Survey Methodology, June 1993
Vol. 19, No. 1, pp. 103–113
Statistics Canada

Stratified Telephone Survey Designs

ROBERT J. CASADY and JAMES M. LEPKOWSKI[1]

ABSTRACT

Two stage random digit dialing procedures as developed by Mitofsky and elaborated by Waksberg are widely used in telephone sampling of the U.S. household population. Current alternative approaches have, relative to this procedure, coverage and cost deficiencies. These deficiencies are addressed through telephone sample designs which use listed number information to improve the cost-efficiency of random digit dialing. The telephone number frame is divided into a stratum in which listed number information is available at the 100-bank level and one for which no such information is available. The efficiencies of various sampling schemes for this stratified design are compared to simple random digit dialing and the Mitofsky-Waksberg technique. Gains in efficiency are demonstrated for nearly all such designs. Simplifying assumptions about the values of population parameters in each stratum are shown to have little overall impact on the estimated efficiency.

KEY WORDS: Random digit dialing; Optimal allocation; Coverage; Relative efficiency.

1. THE CURRENT STATUS OF TELEPHONE SURVEY DESIGNS

The two stage random digit dialing design for sampling telephone households, first proposed by Mitofsky (1970) and more fully developed by Waksberg (1978), has been widely employed in telephone surveys. The Mitofsky-Waksberg technique capitalizes on a feature of the distribution of working residential numbers (hereafter referred to as WRNs) in the U.S.: specifically, the WRNs tend to be highly clustered within banks of consecutive telephone numbers. Currently, only about twenty percent of the possible telephone numbers within the known area code, three digit prefix combinations are WRNs for the United States as a whole. However, if a bank of 100 consecutive telephone numbers can be identified that has at least one known WRN then, on average, over 50 percent of the numbers in the bank will be WRNs. A technique which can identify 100-banks containing WRNs will greatly reduce the amount of screening necessary to identify telephone numbers assigned to households.

The two-stage Mitofsky-Waksberg technique starts by obtaining a list of area code, prefix combinations for the study area (available nationally from BellCore Research; see Lepkowski 1988). A frame of telephone numbers, hereafter referred to as the BellCore Research or BCR frame, is generated by appending all 10,000 four digit suffixes (*i.e.*, 0000 to 9999) to the area code-prefix combinations. The telephone numbers in the frame are grouped into banks of 100 numbers using the area code, three digit prefix, and the first two digits of the suffix to specify each bank. For example, the area code, prefix combination 313/764 will have 100 different 100-banks: 313/764-00, 313/764-01, . . . , 313/764-99. Next, a sample of 100-banks is selected and a single complete telephone number is generated for each selected bank by appending a two digit, randomly selected, number to the bank identifier. Each of these generated telephone numbers is dialed in the first sampling stage and the residential status of each number is determined and recorded. All 100-banks for which the randomly generated number is not a WRN are discarded. A second stage sample of WRNs is selected from all 100-banks for which the randomly generated number is a WRN. Typically an equal number of numbers, say k, are generated in each bank to start the second stage sampling process. When one of these second stage numbers is found to be non-residential, it is replaced by another randomly generated number from the same bank. This process is continued until k WRNs are identified in each bank. The result is a two stage sample based on selection of 100-banks with probabilities proportional to the number of residential numbers in the bank. This methodology has proven to be an excellent technique for identifying 100-banks with WRNs.

This technique has obvious advantages. The proportion of residential numbers within the 100-banks retained for second stage sampling is much higher than for the BCR frame in general, which results in a substantial improvement in efficiency over simple random digit dialing (RDD). It only requires that the complete set of area code, prefix combinations for the study area be known, and that the study staff have access to a random number generator for sampling telephone numbers. Finally, it also affords, in principle, complete coverage of all telephone households in the study area.

The Mitofsky-Waksberg technique also has several disadvantages. For example, not every 100-bank has the required k residential numbers so the second stage random number generation can use all 99 remaining numbers and

[1] Robert J. Casady, Bureau of Labor Statistics, U.S. Department of Labor and James M. Lepkowski, Survey Research Center, University of Michigan.

still not achieve the required k WRNs. In addition, determining the residential status of each generated number, especially at the first stage, can be difficult. For instance, in many rural areas recording equipment which notifies the caller that a number is not in service is not used. Calls to unassigned numbers are switched to a "ringing" machine. In these areas it is difficult to distinguish unassigned numbers from residential numbers where no one is at home during the study period. This difficulty is more noticeable at the end of a study period due to the need to replace non-residential numbers. Numbers generated at the end of the study period as replacements for non-residential numbers at the second stage of sampling have less time to be called. A small residual of unresolved numbers accumulates at the end of the study period, and final determination of residential status is impossible within study time constraints. Procedures for handling these unresolved numbers have been proposed (Burkheimer and Levinsohn 1988), but they often detract from the simplicity of the overall method.

Many of the difficulties with the Mitofsky-Waksberg technique can be reduced in importance through prescreening of telephone numbers and the use of computer assisted interviewing systems. However, these difficulties are not eliminated unless departures are made from the basic simplicity and/or underlying probability sampling principles of the method (see for example Potthoff 1987 and Brick and Waksberg 1991).

Alternatively, lists of published telephone numbers have been employed as a frame. These lists of published numbers are available for the entire country from commercial firms such as Donnelley Marketing Information Systems. A straightforward selection of telephone numbers from such lists provides a very high rate of WRNs (typically at least 85%) but unfortunately does not cover households with unpublished numbers. Comparisons of telephone households with and without published numbers (see, for example, Brunner and Brunner 1971) indicates that substantial bias may result.

