) is a non linear estimator and it can express as a function of means $C\hat{S}I_d = g(\bar{Z}_1, \bar{Z}_2)$, where $\bar{Z}_2$ is an estimator of dissatisfaction ratio defined on some specific condition.

The approach proposed by Di Battista (2003) allow us to a consistent estimators of the sampling variances of (9) using approximation and resampling methods which are showed in the previous sections.

6. SOME SIMULATION RESULTS

In this section we report the results of a simulation study performed in order to illustrate the behaviour of the adaptive sampling in the Customer Satisfaction

Survey when the phenomenon of dissatisfaction is rare and concentrated in some clusters. The simulation mimics what we have observed in a real survey of Student Satisfaction carried out at the University of Chieti-Pescara in 2005.

The simulation study consists of generating three different populations with three different levels of dissatisfaction p = 0.05, 0.10 and 0.15 with N_T = 1000. We generated four ordinal variables and one dichotomous variable from a multivariate normal distribution with zero mean and variance and covariance matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.80 & 0.90 & 0.70 & 0.70 \\ 0.80 & 1.00 & 0.85 & 0.75 & 0.75 \\ 0.90 & 0.85 & 1.00 & 0.70 & 0.90 \\ 0.70 & 0.75 & 0.70 & 1.00 & 0.65 \\ 0.70 & 0.75 & 0.90 & 0.65 & 1.00 \end{pmatrix}.$$

The four ordinal variables based on a five point scale (1 very dissatisfied, 2 dissatisfied, 3 neither satisfied nor dissatisfied, 4 satisfied, and 5 very satisfied) are obtained by fixing different thresholds in order to have the univariate distribution reported in Table 1.

TABLE 1

Univariate distribution of four simulated items

	Level of satisfaction				
	1	2	3	4	5
Y_1	0.05	0.05	0.10	0.20	0.60
Y_2	0.05	0.10	0.05	0.30	0.50
Y_3	0.10	0.15	0.05	0.40	0.30
Y_4	0.10	0.10	0.10	0.40	0.30

The threshold for the dichotomous variable (δ) has been chosen differently for each population such as to ensure the three levels of dissatisfaction mentioned above. This variable may represent the overall satisfaction indicator, and it assumes value 1 if an individual is globally dissatisfied and 0 otherwise.

The main characteristics of the three population are showed in Table 2.

TABLE 2

Enrolment status of students by gender (percentage)

	Enrolment status		
	First year	Second year	Third year
M	15.5	21.5	15.0
F	17.0	14.5	16.5

In order to establish the clusters of dissatisfied students, we have considered the population characteristics of Business and Management Faculty students of the University of Chieti-Pescara observed from a Student Satisfaction Survey carried out in 2005. In particular, around the 50% of students who are globally dissatisfied (p = 0.05, 0.10 and 0.15) are male, attending the second year of college and belonging to the age class 20-21. The remaining 50% of students are roughly equally distributed among students with characteristics reported in Table 3.

TABLE 3

Characteristics of dissatisfied students

Characteristics			
Gender	Enrolment	Attendance	Age
Male	1 year	yes	26-ω
Male	2 year	yes	26-ω
Male	2 year	no	20-21
Male	2 year	no	26-ω
Female	1 year	yes	20-21
Female	1 year	yes	22-23
Female	1 year	yes	26-ω
Female	2 year	yes	20-21
Female	3 year	yes	26-ω
Female	1 year	no	26-ω

The CSI_d has been constructed on the four items and the indicator variable δ.

From the three artificial populations, 2,000 samples of size n = 50, 100 and 200 were drawn both by simple random sampling with replacement and without replacement. Each simulated sample was used to start the adaptive selection in which the networks were obtained by joining the "neighbouring" individuals reporting dissatisfaction condition respect to the variable δ. The rule we have chosen in order to construct a network is the following: if a student is completely dissatisfied with respect to the variable δ then we randomly select three students which are the same gender, belong to the same age class and attend the same classes. If any of these added units satisfies the former condition then we continue the procedure of aggregation. Subsequently, for each final sample the ordinary, the jackknife, bootstrap and delta estimators of CSI_d and their variances respectively were computed.

Because of population considered, there was a non zero probability that all the students in the sample were not dissatisfied, especially for small sample sizes. Hence, in the simulation we can obtain these samples for which CSI_d is indeterminate. We discarded these samples in the simulation as they provided no information relevant to the phenomenon that we want to study.

