

Dual Frame Surveys: Recent Developments and Challenges

Sharon L. Lohr

Abstract In a dual frame survey, independent samples are drawn from two frames whose union covers the population. Most estimators, however, have been developed under the assumption that there is no nonresponse or measurement error. We review estimators for dual frame surveys and examine their properties in the presence of nonsampling errors. We also discuss implications of nonsampling errors for survey design.

Key words: Combining Data from Different Sources, Misclassification, Multiple Frame Survey, Mode Effects, Nonresponse, Sampling for Rare Events

1 Introduction

In an ideal sampling world, we have a finite population $\mathcal{U}$ with N units, with y_i a measurement on unit i in the population. A probability sample $\mathcal{S}$ is taken from the frame, and the finite population total $Y = \sum_{i=1}^N y_i$ is estimated by $\hat{Y} = \sum_{i \in \mathcal{S}} w_i y_i$, where $w_i = 1/\pi_i$ is the sampling weight. The estimator $\hat{Y}$ is unbiased for Y if the sampling frame includes everyone in the target population, if all sampled units respond, and if there is no measurement error. In practice, of course, these assumptions are rarely met. Nonresponse rates are increasing, which means that survey estimates rely more on models and often have nonresponse bias. While sampling frames may be improving, undercoverage of specific subpopulations continues to be a problem. With populations having increasingly diverse languages and technological access, multiple modes may be needed to sample the entire population; this may result in measurement error if there are mode effects. Typically, nonresponse and undercoverage are dealt with through weight adjustments—increasing weights of selected

Sharon L. Lohr
School of Mathematical and Statistical Sciences, Arizona State University, Tempe AZ 85287-1804
USA, e-mail: sharon.lohr@asu.edu

respondents in an attempt to reduce bias. Attempts are made to reduce measurement error through careful survey and questionnaire design and through modeling.

Multiple frame surveys can help address some of these problems. The potential advantages of using a multiple frame survey include:

1. Use of multiple list frames from administrative records, to make more efficient use of administrative data.
2. Multiple mode sampling (for example, using independent samples from a mobile telephone frame and a landline telephone frame).
3. Future use of the internet for data collection. Although the internet presents many coverage and domain specification challenges, it is worthy of consideration because of the potential cost savings and ease of data collection and processing.
4. Improved small area estimation. A national survey could be supplemented with smaller, localized surveys to obtain more precision in those areas.
5. Improved estimation for rare populations. A general population survey may be supplemented by a survey from a frame with a high concentration of members of the rare population.
6. Modular survey design. A multiple frame approach can give more flexibility for design of continuing surveys. As particular frames become less expensive to sample from, the relative allocation of sample size to the different frames can be modified. The modular approach also allows more flexibility in responding to changing needs for data.

At the same time, however, multiple frame surveys are also sensitive to nonsampling errors, and are subject to more types of nonsampling errors than their single frame counterparts. In this paper, we summarize estimators that have been proposed for multiple frame surveys, and discuss some of the research that is completed and research that may be needed for their practical implementation.

We consider multiple frame surveys in which a unit may belong more than one frame; screening multiple frame surveys, in which units belonging to multiple frames are removed from all except one of the frames, are a special case. Fig. 1 depicts two frames that are both incomplete, as considered by Hartley [8]. Frame A might be a frame of landline telephones and frame B might consist of mobile telephone numbers [18]. It is often unknown in advance whether a household member sampled using one frame also belongs to the other frame [3]. There are three domains: domain a consists of units in frame A but not in frame B, domain b consists of units in frame B but not in frame A, and domain ab consists of units in both frames. In the telephone context, domain a contains individuals belonging to a landline-only household, domain b consists of individuals who have only a mobile telephone, and domain ab consists of individuals who have both mobile and landline telephones.

A special case of Fig. 1 occurs if frame B is a subset of frame A, so that domain b is empty. Such is the case in surveys in which frame A is an area frame and frame B is a list frame; every unit on the list is also in the area frame. A dual frame survey is then often much less expensive than a single frame survey from frame A with equivalent precision since the list frame is often much less expensive to sample than

the area frame. Multiple frame samples can often result in substantial cost savings when sampling rare populations [9]; frame B can sometimes be constructed so that it has a high percentage of persons from the rare population, while frame A may be costly but have complete coverage.