Lists of published numbers can be employed in a manner to provide coverage of households with unlisted numbers as well. Groves and Lepkowski (1986) describe a dual frame approach in which a sample of listed numbers is combined with a random digit dialed sample through post-stratification estimation. If coverage of the population is less important, lists of published numbers can be used to identify 100-banks with at least one listed residential number, and sampling can be restricted to such banks. Survey Sampling Inc. (1986), and previously Stock (1962) and Sudman (1973) using reverse directories, selected samples of telephone numbers from this type of 100-bank. Clearly this approach does not provide complete coverage of unlisted telephone households, but it can greatly improve sampling efficiency. In fact these "truncated frame" methods have rates of residential numbers comparable or higher than the Mitofsky-Waksberg technique, and the troublesome replacement of non-residential numbers is not needed. Unfortunately, for many survey organizations, the coverage deficiency caused by truncating the frame is considered to be unacceptable.

The purpose of this paper is to examine stratified designs for the BCR frame as an alternative to frame truncation and Mitofsky-Waksberg designs. As an example of frame stratification, the BCR frame could be partitioned into two strata: a "high density" stratum consisting of residential numbers in 100-banks with one or more listed numbers and a "low density" stratum consisting of all the remaining numbers in the BCR frame. The "cut-off" point between high and low density strata is somewhat arbitrary; a cut-off of two or more listed numbers could reduce the chance that 100-banks are inadvertently included due to a keying error in a telephone number. Direct access to all listed numbers is not required for this stratification scheme. Counts of listed numbers, or any other indicator of the presence of listed telephone numbers in a 100-bank obtained from a reverse directory (in metropolitan areas with such commercial services) or a commercial list for the entire country, would be sufficient. Preliminary work indicates that approximately 50% of the numbers in the high density stratum are WRNs while only about 2% of the numbers in the low density stratum are WRNs. The obvious cost difference of sampling from the two strata can be exploited through differential sample allocation. The telephone numbers in the low density stratum could be further stratified by careful examination of the characteristics of the 100-banks as determined by other data available from the BCR frame and/or the Donnelley list which may result in even further sampling efficiency.

The next section examines the question of the appropriate allocation of sample between the strata when simple random sampling is utilized within each stratum. A key feature of the stratified telephone sample approach is that it permits alternative approaches to sample selection within in the different strata. Several alternatives are presented and discussed in Section 3. Section 4 presents a study of the impact of "non-optimal" sample allocation on design efficiency. The paper concludes with a general discussion contrasting the Mitofsky-Waksberg procedure and stratified designs.

2. THE ALLOCATION PROBLEM FOR STRATIFIED TELEPHONE DESIGNS

2.1 Background

We assume that the basic sampling frame is the collection of all telephone numbers generated by appending four digit suffixes to the BCR list of area-prefix codes. Further, we assume that each household in the target population is "linked" to one and only one telephone number in the basic sampling frame (to avoid complications of unequal probability of selection).

We also assume that we have access (possibly only indirect) to a directory based, machine readable list of telephone numbers. It should be noted that because many households choose not to list their telephone numbers in a directory, any such directory based frame will not contain all of the WRNs. Directory based lists are also by nature out of date so they will omit some numbers that are currently published WRNs while including others that are no longer WRNs.

From a survey design point of view these two frames tend to be radically different. The BCR frame includes all WRNs so it provides complete "coverage" of the households in the target population, but only about 20 percent of the telephone numbers included in the BCR frame are actually WRNs. Thus, the "hit rate" (and hence sampling efficiency) will be quite low for a simple RDD sample design utilizing the BCR frame. In contrast, a typical directory/list frame covers only about 70 percent of the target households, but the hit rate is 85 to 90 percent. In general the sampling efficiency for a simple RDD design using a directory/list frame is far better than can be attained for the BCR frame using the Mitofsky-Waksberg technique. Unfortunately, the low coverage rates associated with directory based frames preclude their use in many cases.

The basic idea of the proposed stratification approach is to utilize information from the directory based frame to partition the BCR frame into two or more strata with disparate hit rates and then allocate the sample to the strata so as to minimize cost (variance) for a specified variance (cost). Hereafter the stratum with the lowest hit rate will be referred to as the residual stratum. The truncated designs discussed earlier can be included in this general type of design if we allow the allocation of no sample to the residual stratum, and use mean squared error in place of variance.

2.2 Basic Notation

Assume that the BCR frame of telephone numbers has been partitioned into H strata based on a 100-bank attribute which can be determined from either the BCR or the directory based frame of telephone numbers. The choice of 100-banks is somewhat arbitrary; banks of from 10 to 500 consecutive numbers could be considered. For the ith stratum let

P_i = proportion of the frame included in the stratum,

h_i = proportion of the telephone numbers in the stratum that are WRNs (*i.e.* the hit rate),

w_i = average proportion of WRNs in the non-empty 100-banks (*i.e.* the average hit rate for non-empty banks),

z_i = proportion of the target population included in the stratum, and

t_i = proportion of 100-banks in the stratum that contain no WRNs.

The average hit rate for the frame is given by $\bar{h} = \sum_{i=1}^{H} h_i P_i$ and the proportion of empty 100-banks in the frame is given by $\bar{t} = \sum_{i=1}^{H} t_i P_i$.

In general only the P_i's will be known with certainty. Data from a joint research project involving the Bureau of Labor Statistics and the University of Michigan were used to provide approximate values for the parameters h_i and w_i for the two strata in the example. Values for the remaining parameters were calculated using the algebraic relationships $t_i = 1 - (h_i/w_i)$ and $z_i = h_i P_i/\bar{h}$. The approximations for all of the frame parameters for the two stratum design are given in Table 1 below; note that for the BCR frame and $\bar{h} \cong .211$ and $\bar{t} \cong .605$. The value of $\bar{h}$ is in close agreement with that given in Waksberg (1978) but the value of $\bar{t}$ is somewhat smaller than the .65 provided by Groves (1977). At this time it is impossible to determine which value of $\bar{t}$ is more accurate; in fact, the value may have changed since 1977. More recently, Tucker, Casady and Lepkowski (1992) estimated the value of $\bar{t}$ to be .616 for 10-banks which supports the lower estimate $\bar{t}$ of for 100-banks.