Considering 2,000 simulations, Table 4 and Table 5 list the expected (effective) sample sizes $E(\nu)$, variances and relative efficiencies, $eff(C\hat{S}I_{d(adap)}) = \text{var}(C\hat{S}I_d) / \text{var}(C\hat{S}I_{d(adap)})$, for the different sampling strategies for a selection of initial sample sizes n. The variance of $C\hat{S}I_d$ is computed considering the sample sizes $E(\nu)$ in place of n. Thus, $\text{var}(C\hat{S}I_d)$ offers one way to compare the adaptive strategies with simple random sampling of equivalent sample size.

The adaptive strategies have a relative advantage respect to random sampling strategies as shown in the last column of Tables 4 and 5. In fact, considering the first example population (p=0.05) with an initial sample size of 50 (initial sampling fraction 0.05), the adaptive sampling strategy increases the expected size by 14.3%, but is almost 1.6 times as efficient as the equivalent non adaptive strategy.

In the real survey, in order to build a network we can ask "interesting" individuals to name other individuals that share the same characteristics such as sex and/or class attendance.

Whereas in our simulation, we constructed the network under the "worst conditions" associating to interesting units other people sharing the same characteristics randomly. Hence, in practice if there are clusters of dissatisfied units then we will expect that the adaptive sampling will perform better than the simulation study shows.

TABLE 4

Variance comparisons with adaptive cluster sampling and initial sample size n (with replacement)

n	$E(\nu)$	$\text{var}(C\hat{S}I_d)$	$\text{var}(C\hat{S}I_{d(adap)})$	$eff(C\hat{S}I_{d(adap)})$
p=0.05				
50	57.14	0.307	0.192	1.599
100	111.69	0.151	0.093	1.614
200	219.05	0.065	0.046	1.416
p=0.10				
50	62.84	0.127	0.098	1.292
100	122.90	0.059	0.044	1.342
200	235.35	0.027	0.020	1.337
p=0.15				
50	69.43	0.097	0.065	1.506
100	133.64	0.041	0.029	1.408
200	251.60	0.023	0.014	1.655

TABLE 5

Variance comparisons with adaptive cluster sampling and initial sample size n (without replacement)

n	$E(\nu)$	$\text{var}(C\hat{S}I_d)$	$\text{var}(C\hat{S}I_{d(adap)})$	$eff(C\hat{S}I_{d(adap)})$
p=0.05				
50	56.887	0.296	0.190	1.556
100	112.456	0.141	0.092	1.536
200	220.317	0.057	0.039	1.453
p=0.10				
50	63.371	0.118	0.097	1.218
100	123.492	0.053	0.041	1.295
200	237.437	0.024	0.018	1.333
p=0.15				
50	69.809	0.089	0.057	1.545
100	134.455	0.039	0.026	1.500
200	254.169	0.018	0.012	1.545

Tables 6 and 7 list the coverage probabilities of the lower confidence bound (CPL), the upper confidence bound (CPU), the two-sided confidence interval (CPI) and the standardized length (length of interval estimate divided by $2z_{1-\alpha/2}\sqrt{\text{mse}}$) of the linearization (LIN), the jackknife (JACK) and bootstrap (BOOT) variance estimators. The relative bias (RB) and the relative stability (RS) are also given, where

$$RB = \frac{\text{simulation mean of variance estimator}}{\text{the true mse}} - 1$$

and

$$RS = \frac{(\text{simulation mse of variance estimator})^{1/2}}{\text{the true mse}}$$

The bootstrap estimators are approximated by the simple Monte Carlo approximation with B=500.

The jackknife has a small relative bias but is quite unstable. The bootstrap and linearization variance estimators tend to underestimate and the phenomenon becomes more pronounced as the sampling number decreases. The bootstrap variance estimator perform slightly better than delta method in terms of relative bias and relative stability. Even though the coverage probabilities for all of the two-sided confidence intervals are not close to the nominal level and the left tail is understated, the right tail is closer to the nominal level for all the three methods especially when sample size increases. The jackknife variance estimator perform better than the other two methods when the sample size increases. In addition, from the simulation results, we can conclude that the jackknife procedure is recommended for estimating the variance for the statistic defined in (8).