More than two frames can be employed as well, as illustrated in Fig. 2 for a three-frame survey in which all frames are incomplete. In this situation, there are seven domains.

2 Estimators in Multiple Frame Surveys

One goal in analyzing data from any survey is often to estimate a population total Y . In a dual frame survey, we can write

Fig. 1 Frames A and B overlap, creating the three domains a , b , and ab .

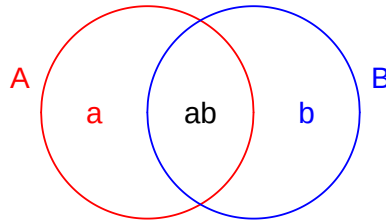
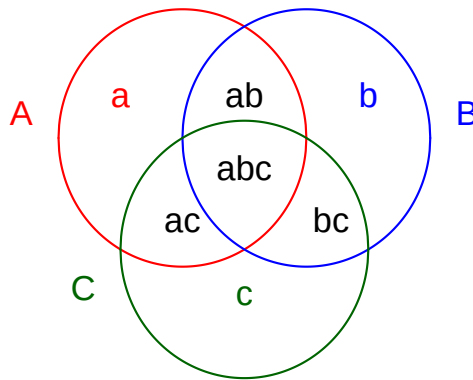


Fig. 2 Frames A, B, and C are all incomplete and overlap.



$$Y = Y_a + Y_{ab} + Y_b,$$

where Y_a is the total of the population units in domain a , Y_{ab} is the total of the population units in domain ab , and Y_b is the total of the population units in domain b . A special case of this is estimating the population size $N = N_a + N_{ab} + N_b$. As multiple frame surveys become more prevalent, the goals expand to include estimation of general population characteristics, fitting models thought to describe the superpopulation, and employing multiple frame surveys in small area estimation. In this section we review estimators for population totals in overlapping multiple frame surveys, along with the weight modifications induced by these estimators. The modified weights can then be used to find point estimates of other population quantities [12].

We first set out some desirable properties for estimators from multiple frame surveys, and then evaluate estimators that have been proposed according to these criteria.

1. An estimator should be approximately unbiased for the corresponding finite population quantity.
2. Estimators should be internally consistent: that is, if $\hat{Y}_1$ estimates the number of female engineers in the population, $\hat{Y}_2$ estimates the number of male engineers in the population, and $\hat{Y}_3$ estimates the total number of engineers in the population, then we should have $\hat{Y}_1 + \hat{Y}_2 = \hat{Y}_3$. Internal consistency preserves multivariate relationships in the data. In practical terms, internal consistency requires that one set of weights be used for all estimates.
3. An estimator should be efficient, with low mean squared error.
4. An estimator should be of a form that can be calculated with standard survey software. In practical terms, one data file would be created from the multiple frame survey. The file would include one column of weights to be used for calculating point estimates, and it would contain either variables describing the survey designs for formula-based variance estimation, or columns of replicate weights for replication-based variance estimation [13].
5. An estimator should, if possible, be robust to nonsampling errors that might occur with multiple frame surveys.

For notational simplicity, we summarize estimators for the overlapping dual frame survey depicted in Fig. 1, where domain ab is nonempty. Let $\mathcal{S}(A)$ and $\mathcal{S}(B)$ denote the probability samples from frames A and B , respectively. Unit i in sample $\mathcal{S}(A)$ has probability of inclusion π_i^A and weight w_i^A , and unit i in sample $\mathcal{S}(B)$ has probability of inclusion π_i^B and weight w_i^B . The weights may be the inverses of the inclusion probabilities, or they may be poststratified to agree with population counts; it is assumed that estimators of population totals from the different samples are approximately unbiased.