Table 1

Approximate values of the frame parameters for a two stratum design based on the BCR frame and Donnelley directory list. Stratum 1 consists of all telephone numbers in 100-banks with at least one telephone number on the Donnelley list frame; stratum 2 contains all remaining numbers

Stratum	Proportion of Frame (P_i)	Proportion of Population (z_i)	Hit Rate (h_i)	Proportion of Empty 100-Banks (t_i)	Hit Rate Within Non-empty Banks (w_i)
1	.3804	.9402	.5210	.0300	.5371
2	.6196	.0598	.0204	.9584	.4900

2.3 The Basic Estimation Problem, Sample Designs and Estimators

We assume the telephone numbers in the ith stratum are labeled 1 through M_i. Let

$$d_{ij} = \begin{cases} 1 & \text{if the } j\text{th telephone number in the } i\text{th stratum is a WRN,} \\ 0 & \text{otherwise.} \end{cases}$$

The variable of interest is the household characteristic Y, and y represents the value of Y for a particular household. The population parameter to be estimated is the population mean $\mu = Y./N.$, where $N. = \sum_{i=1}^{H} \sum_{j=1}^{M_i} d_{ij} = \sum_{i=1}^{H} N_i$ and $Y. = \sum_{i=1}^{H} \sum_{j=1}^{M_i} d_{ij}\, y_{ij}$. The term N_i denotes the number of WRNs in the ith stratum and $N.$ denotes the number of WRNs in the population.

Consider two sample designs: (1) simple random sampling without replacement (*i.e.* simple RDD) from the telephone numbers in the BCR frame, denoted as design D_0 and (2) stratified simple random sampling from the BCR frame (*i.e.* independent simple RDD samples are selected from each stratum), denoted as design D_1. Under design D_0 the standard ratio estimator for μ is given $\bar{Y}_0 = \hat{Y}_0/\hat{N}_0$ where $\hat{Y}_0$ and $\hat{N}_0$ are the usual inflation estimators for $Y.$ and $N.$ respectively. The estimator $\bar{Y}_0$ is asymptotically unbiased for μ and its variance is given by $\text{var}(\bar{Y}_0) \cong \sigma^2/m\bar{h}$ where m is the sample size of telephone numbers and σ^2 is the population variance of the y's. For the design D_1 the standard ratio estimator of μ is given by $\bar{Y}_1 = \hat{Y}_1/\hat{N}_1$ where $\hat{Y}_1$ and $\hat{N}_1$ are the standard inflation estimators for $Y.$ and $N.$ under stratified sampling. The estimator $\bar{Y}_1$ is also asymptotically unbiased for μ and

$$\text{var}(\bar{Y}_1) \cong \sum_{i=1}^{H} \frac{z_i^2\sigma_i^2(1 + (1 - h_i)\lambda_i)}{m_i h_i}, \tag{2.1}$$

where $\lambda_i = (\mu_i - \mu)^2/\sigma_i^2$ and m_i, μ_i, and σ_i^2 are the stratum sample sizes, means, and variances, respectively.

2.4 The Cost Model

There are costs associated both with determining the value of the indicator variable d and the value of the characteristic of interest Y. The cost function for determining the indicator variable is denoted by $C_1(\cdot)$, with

$$C_1(d) = \begin{cases} c_1 & \text{if } d = 1 \\ c_0 & \text{if } d = 0. \end{cases}$$

This model allows for the possibility that the cost of determining that a telephone number is not a WRN may be different than determining that a telephone number is a WRN. In fact the cost of determining the status of telephone numbers that are WRNs is usually less. The cost of determining the value of the characteristic Y includes only the *additional cost* of determining the value of y after the value of d has been determined. Letting $C_2(\cdot,\cdot)$ represent this additional cost, with

$$C_2(d,y) = \begin{cases} 0 & \text{if } d = 0 \\ c_2 & \text{if } d = 1. \end{cases}$$

The sum $c_1 + c_2$ represents the cost of a "productive" sample selection and c_0 represents the cost of an "unproductive" selection, then, following Waksberg (1978), $\gamma = (c_1 + c_2)/c_0$ represents the ratio of the cost of a productive selection to an unproductive selection.

The total cost for sample selection and the determination of the values of Y is a random variable for both design D_0 and D_1. Letting $C(D_0)$ and $C(D_1)$ represent the total cost of conducting a survey under the two respective designs it is straightforward to show that

$$E[C(D_0)] = mc_0(1 + (\gamma - 1)\bar{h}) \tag{2.2}$$

and

$$E[C(D_1)] = c_0 \sum_{i=1}^{H} m_i(1 + (\gamma - 1)h_i). \tag{2.3}$$

2.5 Optimal Allocation for $\bar{Y}_1$

The stratum sample allocation that minimizes var ($\bar{Y}_1$) for a fixed expected total cost C^* (or that minimizes $E[C(D_1)]$ for a fixed variance V^*) is specified up to a proportionality constant by

$$m_i \propto \frac{z_i\sigma_i}{\sqrt{h_i}}\left(\frac{1 + (1 - h_i)\lambda_i}{1 + (\gamma - 1)h_i}\right)^{1/2}, \tag{2.4}$$

where the proportionality constant is determined by substitution into the expected cost equation (or the variance equation, as appropriate). The proportional reduction in variance, relative to RDD sampling, under optimal allocation for fixed cost C^* (or the proportional reduction in cost under optimal allocation for fixed variance V^*) is given by

$$R(\bar{Y}_1,\bar{Y}_0) \cong 1 - \frac{\left[\sum_{i=1}^{H} \frac{z_i\sigma_i}{\sqrt{h_i}}[(1 + (1 - h_i)\lambda_i)(1 + (\gamma - 1)h_i)]^{1/2}\right]^2}{\bar{h}^{-1}\sigma^2(1 + (\gamma - 1)\bar{h})}. \tag{2.5}$$

2.6 Practical Problems Associated With Optimal Allocation

The problem of specifying the values for the parameters in the allocation equations is generic to optimal allocation schemes. For our particular case there are three basic types of parameters: frame related (z_i and h_i), cost related (γ and c_0) and those specific to the variable of interest (λ_i and σ_i^2). Currently, we have a fairly good working knowledge of the frame related parameters for the two stratum example and certain other specific stratification schemes. In Section 5, we will discuss several active research projects which should further expand our knowledge in this area.

