Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Nonsampling errors in dual frame telephone surveys

by J. Michael Brick, Ismael Flores Cervantes,
Sunghee Lee and Greg Norman

June 2011

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2011
Vol. 37, No. 1, pp. 1-12
Statistics Canada, Catalogue No. 12-001-X

Nonsampling errors in dual frame telephone surveys

J. Michael Brick, Ismael Flores Cervantes, Sunghee Lee and Greg Norman [1]

Abstract

Dual frame telephone surveys are becoming common in the U.S. because of the incompleteness of the landline frame as people transition to cell phones. This article examines nonsampling errors in dual frame telephone surveys. Even though nonsampling errors are ignored in much of the dual frame literature, we find that under some conditions substantial biases may arise in dual frame telephone surveys due to these errors. We specifically explore biases due to nonresponse and measurement error in these telephone surveys. To reduce the bias resulting from these errors, we propose dual frame sampling and weighting methods. The compositing factor for combining the estimates from the two frames is shown to play an important role in reducing nonresponse bias.

Key Words: Nonresponse bias; Measurement error; Calibration; Sample allocation; Composite.

1. Introduction

Dual frame telephone surveys that sample from both landline and cell phones have become important in the U.S. to reduce undercoverage bias due to the incompleteness of the landline frame. Blumberg and Luke (2009) show that the percentage of households without a landline telephone but with at least one cell phone has increased dramatically in the last few years, reaching 20 percent by the end of 2008. Other countries also report substantial increases in the percentages of people who have only a cell phone (*e.g.*, Kuusela, Callegaro and Vehovar 2008; Vicente and Reis 2009).

This paper uses data from the California Health Interview Survey (CHIS) and from 8 surveys conducted for the Pew Research Center for the People & the Press to examine the effects of nonsampling errors in dual frame telephone surveys. The CHIS 2007, a survey of California adults, was undertaken in late 2007. It combines a standard landline survey with a screening sample of cell phone numbers, where adults from the cell sample were interviewed only if they indicated that they did not have a landline number in the household. The Pew surveys are national surveys that interviewed an adult at all sampled residential telephone numbers from both landline and the cell samples. These surveys are described in more detail later. A number of important issues associated with the effect of nonsampling errors have been identified as a result of undertaking these dual frame telephone surveys – errors that have not been investigated fully in other studies.

In the next section we review sample design, weighting and variance estimation methods developed for dual frame surveys, and describe CHIS 2007 and Pew dual frame telephone surveys that are used throughout the paper. The third section discusses nonsampling error in dual frame telephone surveys, and the effects these errors may have on the bias of estimates. Nonresponse and measurement errors have special importance in dual frame surveys. The fourth section studies sampling and estimation methods that may be used to alleviate bias in dual frame telephone surveys, and gives conditions under which these sampling and estimation approaches may be most useful. In this section we propose three estimators to reduce the bias due to differential nonresponse within the overlap domain. The final section summarizes some of the findings for dual frame telephone surveys, and speculates on the applicability of these findings for other dual frame surveys.

2. Background

Most of the literature on dual frame surveys deals with the statistical theory related to efficiency in sample design and estimation. We summarize some of the key results in sampling, weighting and variance estimation, and then discuss the application of these methods to dual frame telephone surveys.

2.1 Sampling

The two sampling frames are denoted as A and B, and we assume the samples from these frames, S_A and S_B, are independent. The domain of units that are only in A is a, the domain of units only in B is b, and the intersection containing the overlap units is ab. In our application to telephone surveys, A is the frame of landline numbers, B is the frame of cell phone numbers, a is the domain of households with only landline numbers, b is the domain of households with only cell phone numbers, and ab is the domain of households with both types of telephone service.

1. J. Michael Brick, Westat and Joint Program in Survey Methodology at the University of Maryland. E-mail: mikebrick@westat.com; Ismael Flores Cervantes, Westat; Greg Norman, Westat, 1600 Research Boulevard, Rockville Maryland, 20850 U.S.A.; Sunghee Lee, Institute for Social Research, University of Michigan, 426 Thompson St. Ann Arbor, MI 48104, U.S.A.

Many important features of dual frame surveys depend on how units that could fall into both sampling frames (*ab*) are handled.

A screening dual frame approach attempts to make $ab = \varnothing$ by removing any overlap units before sampling, after sampling but prior to data collection, during data collection, or after data collection. Lohr (2009) gives examples of dual frame surveys using each of these approaches.

Brick, Edwards and Lee (2007) and Fleeman (2007) describe screening in dual frame telephone surveys. While U.S. telephone numbers can be partitioned by whether they are cell or landline numbers, this frame does not identify whether those numbers correspond to households with only landlines (*a*), households with only cell phones (*b*), or households with both types of service (*ab*). In the surveys described by Brick, Edwards and Lee (2007) and Fleeman (2007), households sampled from the cell phone frame (*B*) were screened out during the data collection if they reported having a landline. The CHIS 2007 used this screening approach.

A second approach is called an overlap dual frame survey, and units in the overlap could be sampled from both frames. In this case, estimation methods must be employed to avoid biased estimates because the overlap units have multiple chances of selection. Steeh (2004), Brick, Brick, Dipko, Presser, Tucker and Yuan (2007), and Kennedy (2007) discuss dual frame telephone surveys with overlap. In these cases, all respondents are interviewed irrespective of the frame they are sampled from. The Pew surveys use the overlap approach.

2.2 Estimation

In a screening survey, producing weights for estimating totals and characteristics of the entire population is simple, at least in the absence of nonsampling errors. Since $ab = \varnothing$ and the sampling is independent, the units sampled from each frame are assigned weights that are the inverse of their selection probabilities from the frame from which they were selected. An overall estimate of the total is the sum of the weighted domain estimates, $\hat{y}_{scr} = \hat{y}_A + \hat{y}_b$, where $\hat{y}_A = \sum_{i \in S_A} d_i y_i$ and $\hat{y}_b = \sum_{i \in S_B} d_i \delta_i(b) y_i$, where d_i is the inverse of the selection probability and $\delta_i(b) = 1$ if i is in domain b and 0 otherwise. Variance estimation is also straight-forward since the two frames are strata and variance estimation methods appropriate for stratified samples can be applied. For telephone surveys, the landline sample units are weighted and added to the weighted cell phone sampled units, after the sampled cell phone units that have landlines are given a weight of zero.

Screening during data collection, even in the absence of nonsampling errors, does have implications. For example, screened out households from *B* are not eligible for the interview, and this increases data collection costs and the variance of estimated totals (Kish 1965, Chapter 11). The units that are screened out should also be treated properly as sampled units in variance estimation.