Then $E[\sum_{i \in \mathcal{S}(A)} w_i^A y_i] \approx Y_a + Y_{ab}$ and $E[\sum_{i \in \mathcal{S}(B)} w_i^B y_i] \approx Y_b + Y_{ab}$. Consequently, an estimator that combines the observations from both surveys with the original weights, $\sum_{i \in \mathcal{S}(A)} w_i^A y_i + \sum_{i \in \mathcal{S}(B)} w_i^B y_i$, is biased for the population total Y . The various estimators for the population total Y that have been proposed in the literature

modify the weights so that the estimators are approximately unbiased. The modified weights are of the form $\tilde{w}_i^A = m_i^A w_i^A$ and $\tilde{w}_i^B = m_i^B w_i^B$. The population total is then estimated by

$$\hat{Y} = \sum_{i \in \mathcal{S}(A)} \tilde{w}_i^A y_i + \sum_{i \in \mathcal{S}(B)} \tilde{w}_i^B y_i \quad (1)$$

and the population mean $\bar{Y}$ is estimated by $\hat{\bar{Y}} = \hat{Y}/\hat{N}$ where $\hat{N} = \sum_{i \in \mathcal{S}(A)} \tilde{w}_i^A + \sum_{i \in \mathcal{S}(B)} \tilde{w}_i^B$. The estimators are approximately unbiased, then, if $m_i^A \approx 1$ for $i \in a$, $m_i^B \approx 1$ for $i \in b$, and $m_i^A + m_i^B \approx 1$ for $i \in ab$. All of the estimators reviewed in this section satisfy the criteria needed for approximate unbiasedness [12].

For point estimation, a file can be released with one vector of weights, $\tilde{\mathbf{w}} = (\tilde{w}_i^A, i \in \mathcal{S}(A), \tilde{w}_i^B, i \in \mathcal{S}(B))$. Let $\mathbf{y}$ be the corresponding vector of observations, formed by concatenating the observations from $\mathcal{S}(A)$ with the observations from $\mathcal{S}(B)$. Then $\hat{Y} = \tilde{\mathbf{w}}^T \mathbf{y}$. From a user's perspective, once the modified weights are constructed, the procedure followed to estimate population totals is the same as in a single frame survey.

Fixed weight adjustments. The simplest weight modification to preserve approximate unbiasedness, proposed by [8], takes

$$m_{i,\theta}^A = \begin{cases} 1 & \text{if } i \in a \\ \theta & \text{if } i \in ab \end{cases}, \quad m_{i,\theta}^B = \begin{cases} 1 & \text{if } i \in b \\ 1 - \theta & \text{if } i \in ab \end{cases}, \quad (2)$$

where $\theta \in [0, 1]$. Using the modified weights $\tilde{w}_i^A = m_{i,\theta}^A w_i^A$ and $\tilde{w}_i^B = m_{i,\theta}^B w_i^B$ in (1), the resulting estimator $\hat{Y}(\theta)$ can also be expressed using the estimated domain totals $\hat{Y}_a^A = \sum_{i \in \mathcal{S}(A), i \in a} w_i^A y_i$, $\hat{Y}_{ab}^A = \sum_{i \in \mathcal{S}(A), i \in ab} w_i^A y_i$, $\hat{Y}_{ab}^B = \sum_{i \in \mathcal{S}(B), i \in ab} w_i^B y_i$, and $\hat{Y}_b^B = \sum_{i \in \mathcal{S}(B), i \in b} w_i^B y_i$. The estimator

$$\hat{Y}(\theta) = \sum_{i \in \mathcal{S}(A)} m_{i,\theta}^A w_i^A y_i + \sum_{i \in \mathcal{S}(B)} m_{i,\theta}^B w_i^B y_i = \hat{Y}_a^A + \theta \hat{Y}_{ab}^A + (1 - \theta) \hat{Y}_{ab}^B + \hat{Y}_b^B, \quad (3)$$

thus takes a weighted average of the frame A estimator, $\hat{Y}_{ab}^A$, and the frame B estimator, $\hat{Y}_{ab}^B$, of the domain total Y_{ab} . Mecatti [16] calls this a multiplicity estimator when $\theta = 1/2$, since the original weight is divided by the number of frames to which an observation belongs. If $\theta = 0$, then the observations from $\mathcal{S}(A)$ in the intersection domain are discarded from the analysis.

For a fixed value of θ , the estimator $\hat{Y}(\theta)$ gives internal consistency since the same set of adjusted weights is used for all variables. It is simple to use and implement. Depending on the value of θ , however, it can be inefficient. When $\theta = 0$ or 1, the data in the overlap domain from one of the samples are thrown away.