TABLE 6

Performances of confidence sets and variance estimators for adaptive cluster sampling with an initial random sampling with replacement (CSI_d, $\alpha = 0.05$)

	n	CPL	CPU	CPI	SEL	RB	RS
p=0.05							
	50	0.859	0.704	0.590	0.713	-0.026	1.756
JACK	100	0.931	0.869	0.843	0.970	0.270	1.972
	200	0.948	0.907	0.902	0.968	0.060	1.217
	50	0.829	0.670	0.529	0.496	-0.572	0.825
BOOT	100	0.910	0.843	0.797	0.806	-0.195	0.803
	200	0.940	0.900	0.891	0.913	-0.079	0.669
	50	0.820	0.660	0.512	0.450	-0.645	0.818
LIN	100	0.895	0.829	0.774	0.726	-0.348	0.724
	200	0.934	0.884	0.874	0.856	-0.199	0.574
p=0.10							
	50	0.920	0.861	0.827	0.931	0.236	2.095
JACK	100	0.947	0.915	0.914	0.983	0.090	0.931
	200	0.947	0.927	0.932	0.986	0.019	0.455
	50	0.899	0.835	0.784	0.773	-0.230	0.859
BOOT	100	0.941	0.905	0.901	0.930	-0.038	0.716
	200	0.943	0.923	0.923	0.962	-0.032	0.433
	50	0.888	0.819	0.762	0.704	-0.364	0.774
LIN	100	0.936	0.894	0.883	0.875	-0.156	0.605
	200	0.938	0.919	0.918	0.938	-0.081	0.401
p=0.15							
	50	0.932	0.920	0.896	0.951	0.083	1.363
JACK	100	0.933	0.941	0.926	0.954	-0.024	0.581
	200	0.964	0.948	0.938	0.992	0.018	0.376
	50	0.916	0.905	0.877	0.870	-0.131	0.750
BOOT	100	0.927	0.937	0.917	0.920	-0.094	0.525
	200	0.944	0.943	0.934	0.974	-0.018	0.370
	50	0.903	0.893	0.853	0.810	-0.255	0.650
LIN	100	0.920	0.931	0.911	0.893	-0.149	0.492
	200	0.943	0.938	0.932	0.963	-0.043	0.349

TABLE 7

Performances of confidence sets and variance estimators for adaptive cluster sampling with an intial random sampling without replacement (CSI_d, $\alpha = 0.05$)

	N	CPL	CPU	CPI	SEL	RB	RS
p=0.05							
JACK	50	0.863	0.713	0.604	0.697	-0.101	1.489
	100	0.925	0.882	0.845	0.948	0.220	1.865
	200	0.939	0.911	0.901	0.919	-0.079	0.638
BOOT	50	0.838	0.676	0.544	0.481	-0.606	0.788
	100	0.906	0.859	0.808	0.792	-0.223	0.803
	200	0.932	0.898	0.882	0.902	-0.065	0.541
LIN	50	0.829	0.666	0.557	0.443	-0.664	0.799
	100	0.893	0.842	0.785	0.714	-0.374	0.709
	200	0.928	0.885	0.867	0.859	-0.188	0.503
p=0.10							
JACK	50	0.930	0.873	0.845	0.927	0.177	1.811
	100	0.948	0.924	0.915	0.978	0.068	0.836
	200	0.953	0.922	0.927	0.950	-0.061	0.402
BOOT	50	0.904	0.833	0.787	0.778	-0.232	0.776
	100	0.933	0.913	0.897	0.928	-0.039	0.646
	200	0.943	0.913	0.915	0.951	-0.030	0.388
LIN	50	0.882	0.806	0.754	0.646	-0.377	0.719
	100	0.927	0.904	0.890	0.872	-0.154	0.574
	200	0.942	0.910	0.913	0.934	-0.077	0.377
p=0.15							
JACK	50	0.920	0.927	0.896	0.954	0.079	1.210
	100	0.930	0.939	0.925	0.952	-0.091	0.531
	200	0.955	0.943	0.932	0.985	-0.081	0.331
BOOT	50	0.911	0.914	0.871	0.882	-0.108	0.758
	100	0.923	0.930	0.922	0.921	-0.087	0.443
	200	0.942	0.942	0.928	0.958	-0.082	0.334
LIN	50	0.897	0.898	0.855	0.814	-0.246	0.658
	100	0.917	0.920	0.915	0.900	-0.121	0.438
	200	0.939	0.937	0.921	0.957	-0.095	0.326

6. CONCLUSION

In this article we have illustrated that adaptive resampling designs may be useful to detect dissatisfied people when the characteristics of dissatisfaction are concentrated in some cluster of population defined on some specific features.

We have showed that adaptive strategies have an evident advantage respect to random sampling strategies. Thus, we suggest practitioners to adopt adaptive sampling designs when some a priori information (i.e. statistics from previous studies) about clusterization of dissatisfaction is available. In fact, adaptive designs can give substantial gains in efficiency and then reduce the cost in terms of time, money and labour compared with simple random sampling designs.