Overlap surveys are more complex because units could be sampled from either of the frames. One estimation approach is to combine the two domain estimates, $\hat{y}_a$ and $\hat{y}_b$ with an average of the estimates of the overlap population from the separate frames. If $\hat{y}_{ab}^A$ and $\hat{y}_{ab}^B$ are the weighted estimates of the overlap domain from frame *A* and frame *B*, respectively, then an average or composite estimator is $\hat{y}_{ave} = \hat{y}_a + \hat{y}_b + \lambda \hat{y}_{ab}^A + (1 - \lambda)\, \hat{y}_{ab}^B$, with $0 \le \lambda \le 1$. Following Lohr (2009) we refer to these as average estimators. Assuming $\hat{y}_a$ and $\hat{y}_b$ are unbiased for domain a and domain *b*, and $\hat{y}_{ab}^A$ and $\hat{y}_{ab}^B$ are both unbiased for domain *ab*, then $\hat{y}_{ave}$ is an unbiased estimator of the total. Estimates of means and other quantities can be produced using weights, where the weights for units in *ab* that are sampled from *A* are multiplied by λ and the weights for overlap units sampled from *B* are multiplied by $(1 - \lambda)$. The choice of the compositing factor, λ, has been investigated by many researchers and specific choices to reduce the variance of the estimates have been suggested by Hartley (1962, 1974) and Fuller and Burmeister (1972). All of average estimators require that the domain for all sampled units can be identified.

Variance estimation with the average estimator is relatively simple if λ is a fixed and not dependent on the selected sample. In this case, $V(\hat{y}_{ave}) = V(\hat{y}_a + \lambda \hat{y}_{ab}^A) + V(\hat{y}_b + (1 - \lambda)\, \hat{y}_{ab}^B)$, and each of these variances can be computed using variance estimation methods appropriate for the separate samples. If λ is sample dependent, as with the Hartley and Fuller and Burmeister estimators, then variance estimation is more complicated. The average estimators with a fixed λ have been used in most dual frame telephone surveys with overlap. This approach is discussed below for the Pew surveys.

Other estimation approaches that have been considered for an overlap survey include the single frame estimator (Bankier 1986; Kalton and Anderson 1986; and Skinner 1991), and the pseudo-maximum-likelihood estimator (Skinner and Rao 1996; Lohr and Rao 2000; and Lohr and Rao 2006). Lohr (2009) reviews these estimators. Nearly all telephone surveys with overlap that we have seen use some versions of the average estimator, and it is the focus of this research.

2.3 Telephone survey applications

Data from CHIS 2007 are used to illustrate issues that arise in dual frame telephone survey that use a screening approach. The CHIS 2007 is a telephone survey of

California's population conducted by the UCLA Center for Health Policy Research in collaboration with the California Department of Public Health, the California Department of Health Care Services, and the Public Health Institute. Data collection for CHIS 2007 was carried out by Westat in late 2007 through early 2008.

In the CHIS 2007 landline sample, one adult was sampled and interviewed in each household. In the cell phone sample, persons living in households with landline phones were screened out; an adult was sampled and interviewed in the cell sample if they lived in a household classified as cell-only. All responding households, including those screened out from the cell phone frame, were asked questions about telephone status and usage. Nearly 49,000 adult interviews were completed from the landline sample, and 825 interviews were completed with cell-only adults. The landline sample response rate was 35.5% in the interview conducted with a household informant, and a 59.4% for the sampled adult. Respective response rates for the sample from the cell frame were 22.1% and 52.0%. Since CHIS 2007 used a screening approach, the reported response rate for the cell-only household informant interview is 30.5%. California Health Interview Survey (2009) discusses details of the study design, including differences between the overall cell phone response rate and the cell-only rate.

In the CHIS 2007, the estimates from the cell phone sample are calibrated to the cell-only adult population in California at the screening stage (prior to nonresponse weight adjustment for the sampled adult). There are some difficulties with obtaining reliable control totals for the calibration at the state level that are discussed later. The two samples from the two frames are independent samples and are treated as such, until the ultimate stage where the two are combined and calibrated to independent totals of the entire adult population of California. This last calibration stage does not include telephone status as a domain.

For dual frame telephone surveys with overlap, we use data aggregated from 8 surveys conducted for the Pew Research Center for the People & the Press in late 2008 through early 2009. (The data for the Pew surveys were provided by Scott Keeter of the Pew Research Center for the People & the Press). All of these are surveys of the entire U.S. adult population. The surveys interview one adult in each sampled household from both frames using nearly identical questionnaires. Over the 8 surveys, nearly 11,300 landline interviews and 3,800 cell phone interviews were completed. The response rates from the different surveys are very similar for the landline and the cell phone samples, with a median difference of one percentage point between the samples from the two frames. The response rates range across the 8 surveys and two frames from 17% to 24%.

In the Pew surveys, like most dual frame telephone surveys with overlap, a calibrated version of the average estimator is employed. Most surveys calibrate to both the telephone status domain counts (number of adults living in households with only cell phones, the number in household with only landlines, and households with both landlines and cell phones), and to demographic variables. The Pew studies are also calibrated to demographic totals including age, education, race/ethnicity, region, and population density of households with adults 18 years of age or older. In addition, they calibrate to totals of telephone status and, within the overlap domain to relative usage of landline and cell phones.

3. Nonsampling errors

Dual frame theory has been developed for ideal conditions – complete response and the absence of other nonsampling errors. Nonsampling errors affect the bias and precision of the estimates in any survey, but their effects in dual frame surveys may be qualitatively different from those in single frame surveys for three reasons. First, nonsampling error in dual frame surveys often makes it difficult to determine the probability of selection of the sampled unit. This occurs when domain membership is ascertained during data collection, and nonresponse and measurement errors make it difficult to determine if a sampled unit is in the overlap. Second, nonsampling error in dual frame surveys may be linked directly, sometimes causally, to the sampling frame especially when data collection approaches differ by frame. Third, sampling from more than one frame adds complexity and creates more opportunities for nonsampling errors to have differential effects.

3.1 Nonresponse effects

Brick, Dipko, Presser, Tucker and Yuan (2006) show that the over-representation of the number of adults in cell-only households that occurs in almost all dual frame telephone samples may be due to nonresponse error. They suggest that this over-representation might be the result of differential accessibility – adults who rarely use cell phones are less likely to answer their cell phone than those who use their cell phones regularly. They did not find the same type of usage-related differential response rates in the landline sample. Kennedy (2007) further explores this type of nonresponse bias by examining the effects on specific estimates.

To evaluate the differential representation, we compare the CHIS 2007 and Pew survey sample distributions by sampling frame and telephone usage to estimates from the National Health Interview Survey (NHIS). The NHIS is a face-to-face survey sponsored by the National Center for Health Statistics with data collected by the U.S. Bureau of Census (the NHIS data were provided by S. Blumberg and J. Luke as a special tabulation). It is the only federal government survey that provides estimates of telephone status and usage (Blumberg and Luke 2009). We define usage for the dual users (those in households with both types of phone service) as cell-mainly and land-mainly, where cell-mainly are persons who live in households that receive all or almost all their calls on their cell phone and land-mainly are the dual users in households that do not receive all or almost all their calls on their cell phone.

To be more comparable to the CHIS figures, Table 1 restricts the NHIS estimates to those from the West region only (NHIS estimates for California are not available). California accounts for 52 percent of the adults in the West. The NHIS figures are population estimates from the first six months of 2008, which is roughly contemporaneous to the CHIS data collection period. The CHIS figures are the unweighted sample dispositions (the weighted dispositions are nearly identical). Even though CHIS used a screening approach, the telephone usage information was collected for every responding household in the cell phone sample. The table shows that the cell phone frame distribution over-represents the percent of adults in cell-only households and under-represents land-mainly adults when compared to the NHIS estimates. The landline respondents over-represent the land-only users and under-represent the cell-mainly dual users. The landline frame differences are more substantial than observed in a 2004 survey as reported in Brick *et al*. (2006).