Optimal estimators. Hartley [8] proposes choosing θ in (3) so that the variance of $\hat{Y}(\theta)$ is minimized. The optimizing value of θ is

$$\theta_H = \frac{V(\hat{Y}_{ab}^B) + \text{Cov}(\hat{Y}_b^B, \hat{Y}_{ab}^B) - \text{Cov}(\hat{Y}_a^A, \hat{Y}_{ab}^A)}{V(\hat{Y}_{ab}^A) + V(\hat{Y}_{ab}^B)}.$$

Since the variances and covariances are generally unknown, they must be estimated from the data, yielding $\hat{\theta}_H$. The weight modifications for Hartley's estimator $\hat{Y}(\hat{\theta}_H)$ are given by (2), substituting $\hat{\theta}_H$ for θ . Since $\hat{\theta}_H$ is consistent for θ_H , Hartley's estimator is asymptotically optimal among all estimators of the form $\hat{Y}_a^A + \hat{Y}_b^B + \theta \hat{Y}_{ab}^A + (1 - \theta) \hat{Y}_{ab}^B$. The modified weights $\tilde{w}_{i,H}^A$ and $\tilde{w}_{i,H}^B$ are functions of the variances and covariances of estimated domain totals, however. This has two consequences: (1) the modified weights are random variables, and their variability needs to be accounted for in standard errors of estimators, and (2) the optimal weight modifications will differ for different response variables, leading to internal inconsistency.

Fuller and Burmeister [7] modify Hartley's estimator by using additional information about N_{ab} . Lohr and Rao [13] show that the Fuller-Burmeister estimator has the smallest asymptotic variance among the estimators considered. As with the Hartley estimator, however, the modified weights are random variables that differ for different responses, and in complex sampling designs the Fuller-Burmeister estimator also does not achieve internal consistency.

Pseudo-maximum likelihood estimators. To achieve internal consistency Skinner and Rao [17] propose a pseudo-maximum likelihood (PML) estimator that uses the same weights for all variables. When N_{ab} is unknown, it is estimated by $\hat{N}_{ab}^{PML}(\theta)$, which is the smaller of the roots of the quadratic equation

$$\left[\frac{\theta}{N_B} + \frac{1-\theta}{N_A} \right] x^2 - \left[1 + \theta \frac{\hat{N}_{ab}^A}{N_B} + (1-\theta) \frac{\hat{N}_{ab}^B}{N_A} \right] x + \theta \hat{N}_{ab}^A + (1-\theta) \hat{N}_{ab}^B = 0.$$

They suggest using the value θ_P for θ that minimizes the asymptotic variance of $\hat{N}_{ab}^{PML}(\theta)$. Substituting an estimator $\hat{\theta}_P$ for θ_P , the weight adjustments are:

$$m_{i,P}^A = \begin{cases} \frac{N_A - \hat{N}_{ab}^{PML}(\hat{\theta}_P)}{\hat{N}_{ab}^{PML}(\hat{\theta}_P)}, & i \in a \\ \frac{\hat{N}_{ab}^{PML}(\hat{\theta}_P)}{\hat{\theta}_P \hat{N}_{ab}^A + (1-\hat{\theta}_P) \hat{N}_{ab}^B} \hat{\theta}_P, & i \in ab \end{cases}, \quad m_{i,P}^B = \begin{cases} \frac{N_B - \hat{N}_{ab}^{PML}(\hat{\theta}_P)}{\hat{N}_{ab}^{PML}(\hat{\theta}_P)}, & i \in b \\ \frac{\hat{N}_{ab}^{PML}(\hat{\theta}_P)}{\hat{\theta}_P \hat{N}_{ab}^A + (1-\hat{\theta}_P) \hat{N}_{ab}^B} (1 - \hat{\theta}_P), & i \in ab \end{cases}.$$

If the value of θ_P cannot be estimated, for example if the two sampling frames coincide or if domain b is structurally empty, then one can use an average design effect from each survey in the adjustment, as described in [14]. The PML estimator is thus approximately unbiased and internally consistent; it has high efficiency in many survey situations.