Moreover, we have showed that the jackknife procedure is recommended for estimating the variance of the estimator proposed and building the related confidence interval.

Department of Quantitative Methods and Economic Theory TONIO DI BATTISTA
University of Chieti-Pescara PASQUALE VALENTINI

REFERENCES

T. DI BATTISTA, (2003), *Resampling Methods for Estimating Dispersion Index in Random and Adaptive Designs*, "Environmental and Ecological Statistics", 10, pp. 83-93.

T. DI BATTISTA, D. DI SPALATRO, (1998), *A jackknife method for adaptive cluster sampling*. "Funzioni speciali e applicazioni", A. Bellacicco and A. Laforgia (eds), F. Angeli, Milano, pp. 76-84.

T. DI BATTISTA, D. DI SPALATRO, (1999), *A bootstrap method for adaptive cluster sampling*. in "Classification and data analysis. Theory and Application, Series Studies in Classification, Data Analysis, and Knowledge Organization", M. Vichi and O. Opitz (eds), Springer-Verlag, Heidelberg, pp. 19-26.

W.G. COCHRAN, (1977), *Sampling Techniques*, John Wiley & Sons, New York.

J. DURBIN, (1959), *A note on the application of Quenouille's method of bias reduction to the estimation of ratios*, "Biometrika", 46, pp. 477-80.

B. EFRON, (1979), *Bootstrap methods: Another look at the Jackknife*. "The Annals of Statistics", 7, pp. 1-26.

B. EFRON, (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, "Society for Industrial and Applied Mathematics", Philadelphia, PA.

B. EFRON, C. STEIN, (1981), *The jackknife estimate of variance*, The Annals of Statistics, 9, pp. 586-596.

L.A. GOODMAN, (1961), *Snowball sampling*, "Annals of Mathematical Statistics", 32, p. 148-170.

L. KISH, M.R. FRANKEL, (1974), *Inference from Complex Samples*, Journal of the Royal Statistical Society B, 36, pp. 1-37.

G. KALTON, D.W. ANDERSON, (1986), *Sampling Rare Populations*, "Journal of the Royal Statistical Society", Ser. A, 149, pp. 65-82.

A.S. KLOVDAHL, (1989), *Urban social networks: Some methodological problems and possibilities*, in M. Kochen, "The Small World," Norwood, NJ: Ablex Publishing, pp. 176-210.

L. KISH, M.R. FRANKEL, (1974), *Inference from Complex Samples*, "Journal of the Royal Statistical Society", Series B, 36, pp. 1-37.

D. KREWSKI, J.N.K RAO, (1981), *Inference from Stratified Samples: Properties of the Linearization, Jackknife and Balance Repeated Replication Methods*, "The Annals of Statistics", 9, pp. 1010-1019.

M.L. OLIVER, (1997), *Satisfaction: A behavioural perspective on the consumer*, McGraw-Hill, America.

J.N.K. RAO, (1988), *Variance Estimation in Sample Surveys*, in "Handbook of Statistics", vol. 6, Amsterdam: Elsevier Science, pp. 427-447.

J.N.K. RAO, C.F.J. WU, K. YUE, (1992), *Some recent work on resampling methods for complex surveys*, "Survey Methodology", 18, pp. 209-217.

J. SHAO, D. TU, (1996). *The Jackknife and Bootstrap*, Springer series in statistics. Springer-Verlag, New York.

S.K. THOMPSON, L.M. COLLINS, (2002), *Adaptive sampling in research on risk-related behaviors*, "Drug and Alcohol Dependence", 68, pp. 57-67.

S.K. THOMPSON, G.A.F. SEBER, (1996), *Adaptive Sampling*, Wiley, New York.

SUMMARY

Complex sampling designs for the Customer Satisfaction Index estimation

In this paper we focus on sampling designs best suited to meeting the needs of Customer Satisfaction (CS) assessment with particular attention being paid to adaptive sampling which may be useful. Complex sampling designs are illustrated in order to build CS indices that may be used for inference purposes. When the phenomenon of satisfaction is rare, adaptive designs can produce gains in efficiency, relative to conventional designs, for estimating the population parameters. For such sampling design, nonlinear estimators may be used to estimate customer satisfaction indices which are generally biased and the variance estimator may not be obtained in a closed-form solution. Delta, jackknfe and bootstrap procedures are introduced in order to reduce bias and estimating variance. The paper ends up with a simulation study in order to estimate the variance of the proposed estimator.