Table 2 shows the same type of comparison of the NHIS national estimates from the second half of 2008 to the aggregated Pew survey unweighted outcomes (all the surveys were equal probability samples). Similar to the CHIS results, the cell frame distribution from the Pew surveys over-represents the percentage in the cell-only group and under-represents the land-mainly group, but the differences are less substantial than in CHIS. The Pew distribution from the landline sample mirrors the NHIS distribution closely, with a slight under-representation of the cell-mainly group.

Table 1
Percentage distribution of adults from CHIS 2007 and NHIS, by telephone usage

Telephone usage	NHIS West adults in landline households	CHIS 2007 landline distribution	NHIS West adults in cell phone households	CHIS 2007 cell phone distribution
Landline-only	23.5% (1.5%)	34.2% (0.2%)	–	–
Dual – land-mainly	56.6% (1.7%)	53.2% (0.2%)	60.9% (1.7%)	18.5% (0.7%)
Dual – cell-mainly	19.9% (1.4%)	12.7% (0.2%)	21.4% (1.4%)	31.2% (0.9%)
Cell-only	–	–	17.7% (1.3%)	50.3% (0.9%)
Total	100.0%	100.0%	100.0%	100.0%

Notes NHIS-West is the National Health Interview Survey, West Region, first 6 months of 2008, with percentages of all households with that type of service (thanks to S. Blumberg and J. Luke for this special tabulation). CHIS 2007 is the California Health Interview Survey, collected in 2007 and early 2008, with unweighted percentages from the landline and cell frames. In the cell phone sample, usage was obtained in the screening interview. Approximate standard errors given in ().

Table 2
Percentage distribution of adults from Pew surveys and NHIS, by telephone usage

Telephone usage	NHIS adults in landline households	Pew surveys landline distribution	NHIS adults in cell phone households	Pew surveys cell phone distribution
Landline-only	19.4% (0.7%)	23.0% (0.4%)	–	–
Dual – land-mainly	58.8% (0.8%)	62.7% (0.5%)	58.8% (0.8%)	42.3% (0.8%)
Dual – cell-mainly	19.3% (0.7%)	14.4% (0.3%)	18.5% (0.7%)	24.0% (0.7%)
Cell-only	–	–	22.7% (0.7%)	33.7% (0.8%)
Total	100.0%	100.0%	100.0%	100.0%

Notes NHIS is the National Health Interview Survey, second 6 months of 2008, with percentages of all households with that type of service. Pew surveys aggregates 8 surveys conducted for the Pew Research Center for the People & the Press from October 2008 through March 2009, with unweighted percentages from the landline and cell frames. (Thanks to S. Keeter for providing these data). Approximate standard errors given in ().

Both of these surveys exhibit response distributions by frame and usage that are consistent with the accessibility conjecture of Brick *et al.* (2006). This conjecture implies an ordering of those that are most accessible and likely to respond – ordering from the most likely to respond to the least likely to respond in the cell frame is cell-only, cell-mainly, and land-mainly. The special problem due to having two frames is that the ordering in the landline frame is different (land-only, land-mainly, cell-mainly), and the overlap units from the two frames could have very different response rates and biases.

To examine nonresponse bias for a dual frame survey with overlap, suppose both the landline and cell samples are poststratified to telephone status domain totals prior to forming an average overall estimate. The poststratified estimator is

$$\hat{y}_{ps} = \frac{N_a}{\hat{N}_a}\hat{y}_a + \frac{N_b}{\hat{N}_b}\hat{y}_b + \lambda g^A \hat{y}^A_{ab} + (1-\lambda) g^B \hat{y}^B_{ab}, \quad (1)$$

where the poststratification factor for the land-only sample is $N_a / \hat{N}_a$, for the cell-only sample it is $N_b / \hat{N}_b$, and the frame specific poststratification factors for the overlap are $g^A = N_{ab} / \hat{N}^A_{ab}$ and $g^B = N_{ab} / \hat{N}^B_{ab}$ for the landline and cell samples, respectively. The Horvitz-Thompson (HT) estimators of the number of units are $\hat{N}_a$ for the land-only domain, $\hat{N}_b$ for the cell-only domain, and $\hat{N}^A_{ab}$ and $\hat{N}^B_{ab}$ for the overlap domain from the two samples. Since we focus on the overlap, we write

$$\hat{y}_{ps,ab} = \lambda g^A \hat{y}^A_{ab} + (1-\lambda) g^B \hat{y}^B_{ab}. \quad (2)$$

This poststratified estimator differs from the approach suggested by Lohr and Rao (2000), who average and then poststratify rather than poststratify and then average. Both approaches are consistent and approximately unbiased when there are no nonsampling errors.

If we allow for differential response rates by telephone usage within the overlap such as those observed in dual frame telephone surveys, (2) is biased. Let W be the proportion of the overlap that are land-mainly, and let $\overline{Y}_{ml}$ and $\overline{Y}_{mc}$ be the population means for a characteristic for land-mainly and cell-mainly dual users, respectively. The bias of $\hat{y}_{ps,\,ab}$ is

$$b(\hat{y}_{ps,\,ab}) \doteq W N_{ab} (\overline{Y}_{ml} - \overline{Y}_{mc})$$
$$(\lambda r_{l1} r_l^{-1} + (1-\lambda)\, r_{c1} r_c^{-1} - 1), \quad (3)$$

where r_l is the dual user's response rate for the landline sample, r_{l1} is the landline sample response rate of the land-mainly, r_c is the dual user's response rate for the cell sample, and r_{c1} is the cell phone sample response rate of the land-mainly.

To derive (3), we first define land-mainly and cell-mainly domain estimators from the landline sample as $\hat{y}^A_{ab}(ml) = \hat{N}^A_{ml} \overline{y}^A_{ab}(ml)$ and $\hat{y}^A_{ab}(mc) = \hat{N}^A_{mc} \overline{y}^A_{ab}(mc)$, and from the cell sample as $\hat{y}^B_{ab}(ml) = \hat{N}^B_{ml} \overline{y}^B_{ab}(ml)$ and $\hat{y}^B_{ab}(mc) = \hat{N}^B_{mc} \overline{y}^B_{ab}(mc)$. Now assume (a) $E\overline{y}^A_{ab}(ml) = E\overline{y}^B_{ab}(ml) = \overline{Y}_{ml}$ and $E\overline{y}^A_{ab}(mc) = E\overline{y}^B_{ab}(mc) = \overline{Y}_{mc}$; (b) covariances such as $\mathrm{cov}(\hat{N}^A_{ml} / \hat{N}^A_{ab}, \overline{y}^A_{ab}(ml)) = 0$; and, (c) the expected domain totals are simple expressions such as $E\hat{N}^A_{ml} = r_{l1} N_{ml}$, $E\hat{N}^A_{ab} = r_l N_{ab}$, *etc.* Since $E(N_{ab} / \hat{N}^A_{ab})\, \hat{y}^A_{ab} = N_{ab}\, E\{(\hat{N}^A_{ml}\, \overline{y}^A_{ab}(ml) + \hat{N}^A_{mc}\, \overline{y}^A_{ab}(mc)) / \hat{N}^A_{ab}\}$, we can write $E(N_{ab} / \hat{N}^A_{ab})\, \hat{y}^A_{ab} \doteq r_{l1} r_l^{-1} N_{ml} \overline{Y}_{ml} + r_{l2} r_l^{-1} N_{mc} \overline{Y}_{mc} = N_{ab} (r_{l1} r_l^{-1} W(\overline{Y}_{ml} - \overline{Y}_{mc}) + \overline{Y}_{mc})$. A corresponding expression can be written for $Eg^B \hat{y}^B_{ab}$. Combining the two gives (3).