It is clear that $\gamma \geq 1$, but the actual value can vary widely. For example, in the case of a multipurpose survey information is collected for several variables, so the costs of determining the status of a telephone number, c_0 and c_1, are in effect amortized over the variables of interest, and γ will probably be considerably larger than unity. On the other hand, if the survey is intended to collect information on only a single variable then the value of γ is probably not much larger than two or three. Waksberg (1978) considers values of γ between 2 and 20.

Potentially the variable specific parameters pose the most serious problem. Usually our knowledge regarding the values of these parameters is limited and, in the case of multipurpose surveys, we must decide which variable(s) to use for the purposes of allocation. Fortunately, in many practical applications, two factors combine to somewhat lessen this problem. First, the allocation tends to be relatively "flat" in a neighborhood of the optimum allocation so that the reduction in variance is relatively robust with respect to allocation. Secondly, in most cases the variables of interest will not be highly related to variables of the type we are using for stratification. Therefore, with caution, we assume that $\lambda_i = 0$ and $\sigma_i^2 = \sigma^2$ for $i = 1, 2, \ldots, H$. Optimal allocation is achieved by

$$m_i \propto \frac{z_i}{\sqrt{h_i}}(1 + (\gamma - 1)h_i)^{-1/2} \tag{2.6}$$

and the proportional reduction in variance is

$$R(\bar{Y}_1, \bar{Y}_0) \cong 1 - \bar{h}\frac{\left[\sum_{i=1}^{H} z_i\left(\frac{1 + (\gamma - 1)h_i}{h_i}\right)^{1/2}\right]^2}{(1 + (\gamma - 1)\bar{h})}. \tag{2.7}$$

In the case of the two stratum example, the allocation specified by (2.6) implies that allocation relative to the residual stratum (*i.e.* m_1/m_2) is 2.54 when $\gamma = 2$ and 1.42 when $\gamma = 10$. In the first case the projected proportional reduction in variance is $R = .283$ and in the second $R = .077$. In fact, it follows from (2.7) that as the relative cost of determining the value of the variable of interest increases, the relative benefit of optimal allocation decreases.

The Mitofsky-Waksberg sample design, denoted by D_3, employs two stages of sample selection (*i.e.* non-empty 100-banks are selected in the first stage and WRNs are selected in the second stage). Following Waksberg (1978), we let $(k + 1)$ be the total number of WRNs selected from each sample 100-bank. The Mitofsky-Waksberg estimator, denoted by $\bar{Y}_3$, is unbiased for μ, and its variance is minimized when

$$k + 1 = \max\left\{1, \left(\frac{(1 - \rho)\bar{t}}{(1 + (\gamma - 1)\bar{h} - \bar{t})\rho}\right)^{1/2}\right\}, \tag{2.8}$$

where ρ is intra-bank correlation. Under this "optimal" within 100-bank sample allocation the reduction in variance, relative to simple RDD, for the estimator $\bar{Y}_3$ is given by

$$R(\bar{Y}_3, \bar{Y}_0) \cong 1 - \frac{\left[(1 + (\gamma - 1)\bar{h} - \bar{t})^{1/2}(1 - \rho)^{1/2} + (\rho\bar{t})^{1/2}\right]^2}{1 + (\gamma - 1)\bar{h}}. \tag{2.9}$$

At the national level Groves (1977) reports that $\rho \cong .05$ for economic or social statistics. Using this value of ρ, together with the values of $\bar{h}$ and $\bar{t}$ from the two stratum example, the projected proportional reduction in variance for the Mitofsky-Waksberg procedure is $R = .281$ when $\gamma = 2$ and $R = .060$ when $\gamma = 10$.

The two methodologies appear to produce essentially identical variance reduction for both values of the cost ratio. However, too much should not be read into this simple comparison as the projected reduction for each of the procedures is based on simplifying assumptions that will not be strictly true for any application. The only inference intended is that the two procedures appear to highly competitive under a general set of circumstances typically encountered in application.

3. ALTERNATIVE SAMPLE DESIGNS

3.1 Truncated Designs

The designs presented in the previous section produce unbiased estimates of the population mean. Incorrect assumptions regarding the various frame, cost, and population parameters only affect the efficiency of the estimators, not their expectations. Unfortunately an extremely high price is paid for the assurance of unbiasedness because sampling from the residual stratum provides information on only a small proportion of the population and at a relatively high cost. For example, suppose we are willing to settle for an estimate of the population mean exclusive of those households linked to telephone numbers in the residual stratum (*i.e.* we "truncate" the original frame by eliminating the residual stratum and select a stratified RDD sample from the remaining telephone numbers). For the two stratum example the "truncated frame" would consist only of those telephone numbers in the first stratum. The hit rate for the sample from the truncated frame would be .521, in contrast to a hit rate of .211 for the entire frame. However, only about 94% of the target population would remain in scope.

In what follows we assume that the truncated frame is simply the original BCR frame less the residual stratum which (without loss of generality) we assume to be stratum H. Accordingly, for the truncated frame $\bar{h}^* = (\bar{h} - P_H h_H)/(1 - P_H)$ is the hit rate, $\bar{t}^* = (\bar{t} - P_H t_H)/(1 - P_H)$ is the proportion of empty 100-banks and $\mu^* = (\mu - z_H\mu_H)/(1 - z_H)$ is the population mean. Let design D_4 be stratified simple random sampling from the truncated frame, and $\bar{Y}_4$ the standard ratio estimator of the population mean. The estimator $\bar{Y}_4$ is asymptotically unbiased for μ^*, and, in general, it is biased for μ. The (asymptotic) bias is given by

$$B(\bar{Y}_4) = \mu^* - \mu = \frac{z_H(\mu - \mu_H)}{(1 - z_H)}. \qquad (3.1)$$

In most practical circumstances the bias tends to zero monotonically as the proportion of the target population in the residual stratum becomes small, although, as indicated by (3.1), this is not necessarily the case. In any event, since the value of $\mu - \mu_H$ is never known, an upper limit on the proportion of the population in the residual stratum is usually the key specification to be determined when considering the use of a truncated frame. For the two stratum example approximately 6% of the target population is excluded from the sampling frame and, in almost all cases, this would not be tolerable for Federal agencies.