Single frame estimators. Bankier [1] and Kalton and Anderson [9] propose estimators of the form in (1) that treat all the observations as though they had been sampled from one frame, with adjusted weights in the intersection domain relying on the inclusion probabilities for each frame. The weight adjustments for the single frame estimator in [9] are:

$$m_{i,S}^A = \begin{cases} 1, & i \in a \\ w_i^B / (w_i^A + w_i^B), & i \in ab \end{cases}, \quad m_{i,S}^B = \begin{cases} 1, & i \in b \\ w_i^A / (w_i^A + w_i^B), & i \in ab \end{cases}.$$

If $w_i^A = 1/\pi_i^A$ and $w_i^B = 1/\pi_i^B$, the single frame estimator uses $\tilde{w}_{i,S}^A = \tilde{w}_{i,S}^B = 1/(\pi_i^A + \pi_i^B)$ for units in ab . The weight adjustment requires knowledge of the inclusion probability for a population unit for both samples. Thus if a stratified random sample with unequal sampling fractions is taken from from B , one needs to know the frame-B stratum membership for units sampled in $\mathcal{S}(A)$.

The single frame weight modifications depend only on the original weights for the two samples. The weight modifications thus are the same for all response variables, so estimators are internally consistent. For complex surveys, however, single frame estimators may not be as efficient as the optimal estimators. The adjusted weights from the single frame estimator, while not necessarily optimal for estimating a particular response, can be interpreted in terms of inclusion probabilities for sampled units. If the sampling fractions are small, $\tilde{w}_{i,S}^A$ is approximately $1/P(\text{unit } i \text{ is included in one of the samples})$. Like the other estimators, the single frame estimator is approximately unbiased for population totals.

Weight Adjustments with More than Two Frames. In the general case, suppose there are Q frames, denoted $A_1, \dots, A_Q$. Let $\mathcal{S}_q$ denote the probability sample from frame A_q , for $q = 1, \dots, Q$. Unit i in sample $\mathcal{S}_q$ has probability of inclusion $\pi_i^{A_q}$ and weight $w_i^{A_q}$. There are a total of D nonoverlapping domains.

A multiple frame estimator generalizing (1) is of the form

$$\hat{Y} = \sum_{q=1}^Q \sum_{i \in \mathcal{S}_q} m_i^{A_q} w_i^{A_q} y_i.$$

If $m_i^{(A_q)}$ does not depend on the data, the estimator $\hat{Y}$ is approximately unbiased when $\sum_{q=1}^Q \delta_i(A_q) m_i^{A_q} \approx 1$ for each i , where $\delta_i(A_q) = 1$ if unit i is in frame A_q and 0 otherwise. This clearly leaves many choices for the multiplicity weight adjustment $m_i^{A_q}$. A simple choice, which generalizes the fixed weight dual frame estimator $\hat{Y}(1/2)$ in (3),

$$m_i^{A_q} = \frac{1}{\sum_{q=1}^Q \delta_i(A_q)} = \frac{1}{\text{number of frames that contain unit } i} \quad (4)$$

as suggested in [16]. Other choices include setting $m_i^{A_q} = 1$ for one of the frames and 0 for all of the other frames.

Lohr and Rao [14] develop optimal and pseudo-maximum likelihood (PML) estimators for three or more frames. They derive maximum likelihood estimators of the vectors of domain sizes and domain means under simple random sampling, and then adjust the estimators so that they are design consistent. The PML estimator has been shown theoretically and through simulation studies to have small mean squared error and to work well in a wide variety of survey designs [17, 14].

Many of the properties from the dual frame situation extend to the case of three or more frames. How do the multiple frame estimators satisfy the criteria set out in Section 2? All of the estimators—fixed weight, optimal, PML, and single frame—

are approximately unbiased for population totals when sufficiently large samples are taken in the frames. The fixed weight, PML, and single frame estimators are internally consistent; the optimal Hartley-type and Fuller-Burmeister-type estimators in [14] are not internally consistent. While the optimal estimators are asymptotically efficient, they are often unstable in small or moderate samples with three or more frames: optimal estimated weight modifications are functions of large estimated covariance matrices, and can be outside the range $[0, 1]$. The empirical mean squared error of the optimal estimators can be far from the asymptotic mean squared error when covariances are estimated. The optimal estimators are also ill suited for use with standard survey software. Because of the instability and internal inconsistency of optimal estimators, we do not recommend their use in practice.