These expressions assume that $E\overline{y}^B_{ab}(ml) = \overline{Y}_{ml}$ and $E\overline{y}^B_{ab}(mc) = \overline{Y}_{mc}$. An alternative approach that does not require this assumption is to posit that there is response propensity associated with telephone usage. The bias in this case would be a function of the response propensities from each frame. We do not examine the response propensity approach here.

Expression (3) shows that when $0 < W < 1$, the bias of $\hat{y}_{ps,ab}$ is zero if (a) $\overline{Y}_{ml} = \overline{Y}_{mc}$; or (b) $\lambda r_{l1} r_l^{-1} + (1-\lambda)\, r_{c1} r_c^{-1} = 1$. Condition (a) is basically the well-known condition from single frame methodology. Condition (b) differs from single frame expressions because the bias depends on both the relative response rates and the compositing factor, λ. The exception is when $r_{l1} r_l^{-1} = r_{c1} r_c^{-1}$, or equivalently $r_{l1} r_{l2}^{-1} = r_{c1} r_{c2}^{-1}$, where r_{l2} is the landline sample response rate of the cell-mainly and r_{c2} is the cell sample response rate of the cell-mainly. In this form, this expression is comparable to the single frame bias expression that shows no bias exists when response rates are constant.

More generally, the value of λ affects the bias of the estimate, not just its variance. The bias can be eliminated by choosing

$$\lambda_0 = \frac{r_l (r_c - r_{c1})}{r_c r_{l1} - r_l r_{c1}}. \quad (4)$$

Since the proportion of the total population covered by the landline frame is approximately equal to the proportion covered by the cell phone frame, most applications have used $\lambda = 0.50$ without considering its effect on bias.

We can now apply these expressions to evaluate the bias of dual frame telephone estimator for CHIS, assuming the bias is only from differential nonresponse in the overlap. Using the data in Table 1, $W = 0.74$ for the NHIS West region. We approximate $r_{l1} r_l^{-1}$ by the relative poststratification factor that is the ratio of the percentage of the CHIS landline sample classified as land-mainly to the percentage of the NHIS adults in landline households that are land-mainly; $r_{c1} r_c^{-1}$ is computed similarly for the cell phone quantities. The quantities estimated from CHIS 2007 are given in Table 3, $r_{l1} r_l^{-1} \doteq 1.09$ for the landline sample, and $r_{c1} r_c^{-1} \doteq 0.50$ for the cell sample. As an example, suppose $\bar{Y}_{ml} = 0.3$ and $\bar{Y}_{mc} = 0.5$, then the bias of the estimated percentage based on (3) is approximately 3 percentage points (a relative bias of about 9%) if $\lambda = 0.5$. Using (4), the bias is zero when $\lambda \doteq 0.84$; the bias becomes negative for larger values of λ.

Table 3
Within overlap, relative poststratification factors for CHIS 2007 and Pew surveys

Relative poststratification factors*	CHIS 2007	Pew surveys
$r_{l1} r_l^{-1} \doteq g^A / g^A{}_{ml}$	1.09	1.07
$r_{l2} r_l^{-1} \doteq g^A / g^A{}_{mc}$	0.50	0.84
$r_{c1} r_c^{-1} \doteq g^B / g^B{}_{ml}$	0.74	0.78
$r_{c2} r_c^{-1} \doteq g^B / g^B{}_{mc}$	2.42	1.51

* Poststratification adjustment factor for telephone usage domain within overlap divided by overlap poststratification factor.

The same computations can be done using the data from the Pew surveys, and the estimates are also shown in Table 3. The parameters differ substantially from those computed from CHIS. Since the Pew studies are national, the NHIS estimate is $W = 0.81$. The ratios of the Pew figures to the NHIS also have lower variability than those from the CHIS, with $r_{l1} r_l^{-1} \doteq 1.07$ and $r_{c1} r_c^{-1} \doteq 0.84$. As a result, the bias is only approximately 1 percentage points when $\lambda = 0.5$. The bias is zero when $\lambda \doteq 0.7$.

To evaluate the biases more completely, estimates of $\bar{Y}_{ml} - \bar{Y}_{mc}$ are needed for characteristics from a dual frame telephone survey rather than making arbitrary assumptions as done in the example above. Blumberg and Luke (2009) give estimates that suggest these differences may be as substantial as the differences between the cell-only and landline population that have been documented extensively elsewhere. However, the NHIS estimates are from a face-to-face survey, not a dual frame telephone survey.

Keeter, Dimock and Christian (2008) give estimated characteristics for dual telephone users by sampling frame, but not in sufficient detail to compute the biases. Keeter's estimates indicate the estimates of dual users from the cell frame might be closer to the NHIS overlap estimates than those from the landline frame. However, since the response rates within the overlap are more variable from the cell frame than from the landline frame, a screening design that aims to reduce bias should exclude dual users from the cell phone frame rather than the landline frame when the cell frame has more variable response rates by frame.

Because of the potential bias in the overlap design, Brick *et al*. (2006) suggest using a screening design that excludes adults in dual usage households if they were sampled from the cell frame. In a screening design, a bias still exists due to the differential nonresponse in the landline sample of dual users by telephone usage. Substituting $\lambda = 1$ into (2) and (3), the bias of $\hat{y}_{scr,ab} = g^A \hat{y}^A_{ab}$ is

$$b(\hat{y}_{scr,ab}) = WN_{ab}(\bar{Y}_{ml} - \bar{Y}_{mc})(r_{l1} r_l^{-1} - 1). \quad (5)$$

The bias for this design and estimator is equivalent to single frame estimators, with the bias vanishing when either $\bar{Y}_{ml} = \bar{Y}_{mc}$ or the landline response rates are the same for the land-mainly and the cell-mainly. Notice that in this design, there is no compositing factor that can be used to control the bias.

The bias of the screener estimator for CHIS 2007 is about half that of the average estimator using $\lambda = 0.50$ (the screener bias is 1.3 percentage points compared to the post-stratified average estimator using $\lambda = 0.50$ with bias of -3.3 points). With the Pew parameters, the bias of the post-stratified average estimator and the screener estimator are nearly equal, with the bias of the screener slightly greater than the poststratified estimator (the screener bias is 1.1 percentage points compared to -0.7 points for the post-stratified overlap).