The equations for cost, variance, allocation, and proportional reduction in variance (or cost) are essentially the same as those presented in Section 2. In fact the only modifications required for equation (2.1) and equations (2.3) through (2.7) are to replace μ by μ^* and, for $i = 1, 2, \ldots, H - 1$, replace z_i with $z_i^* = z_i/(1 - z_H)$, and replace λ_i with $\lambda_i^* = (\mu_i - \mu^*)^2/\sigma_i^2$. Obviously all sums are only over the remaining $H - 1$ strata. For the special case where only one stratum remains after truncation the proportional reduction in variance (cost) reduces to

$$R(\bar{Y}_4, \bar{Y}_0) = 1 - \frac{\bar{h}(1 + \bar{h}^*(\gamma - 1))}{\bar{h}^*(1 + \bar{h}(\gamma - 1))}. \qquad (3.2)$$

Thus for the two stratum design, the proportional reduction in variance (cost) is approximately $R = .492$ when $\gamma = 2$ and $R = .206$ when $\gamma = 10$. In both cases the reduction is substantially greater than achieved by the two methods in the previous section. However, nearly 6% of the population is not covered by the frame.

In an attempt to retain the relative efficiency of truncation while reducing the magnitude of the coverage problem, BLS and the University of Michigan are investigating several alternative stratification plans in an effort to reduce the proportion of the population in the residual stratum. One promising approach calls for the partition of the residual stratum into two or more residual strata. For example, the partitioning could create a residual stratum 3 consisting of telephone numbers in 100-banks thought to be primarily assigned to commercial establishments or not yet activated for either residential or commercial use. Residual stratum 2 will now contain all other telephone numbers in the residual stratum from the two stratum design D_2. Estimated frame parameters for the resulting three stratum design are given in Table 2.

Table 2

Estimated frame parameters for a proposed three stratum design based on the BCR frame and the Donnelley list frame

Stratum	Proportion of Frame (P_i)	Proportion of Population (z_i)	Hit Rate (h_i)	Proportion of Empty 100-Banks (t_i)	Hit Rate Within Non-empty Banks (w_i)
1	.3804	.9402	.5210	.0300	.5371
2	.2000	.0399	.0420	.9143	.4900
3	.4196	.0199	.0100	.9796	.4900

These data were used to compute the projected proportional reduction in variance for both a three stratum design and a truncated three stratum design in which Stratum 3 is excluded. These results, together with a summary of the results for the two stratum designs and the Mitofsky-Waksberg design, are presented in Table 3 below. (Although not discussed in the text, Table 3 also includes the projected reduction in variance for a cost ratio of 20.)

Table 3

Projected proportional reduction in variance (or cost) relative to simple RDD sampling for five alternative telephone sample designs

Sample Design	Proportional Reduction in Variance or Cost			Proportion of Frame not in Scope
	$\gamma = 2$	$\gamma = 10$	$\gamma = 20$	
Two Stratum	.2829	.0766	.0320	.0000
Two Stratum (Truncated)	.4917	.2055	.1189	.0598
Mitofsky-Waksberg	.2811	.0597	.0135	.0000
Three Stratum	.3001	.0866	.0389	.0000
Three Stratum (Truncated)	.4095	.1574	.0879	.0199

The proposed partitioning strategy successfully reduces the percent of the population out of scope from nearly 6% to approximately 2%. The projected proportional reduction in variance for the truncated three stratum design is approximately $R = .410$ when $\gamma = 2$ and $R = .157$ when $\gamma = 10$. From an efficiency point of view, it occupies the middle ground between the highly efficient truncated two stratum design and unbiased designs.

Of course the issue to be faced when considering such a design is the coverage problem. The design is already subject to non-coverage of the non-telephone household population. Truncating the frame may add to any non-coverage bias already due to this source. For any particular application the risk inherent in sampling from a frame that does not include all of the target population must be weighed against the potential gain in efficiency. As expected, the standard three stratum design is slightly more efficient than the two stratum design. However, the increase in efficiency is so small that it is doubtful that the added cost of partitioning the BCR frame into an additional stratum is justified except for the purpose of truncation.

3.2 Designs Using Optimal Allocation and the Mitofsky-Waksberg Procedure

The final design to be considered is based on the stratified BCR frame. Depending on the proportion of empty 100-banks in the stratum, we use simple RDD sampling in some strata and Mitofsky-Waksberg sampling in others. The motivation for this type of design is based on the following two considerations:

(a) Mitofsky-Waksberg sampling tends to be "administratively complex", and if the gain in efficiency is small, simple RDD is preferred.

(b) It follows from (2.9), applied at the stratum level, that if the proportion of empty banks in a stratum is "small" then Mitofsky-Waksberg sampling offers little, if any, increase in efficiency.

Thus, we propose to utilize simple RDD sampling in strata with a "small" proportion of empty hundred banks and Mitofsky-Waksberg sampling in the remaining strata. The criterion for determining the type of sampling to be utilized is based on equation (2.8) applied at the stratum level. Specifically, if the "optimal" total number of WRNs, as determined by equation (2.8), to be selected from sample 100-banks in a particular stratum is equal to one, then the stratum is designated a simple RDD stratum; otherwise it is designated a Mitofsky-Waksberg stratum. In terms of the proportion of empty hundred banks, the ith stratum will be an RDD stratum if

$$t_i \leq \frac{2.25\rho(1 + h_i(\gamma - 1))}{(1 + 1.25\rho)} \tag{3.3}$$

and a Mitofsky-Waksberg stratum otherwise. For the two stratum example, the first stratum is a RDD stratum, and the second is a Mitofsky-Waksberg stratum for γ equal either 2 or 10.