Depending on the survey circumstances, the fixed weight, PML, or single frame estimators should be used. The PML estimator has small mean squared error in many survey circumstances, and thus is a good choice for general surveys [14]. In some cases, though, the simplicity and transparency of a fixed weight estimator may be preferred. Fixed weight adjustments may make year-to-year domain comparisons easier in a survey where the domain proportions are relatively constant over time. A compromise may be reached by setting fixed weights approximately equal to the near-optimal weights from the PML method.

3 Nonsampling Errors

One major reason to use multiple frame surveys is that they often have better coverage properties than a single frame survey. When all frames are incomplete, as in Fig. 2, any one of frames A, B, or C, if used as the sole sampling frame, would have severe undercoverage. The multiple frame survey design ensures that all units in the overlapping frames have a positive probability of inclusion in the survey.

Like all surveys, multiple frame surveys are subject to nonsampling errors. They have nonresponse, which may differ in the different frames. While the union of the frames may have better coverage than a single frame, there may still be undercoverage of the target population. Estimators for multiple frame surveys are also sensitive to domain misclassification and biases that might result from different administration methods or modes in the component surveys.

3.1 Domain Misclassification

The estimators discussed in Section 2 construct weights for the observations based on domain membership. Thus in the estimator $\hat{Y}(\theta)$ in (3), the weight multiplier of an observation from sampling frame A is 1 if the observation is in domain a , and is θ if the observation is in domain ab , in order to account for the multiplicity of sampling.

In practice, domain membership may not be clear. It may be unknown whether a respondent in area frame A also belongs to a list frame B. If frame A is an area frame and frame B is an internet frame, for example, the only way to determine whether an individual sampled from frame A is also in frame B might be to ask the person about internet access, and the person might not give the correct response.

If matching or record linkage is used to determine frame membership, imperfect matches can also result in observations being misclassified. Lesser and Kalsbeek [11] discuss nonsampling errors in dual frame surveys that have been conducted by the U.S. National Agricultural Statistics Service. Domain misclassification can occur if a farm sampled in the area frame is incorrectly classified with respect to its list frame membership. In mobile/landline dual frame telephone surveys, it is challenging to determine whether a person in one frame is also in the other frame [10]. A person reached in a landline telephone sample may also have a mobile phone, but usually have the mobile phone turned off. While technically in the overlap domain, that person is virtually unreachable in the mobile phone survey.

Lohr and Rao [14] find in simulation studies that the PML estimator has smaller mean squared error than the other estimators when misclassification is present, but this is due largely to the smaller variance of that estimator. In dual frame surveys when a simple random sample is taken in each frame, for any given estimator it is possible to construct a population for which that estimator of the population total has the largest bias among the set of estimators considered. Thus, while in specific populations one estimator may be less sensitive to domain misclassification than another, one cannot conclude that one estimator is better in all situations.

If domain misclassification is severe, any method for modifying the survey weights to adjust for multiplicity can result in biased estimates. To see possible effects of domain misclassification, consider a simple model for fixed weight adjustments under which every observation in a domain has the same misclassification probabilities. Let $q_{d,e}^F = P(\text{observation from frame } F \text{ sample that is in domain } d \text{ is classified into domain } e)$. Then for a dual frame survey where $m_i^A \approx 1$ for i classified in domain a , $m_i^A = m_{ab}^A$ for i classified in ab , $m_i^B \approx 1$ for i classified in b , and $m_i^B = m_{ab}^B$ for i classified in ab , the approximate bias of $\hat{Y}$ for any of the estimators in Section 2 is:

$$\text{bias}(\hat{Y}) \approx [q_{ab,a}^A(1 - m_{ab}^A) + q_{ab,b}^B(1 - m_{ab}^B)]Y_{ab} - (1 - m_{ab}^A)q_{a,ab}^A Y_a - (1 - m_{ab}^B)q_{b,ab}^B Y_b. \quad (5)$$