An issue mentioned earlier is that domain totals for poststratification, even for telephone status alone (land-only, cell-only, and dual domains), are not generally available for state or local area surveys. While small area estimates of the percentage of adults who are cell-only at the state level have been published (Blumberg, Luke, Davidson, Davern, Yu and Soderberg 2009), these do not give small area estimates for all three domains. The situation for telephone usage control totals is even more limited, with only national NHIS estimates published. Since the response rates in the cell frame typically vary by usage, some assumptions about the response rates in the cell sample may be useful to avoid substantial over-representation of cell-only and cell-mainly adults from the cell frame sample when using the overlap design.

3.2 Measurement error effects

In addition to nonresponse, some of the differences in the distributions shown in tables 1 and 2 could be due to measurement error. Before we discuss hypotheses related to measurement error, some of the key procedures in the surveys that could be related to measurement error are discussed. There are fundamental differences in the surveys, such as mode and topic. The NHIS is a face-to-face survey; the CHIS and Pew surveys are telephone surveys. Both NHIS and CHIS are health surveys, while the Pew surveys cover a broad range of topics.

The surveys also use different methods for collecting telephone status and usage. In the NHIS an adult family member is asked to answer questions about telephone status and usage for the entire family in a section of the interview about family characteristics. In the cell phone sample in CHIS 2007, the telephone status items are asked during the household screening, but the usage items are in the sampled adult interview. In the CHIS landline sample and the Pew surveys, the status and usage items are all in one of the last sections of the adult interview. This later placement is possible because no screening is involved.

The sampling of an adult is another procedure that may interact with the measurement process. In the CHIS 2007, an adult is sampled from all adults who share the same cell phone. In the Pew surveys, and most other cell phone surveys, the cell phone is considered a personal device, and the person answering the phone is interviewed. In dual use households, the CHIS and Pew methods may result in different samples of adults.

The greatest potential source of measurement error may be related to differences in the questionnaire items for telephone status and usage in the surveys. The items asked in each survey are given in the appendix. The approaches are quite varied. At least part of the difference in the studies is because the CHIS and Pew surveys are conducted by telephone and have prior information about telephone status.

The items used in all three surveys are derived from items used in a supplement to the Current Population Survey (CPS) in 2004. As discussed in Tucker, Brick and Meekins (2007), cognitive testing and behavioral coding for the supplement identified a number of concerns with the CPS items, especially the usage item. Their testing found that a lack of a specific reference period, not having a code for ''half the time,'' and difficulty in reporting for other members of the household made the usage item susceptible to measurement error. Tucker *et al*. (2007) also highlight the difficulty respondents had in reporting telephone status and usage for all household members in a single item. In addition, respondents had difficulty with understanding the meaning of "landline," "regular," a "working" cell phone, and the difference between using and answering a cell phone.

These issues could affect domain classification, and thus bias estimates. For example, a 23-year-old living with parents might report being cell-only, while the parents might report dual usage. The effects on the estimates of these types of measurement errors in the NHIS and telephone surveys are difficult to predict, but inconsistent reporting in telephone and face-to-face administrations is not unexpected.

Another possible measurement problem is the relationship between reporting telephone usage and the sampling frame from which respondents were selected. The hypothesized error arises if the respondent, when asked which device they use to receive most of their calls, is more likely to choose the device they are using to do the interview. We do not believe this hypothesis has been tested, but any device effect of this nature would be expected to be in the same direction as the nonresponse effect. A dual user should have a greater likelihood of reporting as cell-mainly if sampled from the cell frame; they should be more likely to report as landmainly if sampled from the landline. Thus, the bias discussed earlier in the context of nonresponse could be arising due to the combined effect of nonresponse and device effect. Without being able to identify the magnitude of these sources of the bias, methods for reducing bias are unclear.

4. Design and estimation approaches with nonsampling errors

Because of the additional issues at play in dual frame surveys, sampling and estimation methods should be designed to account for the most important sources of error rather than focusing solely on sampling error. In this section we address sample design and estimation choices for dual frame telephone surveys within this larger error structure setting.

4.1 Sample design approaches

A key design decision for a dual frame telephone survey is whether to use a screening or full overlap sample design. We begin by exploring the optimal allocation of the sample for overlap and screening designs appropriate for dual frame telephone surveys when simple random samples are selected independently from the two frames and $N_a > 0$, $N_b > 0$, and $N_{ab} > 0.$ We assume throughout that the sample sizes are large enough to ignore the finite population correction factors.

We use a linear expected cost function $E(C) = c_A(n_A + n_B\, c_B\, c_A^{-1}),$ where c_A is the cost of a landline interview, c_B is the cost of a cell phone interview, and n_A and n_B are the number sampled from frames A and B, respectively. Assuming a constant element variance, $\sigma^2,$ the variance of the overlap estimator is $v_{ov}^2 = \sigma^2(N_A(N_a + \lambda^2 N_{ab})n_A^{-1} + N_B(N_b + (1-\lambda)^2 N_{ab})n_B^{-1}).$ The allocation that minimizes the variance with this cost function can be found by standard Lagrangian methods, and is

$$n_{o,\,A} = E(C)\,\tau^{-1}\sqrt{c_A^{-1} N_A (N_a + \lambda^2 N_{ab})}$$

$$n_{o,\,B} = E(C)\,\tau^{-1}\sqrt{c_B^{-1} N_B (N_b + (1-\lambda)^2 N_{ab})}, \quad (6)$$

where

$$\tau = \sqrt{c_A N_A (N_a + \lambda^2 N_{ab})} + \sqrt{c_B N_B (N_b + (1-\lambda)^2 N_{ab})}.$$

For a screening design, a linear cost function appropriate for dual frame telephone surveys is $E(C) = c_A n_A + n_b c_b$, where $c_b = c_B + N_B N_b^{-1} c_s$, n_b is the sampled number of cell-only, and c_s is the cost of screening. The variance of the screening estimator is $v_{sc}^2 = \sigma^2 (N_A^2 n_A^{-1} + N_B N_b n_B^{-1})$. The optimal allocation is just the stratified allocation given by $n_{s,A} = E(C) N_A (c_A N_A + \sqrt{c_A c_b} N_b)^{-1}$ and

$$n_B = \frac{E(C) N_B}{\sqrt{c_A c_b} N_A + c_b N_b},$$

yielding

$$n_b = \frac{E(C) N_b}{\sqrt{c_A c_b} N_A + c_b N_b}$$

cell-only interviews.

With no nonsampling error and a fixed expected cost, the variance for the optimally allocated overlap design is smaller than the variance for the optimally allocated screener design when the cost of screening is large enough so that $\sqrt{c_b} > N_b^{-1}(\tau - N_A \sqrt{c_A})$. When bias is included, the screening design may have smaller mean square error than the overlap design even when this condition holds. In the analysis below, we consider bias but do not account for all the effects of nonsampling error. For example, differential response affects the yield by the sampling frame from which the units are selected thus affecting the allocation and variance of the estimate.