Formally the proposed sample design is as follows. The BCR frame has been partitioned into H strata and, according to the criteria given in (3.3), simple RDD sampling is specified for the first H_1 strata and Mitofsky-Waksberg sampling is specified for the remaining strata.

Let:

m_i = the number of telephone numbers selected from the ith RDD stratum,

m_i' = the number of WRNs in the sample from the ith RDD stratum,

$\tilde{m}_i$ = the number of 100-banks selected from the ith Mitofsky-Waksberg stratum,

$\tilde{m}_i'$ = the number of retained 100-banks in the ith Mitofsky-Waksberg stratum,

k_i = number of additional WRNs selected from each retained 100-bank, and

$y_{i\cdot}$ = aggregate of y values for the sample WRNs from the ith stratum.

The combined ratio estimator $\bar{Y}_5 = \hat{Y}_5/\hat{N}_5$, where $\hat{Y}_5 = \sum_{i=1}^{H_1} M_i/m_i\, y_{i\cdot} + \sum_{i=H_1+1}^{H} M_i/\tilde{m}_i\, (y_{i\cdot}/k_i + 1)$ and $\hat{N}_5 = \sum_{i=1}^{H_1} M_i/m_i\, m_i' + \sum_{i=H_1+1}^{H} M_i/\tilde{m}_i\, \tilde{m}_i'$, is utilized to estimate the population mean μ and the values of m_i, $\tilde{m}_i$ and k_i are to be chosen to minimize var($\bar{Y}_5$) or the expected cost as specified.

The estimator $\bar{Y}_5$ is asymptotically unbiased for μ and it is straightforward to show that

$$\begin{aligned} \operatorname{var}(\bar{Y}_5) \cong \sum_{i=1}^{H_1} \frac{z_i^2\sigma_i^2}{m_i h_i}(1 + (1 - h_i)\lambda_i) \\ + \sum_{i=H_1+1}^{H} \frac{z_i^2\sigma_i^2}{\tilde{m}_i h_i}\big[1 + (1 - h_i)\lambda_i \\ - k_i(1 - \rho)(k_i + 1)^{-1}\big] \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} E[C(D_5)] = c_0 \Big\{ \sum_{i=1}^{H_1} m_i\big[1 + h_i(\gamma - 1)\big] \\ + \sum_{i=H_1+1}^{H} \tilde{m}_i\big[1 + k_i(1 - t_i) \\ + h_i(k_i + 1)(\gamma - 1)\big]\Big\}. \end{aligned} \tag{3.5}$$

The optimal values of m_i and $\tilde{m}_i$, specified up to a proportionality constant, are given by

$$m_i \propto z_i\sigma_i\left(\frac{1 + (1 - h_i)\lambda_i}{h_i(1 + h_i(\gamma - 1))}\right)^{1/2}, \tag{3.6}$$

for $i = 1, \ldots, H_1$ and

$$\tilde{m}_i \propto z_i\sigma_i\left(\frac{\lambda_i(1 - h_i) + \rho}{h_i t_i}\right)^{1/2}, \tag{3.7}$$

for $i = H_1 + 1, \ldots, H$. The optimal value of $(k_i + 1)$, for $i = H_1 + 1, \ldots, H$, is given by

$$k_i + 1 = \max \left\{ 1, \left(\frac{t_i(1-\rho)}{(1 + h_i(\gamma - 1) - t_i)(\lambda_i(1 - h_i) + \rho)} \right)^{1/2} \right\}. \tag{3.8}$$

The proportionality constant for (3.6) and (3.7) is found by substitution into the expected cost equation or the variance equation as appropriate.

Under optimal allocation the reduction in variance (or cost) relative to simple RDD, is given by

$$R(\bar{Y}_5, \bar{Y}_0) = 1 - \frac{\bar{h}\Phi^2}{\sigma^2(1 + (\gamma - 1)\bar{h})}, \tag{3.9}$$

where

$$\Phi = \sum_{i=1}^{H_1} \frac{z_i \sigma_i}{h_i^{1/2}} (1 + (1 - h_i)\lambda_i)^{1/2} (1 + (\gamma - 1)h_i)^{1/2} + \sum_{i=H_1+1}^{H} \frac{z_i \sigma_i}{h_i^{1/2}} \Big[(\rho + (1 - h_i)\lambda_i)^{1/2} t_i^{1/2} + (1 - t_i + (\gamma - 1)h_i)^{1/2} (1 - \rho)^{1/2} \Big]. \tag{3.10}$$

Under the simplifying assumptions $\lambda_i = 0$ and $\sigma_i^2 = \sigma^2$ for $i = 1, 2, \ldots, H$,

$$\Phi = \sigma \left[\sum_{i=1}^{H_1} \frac{z_i}{h_i^{1/2}} (1 + (\gamma - 1)h_i)^{1/2} \right] + \sigma \left[\sum_{i=H_1+1}^{H} \frac{z_i}{h_i^{1/2}} \Big((\rho t_i)^{1/2} + ((1 - t_i + (\gamma - 1)h_i)(1 - \rho))^{1/2} \Big) \right]. \tag{3.11}$$

When applied to the two stratum frame, this combined sampling strategy yields a proportional reduction in variance of approximately $R = .440$ for $\gamma = 2$ and $R = .157$ for $\gamma = 10$. For both of the cost ratios, the reduction in variance is considerably larger than achieved by any of the unbiased procedures considered previously. In fact, the variance reduction is essentially equivalent to that attained by the three stratum truncated design (which is subject to a bias of unknown magnitude). Thus, on first consideration, this combined sampling strategy appears to be superior to all of the other methods.

Unfortunately there are practical problems which may preclude the use of this sampling design in certain situations. For example, the hit rate in the Mitofsky-Waksberg stratum is very low (only .02) so the number of first stage sample 100-banks must be fairly large in order that the expected number of retained 100-banks is not too small. On the other hand, the *relative* number of first stage sample units allocated to the RDD stratum is considerably larger than allocated to the Mitofsky-Waksberg stratum, therefore a large overall sample size is required (see Table 4). Also, from Table 4, the number of WRNs required from each of the retained 100-banks is relatively large and may actually exceed the number of WRNs in some banks. Clearly both of these problems are more acute for $\gamma = 2$ than for $\gamma = 10$. Therefore, the use of this design is restricted to situations where resources can support a "large" sample, and the cost ratio is moderate to large.