If the misclassification probabilities are known or can be estimated from another source, then the bias in (5) can be corrected under this simple model by revising the weight adjustments to $\tilde{m}_{i,ab}^A = [m_{ab}^A - q_{ab,a}^A/q_{a,a}^A]/[q_{ab,ab}^A - q_{ab,a}^A q_{a,ab}^A/q_{a,a}^A]$ for i classified in ab and $\tilde{m}_{i,a}^A = [1 - q_{a,ab}^A \tilde{m}_{i,ab}^A]/q_{a,a}^A$ for i classified in a , with similar adjustments for frame B. If there is no misclassification, the weight adjustments are not revised.

If information about the q 's is known, some estimators may have less misclassification bias than others. Mecatti [16] suggests that if the weight modifications in (4) are used, it is not necessary to know *which* domain an observation belongs to; it is only necessary to know *how many* frames the observation is in. Thus, if the only

misclassification is among domains which belong to the same number of frames, the estimator in (4) will have no misclassification bias. In the dual frame setting, if it is known that $q_{a,ab}^A = q_{ab,a}^A = 0$, then the fixed weight estimator with $\theta = 0$ has no misclassification bias. Both of these accommodations for bias, however, can result in greater mean squared error than if the bias is ignored because the weighting can be very inefficient; if the domain means are similar, then the misclassification might not result in any appreciable loss of precision but discarding observations from domain ab in the frame-A sample would greatly increase the mean squared error.

In practice, we expect domain misclassification to be related to responses of interest; we also expect that in many situations, misclassification is more likely to occur in certain directions. In longitudinal dual frame surveys, domain misclassification can have greater effects than in cross-sectional surveys [15]. Domain misclassification thus needs to be treated analogously to, and in conjunction with, nonresponse. In some situations, the domain indicator can be missing or unavailable. Clark et al. [5] investigate logistic regression and record-linkage methods for predicting the domain of an observation with missing domain information.

3.2 Mode Effects

In some cases, dual frame may also mean dual mode or dual language. De Leeuw [6] compares advantages and disadvantages of different sampling modes and summarizes empirical research on mode biases. Mode effects that occur in single frame surveys will also occur in multiple frame surveys. If different modes are used in different frames, it is challenging to separate mode effects from other nonsampling errors.

Many of the multiple frame survey estimators depend on combining estimates from the overlap domains, and these assume that the estimator of Y_{ab} from the component surveys are both estimating the same quantity. If, however, the frame A survey is conducted in person while the frame B survey is conducted by telephone, it is possible that a census of the domain ab from frame B would give a different total in the domain than a census from frame A. We recommend testing the equivalence of domain estimators from different samples to ascertain if nonsampling errors are a problem.

One possibility to investigate mode effects is to conduct the frame B survey using a split sample, e.g. partly in-person and partly by telephone, but that reduces the cost savings from the dual frames. Careful pretesting can mitigate the mode effects. Research is needed in this area; the same problem of mode effects, of course, occurs in single frame surveys such as the American Community Survey in which nonresponse follow-up is done by different mode than the original sample [4].

3.3 Nonresponse

In any survey, nonresponse can result in biased estimates of population totals and other quantities. Different nonresponse rates in the samples from the two frames can affect the point estimates of the population total given in Section 2; additionally, nonresponse can affect the weight adjustments prescribed by some of the methods.

Kennedy [10] discusses a problem that has occurred when frame A consists of landline phones and frame B is mobile phones: the units in the intersection domain ab who were interviewed by mobile phone differed from those in ab who were interviewed on the landline phone. For example, it was estimated that 18% of the intersection units were aged 18–25 in the frame-B sample, while it was estimated that only 8% of the intersection units were aged 18–25 using the frame-A sample. The difference is ascribed to nonresponse: it is thought that persons who predominantly use cell phones (and thus are difficult to reach through a landline survey) tend to be younger.

Brick et al. [2] propose a method for choosing θ in fixed weight estimators to reduce bias due to nonresponse when the response rates in the different domains and frames are known. They also poststratify to independent estimates of population counts in the different domains.

4 Design Issues

As discussed in Section 1, multiple frame designs can give better coverage and precision than a single frame survey with equivalent cost. The survey needs to be carefully designed to realize those cost savings.