We compare the mean square errors of the screening and overlap designs under the CHIS 2007 parameters given previously. The mean square error is the sum of the variance and the bias squared. The variance is for the overall estimate, but the bias arises only from the overlap under our assumptions. The cost parameters for interviewing and screening cell phones are still not very well-known, but we use $(c_A = 1, c_B = 3, c_s = 2)$ based on information given by Keeter *et al.* (2008) and Edwards, Brick and Grant (2008). The other parameters needed for the comparison are the distribution of the population by telephone status domain, and we approximate national values from the 2008 NHIS national estimates ($N_a = 0.2N$, $N_b = 0.2N$, and $N_{ab} = 0.6N$). In this situation, the variance based on an optimally allocated overlap design with $\lambda = 0.5$ is slightly smaller than the variance for the optimal screening design (the ratio of the variances is 0.976). The variances of the two designs are approximately the same when the cost parameters are such that the screening from frame B is slightly less expensive $(c_A = 1, c_B = 3, c_s = 1.85)$.

The screening approach has smaller mean square error than the overlap design under these conditions because the screening approach reduces the bias of the estimates from -3.3 percentage points to 1.3 points. Even a relatively small bias dominates the mean square error comparison between the two designs, assuming the bias with the screening approach is half the bias under the overlap design. This is the case because the variances of the overlap and screening designs are so similar. If we instead use the parameters from the Pew surveys, then the mean square error for the overlap design is smaller because its bias is lower than the bias of the screener design.

The allocation to the frames with the overlap approach given by (6) assuming only sampling error is determined by the population parameters, the cost parameters, and the compositing factor. While this is not the optimal allocation when differential response rates are admitted, it is still useful to consider this situation since it is likely to be encountered frequently in practice. In this situation, the bias of $\hat{y}_{ps,ab}$ due to differential nonresponse can be eliminated by choosing λ to satisfy (4). Based on the CHIS parameters, the value that eliminates this bias is $\lambda \doteq 0.84$. If we continue with the cost and population assumptions as above, but set $\lambda = 0.84$, then the optimal allocation given by (6) would select about 75% of the sample from the landline frame. This contrasts with the allocation with $\lambda = 0.5$, in which only 63% is from the landline frame. The choice of the compositing factor is critical. When $\lambda = 0.84$ is used in conjunction with the optimal allocation for the CHIS parameters, the estimator is unbiased and has a variance that is about 5 percent less than the estimator from the optimal screener design.

4.2 Estimation approaches

An approach suggested by Brick *et al.* (2006) is to use a full overlap design with an average estimator for the overlap that is poststratified to telephone usage domain totals, as is done in the Pew surveys. This estimator is unbiased and consistent if the estimates within the domains are unbiased and the domain sample sizes are sufficiently large.

The auxiliary data needed for this poststratification for the entire U.S. are now published regularly from the NHIS. As mentioned above, there are some concerns about using these data as control totals that deserve further study. The control totals needed for this estimator are the number of land-only adults, the number of cell-only adults, and the number of adults who are land-mainly and the number who are cell-mainly (N_{ml} and N_{mc}, respectively). This partitions the dual users into its two components.

An alternative estimator of the overlap total using the same auxiliary data is

$$\begin{aligned}\hat{y}_{sep} &= \frac{N_a}{\hat{N}_a}\hat{y}_a + \frac{N_b}{\hat{N}_b}\hat{y}_b + \lambda_1 g_{ml}^A \hat{y}_{ab}^A(ml) \\ &\quad + (1-\lambda_1) g_{ml}^B \hat{y}_{ab}^B(ml) \\ &\quad + \lambda_2 g_{mc}^A \hat{y}_{ab}^A(mc) + (1-\lambda_2) g_{mc}^B \hat{y}_{ab}^B(mc), \qquad (7)\end{aligned}$$

where the detailed poststratification factors are $g_{ml}^{A} = N_{ml} / \hat{N}_{ml}^{A},\ g_{mc}^{A} = N_{mc} / \hat{N}_{mc}^{A},\ g_{ml}^{B} = N_{ml} / \hat{N}_{ml}^{B},\ g_{mc}^{B} = N_{mc} / \hat{N}_{mc}^{B},$ and $0 \le \lambda_1 \le 1;\ 0 \le \lambda_2 \le 1.$ This estimator, like the others considered thus far, is unbiased and consistent in the absence of nonsampling errors. Like (1), the estimates from each frame are poststratified before being averaged. The primary difference between (1) and (7) is that the dual users in (7) are partitioned and poststratified by usage; it also introduces different compositing factors within the overlap.

The estimator $\hat{y}_{sep}$ may be useful when (1) is biased and usage control totals are available for poststratification. If the expected means within the usage domains are approximately equal $(E\bar{y}_{ab}^{A}(ml) = E\bar{y}_{ab}^{B}(ml) = \bar{Y}_{ml}$ and $E\bar{y}_{ab}^{A}(mc) = E\bar{y}_{ab}^{B}(mc) = \bar{Y}_{mc}),$ then (7) is unbiased for any choice of $0 \le \lambda_1 \le 1$ and $0 \le \lambda_2 \le 1.$ Since bias is not affected by the choice, different compositing factors may be used to reduce the variance of the estimates as is traditionally suggested in the dual frame literature. Table 3 shows that the proportion of respondents in the detailed usage domains varies considerably by the sampling frame, and this might make different compositing factors worthwhile.

Because telephone usage control totals often are not available, we explored modifying (2) to use different compositing factors similar to those used in the overlap for (7). In this case, the goal would be to reduce bias rather than variance. A modified estimator of the overlap total is

$$\hat{y}_{\text{mod},ab} = \lambda_1 g^{A} \hat{y}_{ab}^{A}(ml) + (1 - \lambda_1) g^{B} \hat{y}_{ab}^{B}(ml) + \lambda_2 g^{A} \hat{y}_{ab}^{A}(mc) + (1 - \lambda_2) g^{B} \hat{y}_{ab}^{B}(mc). \quad (8)$$

However, this estimator may not be useful for reducing bias. Earlier, we showed that the bias of $\hat{y}_{ps,ab}$ vanishes when $\lambda_0 = r_l (r_c - r_{c1})(r_c r_{l1} - r_l r_{c1})^{-1}.$ The choice of $\lambda_1 = \lambda_2 = \lambda_0$ in (8) eliminates the bias for both land-mainly and cell-mainly estimates, so that different compositing factors are not useful for bias reduction. The bias of the modified estimator is

$$b(\hat{y}_{\text{mod},ab}) = WN_{ab}(\bar{Y}_{ml}(\lambda_1 r_{l1} r_l^{-1} + (1 - \lambda_1) r_{c1} r_c^{-1} - 1) - \bar{Y}_{mc}(\lambda_2 r_{l1} r_l^{-1} + (1 - \lambda_2) r_{c1} r_c^{-1} - 1)), \quad (9)$$

where we make assumptions similar to those used earlier to approximate the bias of $\hat{y}_{ps,\,ab}$.

Another reason for studying an overlap estimator like (8) is because it is appropriate with sample designs that screen out land-mainly adults from the cell frame. This approach has been considered because the number of cell frame respondents that are classified as land-mainly may be small, and the assumption that $E\hat{y}_{ab}^{B}(ml) = \bar{Y}_{ml}$ may not hold and biases might result.