Table 4

First stage allocation ratios and second stage sample sizes for the combined RDD/Mitofsky-Waksberg sample design applied to the two stratum BCR frame

Stratum	$\gamma = 2$		$\gamma = 10$	
	$m_1/\tilde{m}_2$	Sample Size Second Stage	$m_1/\tilde{m}_2$	Sample Size Second Stage
1	28.17	N.A.	14.56	N.A.
2	N.A.	17.00	N.A.	9.00

4. SAMPLE ALLOCATION AND DESIGN EFFICIENCY

In Section 2.6 the problem of specifying the parameters required to optimally allocate the sample to the various strata was considered. It was noted that the variable specific parameters (*i.e.* the λ_i and σ_i^2) tend to pose the most serious problem since we usually have little information regarding their values. For most cases the variables of analytic interest will not be very highly related to the variables used for stratification. Thus it is reasonable to assume that $\lambda_i = 0$ and $\sigma_i^2 = \sigma^2$ for $i = 1, 2, \ldots, H$. Under these assumptions the optimal allocation is given by (2.6) and the proportional reduction in variance is given by (2.7).

It is obvious that for any particular application these assumptions will never be strictly true, so when we allocate according to (2.6) the actual proportional reduction in variance will not be that given exactly by (2.7). Furthermore, allocating according to (2.6) will not provide the maximum reduction in variance which is achieved under the optimal allocation specified by (2.4). Assuming that we plan to allocate according to (2.6) two questions need to be addressed: (1) does (2.7) give a reasonable approximation to the actual reduction in variance, and (2) is the actual reduction in variance reasonably close to the maximum possible reduction in variance? A single simple answer is not possible for either question because the outcome depends on exactly how and to what extent the assumptions failed. In the following we address these question for the two stratum design under three specific cases of model failure which are typical of situations encountered in the "real world". In all three cases the results indicate strongly affirmative answers for both questions.

In the first case we assume that $\sigma_1^2 = \sigma_2^2 \equiv W^2$ but $\lambda_1 \neq \lambda_2$. The projected, the actual, and the maximum reduction in variance were computed for selected values of $\beta = |\sqrt{\lambda_1} - \sqrt{\lambda_2}| = |\mu_1 - \mu_2|/W$ between 0.00 to 0.50 and the results are presented in Table 5 below. Based on our previous discussion regarding the weak relationship between the analytic and stratification variables it would seem highly unlikely that β will ever be larger than 0.50. The results in Table 5 indicate that for both cost ratios and for all selected values of β the actual reduction in variance achieved by allocation under the simplifying assumptions is essentially equivalent to that which would be attained under "optimal" allocation. For both cost ratios the projected reduction in variance is always larger than the reduction actually attained and the difference increases as β becomes larger. However, it should be noted that for $\beta \leq .35$ the percentage difference between the projected reduction and the actual reduction is less than 10% when $\gamma = 10$, and less than 4% when $\gamma = 2$.

Table 5

The projected, the actual, and the maximum proportional reduction in variance for cost ratios of 2 and 10 and values of β between 0.00 and 0.50

β	$\gamma = 2$			$\gamma = 10$		
	Projected Reduction	Actual Reduction	Maximum Reduction	Projected Reduction	Actual Reduction	Maximum Reduction
0.00	.2829	.2829	.2829	.0766	.0766	.0766
0.10	.2829	.2820	.2820	.0766	.0761	.0761
0.20	.2829	.2793	.2794	.0766	.0745	.0746
0.30	.2829	.2748	.2750	.0766	.0720	.0721
0.40	.2829	.2686	.2692	.0766	.0684	.0689
0.50	.2829	.2607	.2619	.0766	.0639	.0649

The second general case considered assumes that the analytic variable is Bernoulli, where p_1 and p_2 represent the proportion of the population with the attribute of interest in stratum 1 and stratum 2, respectively. The projected, the actual, and the maximum proportional reduction in variance were computed for two specific cases of assumption failure, namely $p_2 = .90p_1$ and $p_2 = 1.10p_1$; p_1 was allowed to vary from .05 to .50 and cost ratios of 2 and 10 were considered.

As discussed before it is probably reasonable to assume that p_2 will be within 10% of p_1 in most "real world" situations so these results can be considered general for Bernoulli type analytic variables. The actual reduction in variance was virtually identical to that attained under optimal allocation in all cases; thus, allocation under (2.6) can be considered (near) optimal. The projected reduction in variance was also very close to the actual reduction. When p_2 was smaller than p_1 the actual reduction was always larger than the predicted reduction, and the converse was true when p_2 was larger than p_1. In both cases the maximum difference (which was only about 3.5% of the actual reduction when $\gamma = 2$ and 8.3% of the actual reduction when $\gamma = 10$) occurred when $p_1 = 0.05$ and monotonically decreased as p_1 increased.

In summary; the two cases considered seem to indicate that so long as the assumptions which yield the allocation specified by (2.6) are not radically violated, the variance will be very near that attained under optimal allocation. Furthermore, the proportional reduction in variance given by (2.7) provides an approximation for the actual reduction in variance which is at least accurate enough for the purposes of survey design.

5. CONCLUDING REMARKS

The strengths of the Mitofsky-Waksberg technique for generating telephone samples are clear: high hit rates in the second stage of selection, an efficient method for screening empty banks of telephone numbers, and a conceptually ingenious approach to sample generation. It is a remarkable testimony to the strength of the technique that it is widely considered to be the standard method of random digit dialing with few serious competitors after many years. The weakness of the technique (first stage screening and replacement of non-residential numbers during the data collection) does not, on the surface, seem to be important relative to its general strength. However, these features can cause substantial difficulty, especially in short time-period telephone survey operations.