We have seen that misclassification can create serious biases regardless of estimator choice. Thus, the survey designer must decide whether the additional coverage made possible by including more frames outweighs the potential for bias from misclassifying sampled units or multiple modes. In a survey with Q frames, there are a total of $2^Q - 1$ possible domains. As the number of frames increases, so does the potential for domain misclassification.

In general, the optimal design is a function of sampling variances and nonsampling errors in each frame, as well as of the estimator chosen. In a multiple frame survey, effects of domain misclassification, nonresponse, and mode bias are often confounded. A designed experiment may help disentangle these effects. We are currently studying the relation among these three types of nonsampling errors, and researching the efficiency of different design and estimator combinations in the presence of the nonsampling errors discussed in Section 3.

Acknowledgements This research was partially supported by the U.S. National Science Foundation under grants SES-0604373 and DRL-0909630. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Bankier, M.D.: Estimators based on several stratified samples with applications to multiple frame surveys. *J. Am. Stat. Assoc.*, **81**, 1074–1079 (1986)
2. Brick, J.M., Cervantes, I.F., Lee, S., Norman, G.: Nonsampling errors in dual frame telephone surveys. *Surv. Method.* (in press)
3. Brick, J.M., Dipko, S., Presser, S., Tucker, C., Yuan, Y.: Nonresponse bias in a dual frame survey of cell and landline numbers. *Pub. Opin. Q.*, **70**, 780–793 (2006)
4. Citro, C.F., Kalton, G.: Using the American Community Survey: Benefits and Challenges. National Academies Press, Washington DC (2007)
5. Clark, J., Winglee, M., Liu, B. Handling imperfect overlap determination in a dual-frame survey. *ASA Proc. Surv. Res. Meth. Sec.*, 3233–3238 (2007)
6. de Leeuw, E.D.: Choosing the method of data collection. In: de Leeuw, E., Hox, J., and Dillman, D. (eds.), *International Handbook of Survey Methodology*, pp. 113–135. Lawrence Erlbaum, New York (2008)
7. Fuller, W.A., Burmeister, L.F.: Estimators for samples selected from two overlapping frames. *ASA Proc. Soc. Stat. Sec.*, 245–249 (1972)
8. Hartley, H.O.: Multiple frame surveys. *ASA Proc. Soc. Stat. Sec.*, 203–206 (1962)
9. Kalton, G., Anderson, D.W.: Sampling rare populations. *J. Royal Stat. Soc. A*, **149**, 65–82 (1986).
10. Kennedy, C.: Evaluating the effects of screening for telephone service in dual frame RDD surveys. *Pub. Opin. Q.*, **71**, 750–771 (2007)
11. Lesser, V.M., Kalsbeek, W.D.: Nonsampling errors in environmental surveys. *J. Ag. Biol. Env. Stat.*, **4**, 473–488 (1999)
12. Lohr, S.: Multiple frame surveys. In: Rao, C.R., Pfeffermann, D. (eds.) *Handbook of Statistics*, Vol. 29A, *Sample Surveys: Design, Methods and Applications*, pp. 71–88. North Holland, Amsterdam (2009)
13. Lohr, S.L., Rao, J.N.K.: Inference in dual frame surveys. *J. Am. Stat. Assoc.* **95**, 271–280 (2000)
14. Lohr, S.L., Rao, J.N.K.: Estimation in multiple-frame surveys. *J. Am. Stat. Assoc.* **101**, 1019–1030 (2006)
15. Lu, Y., Lohr, S.L.: Gross flow estimation in dual frame surveys. *Surv. Method.* (in press)
16. Mecatti, F.: A single frame multiplicity estimator for multiple frame surveys. *Surv. Method.* **33**, 151–157 (2007)
17. Skinner, C.J., Rao, J.N.K.: Estimation in dual frame surveys with complex designs. *J. Am. Stat. Assoc.* **91**, 349–356 (1996)
18. Tucker, C., Brick, J.M., Meekins, B.: Household telephone service and usage patterns in the United States in 2004: Implications for telephone samples. *Pub. Opinion Q.*, **71**, 3–22, (2007)