Setting $\lambda_1 = 1,$ (8) reduces to

$$\hat{y}_{\text{mod}\,\lambda=1,ab} = g^{A} \hat{y}_{ab}^{A}(ml) + \lambda_2 g^{A} \hat{y}_{ab}^{A}(mc) + (1 - \lambda_2) g^{B} \hat{y}_{ab}^{B}(mc). \quad (10)$$

In this design, the landline sample alone is used to estimate both the land-only and the land-mainly totals. Both frames are used to estimate totals for the cell-mainly. If we assume $E\bar{y}_{ab}^{A}(ml) = \bar{Y}_{ml}$ and $E\bar{y}_{ab}^{A}(mc) = E\bar{y}_{ab}^{B}(mc) = \bar{Y}_{mc},$ then we no longer need $E\bar{y}_{ab}^{B}(ml) = \bar{Y}_{ml}$ for (10) to be unbiased. As before, setting $\lambda_2 = r_l (r_c - r_{c1})(r_c r_{l1} - r_l r_{c1})^{-1}$ eliminates the bias in the cell-mainly estimate.

5. Discussion

This exploration of nonresponse and measurement errors in dual frame telephone surveys suggests the effects of these errors may be very important. It leads us to believe that research on nonsampling errors to reduce biases may be more important than research that leads to incremental reductions in sampling error.

The research also reveals shortcomings in our knowledge about nonsampling errors in these surveys. The direction and magnitude of the effects of measurement error are especially unclear. The inconsistencies in some of the findings for the CHIS 2007 and Pew surveys may well be due to measurement errors associated with the different approaches to data collection in these surveys, or to interactions due to the procedures. A thorough investigation of the error sources in dual frame telephone surveys is essential to improve the quality of dual frame telephone surveys, and we believe experiments to assess the effects of measurement error would be especially beneficial.

We did find that the CHIS 2007 and Pew surveys consistently over-represented cell-only and cell-mainly users in samples from the cell phone frame, and the surveys had a slight over-representation of the land-only and land-mainly from the landline frame. However, the degree of over-representation of the domains differed by survey. In the CHIS, the over-representation could have led to substantial biases in the estimates if an overlap survey and a simple average estimator were used. The CHIS used a screening approach to reduce this potential bias, and this appears to have been largely successful. In the Pew surveys, the representation was less differential by frame and the potential for bias was smaller. In these conditions, the overlap approach may have smaller mean square error than a screening approach.

Due to the potential for bias in dual frame telephone surveys with response patterns like the CHIS 2007, we examined sampling and estimation methods that could be implemented to deal with these biases. We found that screening approaches may be competitive or even preferable in dual frame telephone surveys when the bias due to differential nonresponse or measurement error is large. If the bias is not negligible, this finding even holds with small sample sizes. However, these results depend on the choice of the compositing factor and the current practice of choosing $\lambda = 0.5$ should be reconsidered. An alternative is to choose the compositing factor to eliminate the bias of the average estimator. In many cases, this approach not only eliminates the bias, but also may be more efficient.

We examined three estimators that deal with the bias due to differential nonresponse within the overlap domain. The first is $\hat{y}_{ps}$, which uses telephone status as domain control totals. This estimator eliminates the bias due to differential nonresponse when λ_0 is used as the compositing estimator. This compositing factor indirectly uses information on the land-mainly and cell-mainly domain totals in computing response rates by domain and frame. A second estimator, $\hat{y}_{sep}$, eliminates this source of bias more directly by post-stratifying to telephone status and usage control totals. This estimator also permits the use of different compositing factors within the overlap domain to reduce the variance of the estimates. The third estimator that might be used to reduce bias is $\hat{y}_{\mathrm{mod}}$, but this estimator is more pertinent for a sample design that interviews the cell-only and the cell-mainly respondents from the cell frame, along with all respondents from the landline sample. This modified screening design and estimator might be especially attractive if there is concern that the mean of the land-mainly respondents from the cell frame sample is subject to nonresponse bias. All of these estimators could also be raked to additional demographic control totals after combining the two samples.

Given our current state of knowledge, we believe there are important advantages with the full overlap design and $\hat{y}_{ps}$ with λ_0 chosen based on other similar surveys. It is worth observing that even though the CHIS and Pew surveys had very different response patterns, choosing a value of $\lambda_0 = 0.75$ would have reduced the bias substantially for both surveys. An advantage of this estimator over $\hat{y}_{sep}$ in general is that $\hat{y}_{ps}$ is not poststratified to usage domain totals. We suspect that usage domain totals estimated from a face-to-face survey (NHIS) may be subject to substantially different errors than the estimates from telephone surveys. These differences could result in telephone survey estimates that are biased and have underestimated variances. For state and local surveys where even telephone status totals are not well-known, control totals for usage domains are likely to be highly suspect.

A screening design with $\hat{y}_{scr}$ as the estimator has the advantage that it only requires control totals for the entire population and for the cell-only component, such as those estimated from the NHIS. A disadvantage is that, unlike the overlap estimators, there is no compositing parameter that can be used to reduce the bias directly. The more elaborate screening design that interviews cell-only and cell-mainly from the cell frame and uses $\hat{y}_{\mathrm{mod}}$ has merit, but there have been no studies that examine the conditions which would favor this estimator.

A more complete analysis of the effects of nonsampling error would include other factors such as the effect of the differential response rates by frame. For example, we noted that samples from the cell phone frame yield more cell-only households than would be expected. These differential response rates can be addressed in allocating the sample, but we have not done so here. Our exploration of this shows that it results in larger allocations to the landline frame, increases the value of the compositing factor, and makes the screening designs more efficient relative to the overlap designs. The screening design and estimator are still subject to the bias noted above.

While this research concentrated on nonsampling errors in dual frame telephone surveys, we suspect that similar issues exist in many other dual frame surveys, but that these issues may not be recognized. Lohr (2009) mentions nonsampling errors in general dual frame surveys and suggests comparing estimates of the overlap from each frame as a simple diagnostic test. We believe this is an excellent way to begin an investigation of problems associated with the overlap.

As we noted earlier, the handling of the overlap is a major concern in dual frame surveys because nonsampling error may be associated with the sampling frame. Our investigation shows that nonresponse and measurement errors are tied to the sampling frame in dual frame telephone surveys. It is very likely that dual frame telephone surveys that use different modes might experience analogous effects. For example, consider a dual frame household survey designed to survey members of a rare population. Suppose it uses an incomplete membership list with telephone numbers for the rare group as frame A, and an area probability sample of households as frame B. Different response rates by sampling frame within the overlap might be expected, and these might be related to characteristics of the respondents leading to biases. Even within the overlap, there may be differences such as those related to how long the person has been a member of the organization used to create frame A and this might be related to characteristics such as age. This type of situation might parallel some of the within overlap domain issues identified in telephone surveys. Differential measurement errors related to the modes are also possible.

Given the potential for bias in a dual frame survey, one of the important findings of our research is that the compositing factor, λ, influences the bias as well as having an effect on the variance. While the choice of λ typically has only a slight effect on the variance if λ is in the vicinity of the optimal value, the bias may be more sensitive to this choice. Thus, in dual frame surveys understanding how the choice of λ affects the bias and the mean square error of the estimates is an important consideration. The other sampling and estimation methods discussed in this paper may also be applicable to other dual frame surveys. The usefulness of these methods depends upon understanding the nature of the nonsampling errors as well as the availability of auxiliary data that could be used in calibration.