In this paper stratified designs, based on commercial lists of telephone numbers, are proposed as alternatives to the Mitofsky-Waksberg technique. Both two and three stratum designs are studied in detail. In addition to simple random sampling within each stratum, two general alternatives are considered:

(1) Simple random sampling from all strata except the low density stratum frame where the Mitofsky-Waksberg method is used.

(2) Simple random sampling from all strata except the low density stratum which is not sampled at all.

The basic thesis of this paper is that stratified sampling methods, using strata based on counts of listed telephone numbers, are at least as efficient as the Mitofsky-Waksberg technique. Furthermore, these designs can eliminate the need for the troublesome replacement of non-residential numbers at the second stage, since the only telephone numbers that must be dialed in the high density stratum are those that are generated at the beginning of the study. Specific conclusions include the following:

- For low cost ratios, the two and three stratum designs are as efficient as the Mitofsky-Waksberg approach.
- When numbers can be dropped from the low density stratum, these alternatives designs are much more efficient, but at the price of unknown bias due to excluding part of the target population.
- When cost ratios are high, the two and three stratum approaches are clearly superior.

A critical issue is the magnitude of the bias introduced by dropping the low density stratum. As noted previously, approximately 7% of U.S. households do not have a telephone and truncating the frame may add to the non-coverage bias. As less than 5% of the U.S. household population is expected to be contained in the low density stratum it is likely that the additional coverage bias will not be substantial for many characteristics of the total population. On the other hand, for some characteristics, and for some subgroups of the population, the magnitude of the additional bias may be large enough to be of concern. Further empirical investigations of this population must be conducted.

There are two costs associated with the use of stratified designs that may detract from their use: the cost of the commercial list used to stratify the BCR frame and the overall lower hit rate. The cost of stratifying the frame into high and low density strata is not addressed in this investigation because the requisite information was derived from a specialized research file. The cost of stratification is a fixed cost and therefore will reduce the resources available for data collection. It is not known what the fixed cost will be in the future as arrangements are made with commercial vendors to routinely provide such data. Furthermore, this fixed stratification cost can be amortized over multiple studies to greatly reduce its impact on any single sample. It is unlikely that data collection for one time surveys will find either the Mitofsky-Waksberg or the stratification method described here to be as cost-effective as indicated. Further investigation is needed into the frame costs before a complete answer can be found.

The second cost issue concerns the lower hit rates presented in this paper. Given the relative competitive efficiencies of the alternatives considered here, it appears that the lower hit rates do not seriously detract from the efficiency of the alternatives. It may be possible to improve the hit rates in the high density stratum if smaller banks of numbers are used. For example, in another investigation we have found that 10-banks will have hit rates in the neighborhood of .57 compared to the .52 reported here for 100-banks. Of course, working with 10-banks substantially increases the size of files and processing operations that must be used to generate samples and the cost of a 10-bank frame is likely to be much higher than the 100-bank frame.

The cost models as shown in (2.2) and (2.3) are relatively simple, ignoring many cost differences in the telephone survey process that may be important for comparisons of relative efficiencies of the designs. These cost models allow the allocations to be expressed in a straightforward way, but they do not specifically address the cost components associated with two features of the Mitofsky-Waksberg technique that the alternative designs address; replacement of nonworking numbers and weighting to compensate for exhausted clusters. Thus, the cost models ignore structural cost differences between the Mitofsky-Waksberg approach and the proposed alternatives that, if properly taken into account, could effect the relative efficiency of the two methods.

Clearly the results presented here are insufficient to draw final conclusions about the overall value of these alternative designs. Further cost data and empirical evidence on the size of the bias caused by eliminating the numbers from the low density stratum is required before a final conclusion can be reached.

ACKNOWLEDGMENTS

The support and assistance of Clyde Tucker and Bob Groves is gratefully acknowledged. The findings and opinions expressed in this article are those of the authors and do not necessarily reflect those of the U. S. Bureau of Labor Statistics or the University of Michigan.

REFERENCES

BRICK, J.M., and WAKSBERG, J. (1991). Avoiding sequential sampling with random digit dialing. *Survey Methodology*, 17, 27-41.

BRUNNER, J.A., and BRUNNER, G.A. (1971). Are voluntarily unlisted telephone subscribers really different? *Journal of Marketing Research*, 8, 121-124.

BURKHEIMER, G.J., and LEVINSOHN, J.R. (1988). Implementing the Mitofsky-Waksberg sampling design with accelerated sequential replacement. In *Telephone Survey Methodology*, (Eds. R. Groves, *et al.*) 99-112. New York: John Wiley and Sons.

GROVES, R.M. (1977). An Empirical Comparison of Two Telephone Designs. Unpublished report of the Survey Research Center of the University of Michigan, Ann Arbor, MI.

GROVES, R.M., and LEPKOWSKI, J.M. (1986). An experimental implementation of a dual frame telephone sample design. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 340-345.

LEPKOWSKI, J.M. (1988). Telephone sampling methods in the United States. In *Telephone Survey Methodology*, (Eds. R. Groves, *et al.*) 73-98. New York: John Wiley and Sons.

MITOFSKY, W. (1970). Sampling of telephone households. Unpublished CBS News memorandum, 1970.

POTTHOFF, R.F. (1987). Generalizations of the Mitofsky-Waksberg technique for random digit dialing. *Journal of the American Statistical Association*, 82, 409-418.

STOCK, J.S. (1962). How to improve samples based on telephone listings. *Journal of Advertising Research*, 2, 55-51.

SUDMAN, S. (1973). The uses of telephone directories for survey sampling. *Journal of Marketing Research*, 10, 204-207.

SURVEY SAMPLING, INC. (1986). Statistical characteristics of random digit telephone samples produced by Survey Sampling, Inc. Westport, CT: Survey Sampling, Inc.

TUCKER, C., CASADY, R.J., and LEPKOWSKI, J.M. (1992). Sample allocation for stratified telephone sample designs. To appear, *Proceedings of the Section on Survey Research Methods, American Statistical Association*.

WAKSBERG, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, 73, 40-46.