Acknowledgements

We would like to thank Scott Keeter, Stephen Blumberg and Julian Luke for providing data for this paper. We would also like to thank many people for helpful comments on earlier drafts including Sherm Edwards, Ralph DiGaetano, David Grant, David Hubble, Paul Lavrakas, Graham Kalton, Scott Keeter, and Courtney Kennedy.

Appendix

Telephone usage items

National Health Interview Survey

N1. *Is there at least one telephone inside your home that is currently working and is not a cellular phone*?

N2. *Does anyone in your family have a working cellular telephone*?

N3. *How many working cellular telephones do people in your family have*?

[If both N1 and N2 are 'yes' ask N4]

N4. *Of all the telephone calls that your family receives, are …*

All or almost all calls received on cell phones?

Some received on cell phones and some on regular phones?

Very few or none received on cell phones?

California Health Interview Survey – Cell phone

CC1. *Is this cell phone your only phone or do you also have a regular telephone at home*?

[If the phone is a cell phone and they have a regular phone then ask CC2]

CC2. *Of all the telephone calls that you receive, are …*

All or almost all calls received on cell phones

Some received on cell phones and some on regular phones, or

Very few or none on cell phones?

[If respondent replies about half, record it]

California Health Interview Survey – Landline

CL1. *Do you have a working cell phone*?

[If yes or they share a cell phone ask CL2]

CL2. *Of all the telephone calls that you receive, are …*

All or almost all calls received on cell phones

Some received on cell phones and some on regular phones, or

Very few or none on cell phones?

[If respondent replies about half, record it]

Pew Research Center for the People & The Press – Cell phone

PC1. *Now thinking about your telephone use… Is there at least one telephone INSIDE your home that is currently working and is not a cell phone*?

[If yes ask PC2]

PC2. *Of all the telephone calls that you receive, do you get*?

[Rotate options−keeping SOME in the middle]

All or almost all calls on a cell phone

Some on a cell phone and some on a regular home phone

All or almost all calls on a regular home phone

Pew Research Center for the People & The Press – Landline

PL1. *Now thinking about your telephone use… Do you have a working cell phone*?

[If yes ask PL2]

PL2. *Of all the telephone calls that you receive, do you get*?

[Rotate options−keeping SOME in the middle]

All or almost all calls on a cell phone

Some on a cell phone and some on a regular home phone

All or almost all calls on a regular home phone

References

Bankier, M.D. (1986). Estimators based on several stratified samples with applications to multiple frame surveys. *Journal of the American Statistical Association*, 81, 1074-1079.

Blumberg, S.J., and Luke, J.V. (2009). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July-December 2008. National Center for Health Statistics. Available at http://www.cdc.gov/nchs/nhis.htm.

Blumberg, S.J., Luke, J.V., Davidson, G., Davern, M.E., Yu, T. and Soderberg, K. (2009). Wireless substitution: State-level estimates from the National Health Interview Survey, January–December 2007. Hyattsville, MD: National Center for Health Statistics. *National Health Statistics Reports*, 14.

Brick, J.M., Brick, P.D., Dipko, S., Presser, S., Tucker, C. and Yuan, Y. (2007). Cell phone survey feasibility in the U.S.: Sampling and calling cell numbers versus landline numbers. *Public Opinion Quarterly*, 71, 29-33.

Brick, J.M., Dipko, S., Presser, S., Tucker, C. and Yuan, Y. (2006). Nonresponse bias in a dual frame sample of cell and landline numbers. *Public Opinion Quarterly*, 70, 780-793.

Brick, J.M., Edwards, W.S. and Lee, S. (2007). Sampling telephone numbers and adults, interview length, and weighting in the California Health Interview Survey cell phone pilot study. *Public Opinion Quarterly*, 71, 793-813.

California Health Interview Survey (2009). CHIS 2007 Methodology Series: Report 4 – Response Rates. Los Angeles, CA: UCLA Center for Health Policy Research. Available at www.chis.ucla.edu/pdf/CHIS2007_method4.pdf.

Edwards, W.S., Brick, J.M. and Grant, D. (2008). Relative Costs of a Multi-frame, Multi-mode Enhancement to an RDD Survey. Presented at the Annual Conference of the American Association for Public Opinion Research, New Orleans, LA.

Fleeman, A. (2007). Survey Research Using Cell Phone Sample: Important Operational and Methodological Considerations. Presented at the Annual Conference of the American Association for Public Opinion Research, Anaheim, CA.

Fuller, W.A., and Burmeister, L.F. (1972). Estimators for samples selected from two overlapping frames. *Proceedings of the Social Statistics Section*, 245-249.

Hartley, H.O. (1962). Multiple Frame Surveys. *ASA Proceedings of the Social Statistics Section*, 203-206.

Hartley, H.O. (1974). Multiple frame methodology and selected applications. *Sankhyā*, C, 36, 99-118.

Kalton, G., and Anderson, D.W. (1986). Sampling Rare Populations. *Journal of the Royal Statistical Society*, A 149, 65-82.

Kish, L. (1965). *Survey Sampling*. New York : John Wiley & Sons, Inc.

Kennedy, C. (2007). Evaluating the effects of screening for telephone service in dual frame RDD Surveys. *Public Opinion Quarterly*, 71, 750-771.

Keeter, S., Dimock, M. and Christian, L. (2008). Calling Cell Phones in ’08 Pre-election Polls. News Release from The Pew Research Center for the People & the Press. Available at www.pewresearch.org/pubs/1061/cell-phones-election-polling.

Kuusela, V., Callegaro, M. and Vehovar, V. (2008) The influence of mobile telephones on telephone surveys. In *Advances in Telephone Survey Methodology*, (Eds., J.M. Lepkowski, C. Tucker, J.M. Brick, E.D. de Leeuw, L. Japec, P.J. Lavrakas, M.W. Link and R.L. Sangster), New York: John Wiley & Sons, Inc., Chapter 4, 87-112.

Lohr, S. (2009). Multiple frame surveys. In *Handbook of Statistics*: *Sample Surveys_Design*, *Methods and Applications*, (Ed., D. Pfeffermann). Elsevier, Amsterdam, Chapter 4, Vol. 29A.

Lohr, S., and Rao, J.N.K. (2000). Inference in dual frame surveys. *Journal of the American Statistical Association*, 95, 271-280.

Lohr, S., and Rao, J.N.K. (2006). Estimation in multiple-frame surveys. *Journal of the American Statistical Association*, 101, 1019-1030.

Skinner, C.J. (1991). On the efficiency of raking ratio estimation for multiple frame surveys. *Journal of the American Statistical Association*, 86, 779-784.

Skinner, C.J., and Rao, J.N.K. (1996). Estimation in dual frame surveys with complex designs. *Journal of the American Statistical Association*, 91, 349-356.

Steeh, C. (2004). A New Era for Telephone Surveys. Presented at the Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ.

Tucker, C., Brick, J.M. and Meekins, B. (2007). Household telephone service and usage patterns in the U.S. in 2004: Implications for telephone samples. *Public Opinion Quarterly*, 71, 3-22.

Vicente, P., and Reis, E. (2009). The mobile-only population in Portugal and its impact in a dual frame telephone survey. *Survey Research Methods*, 3, 105-111.