Venn Diagrams, Probability 101 and Sampling Weights Computed for Dual Frame Telephone RDD Designs

Trent D. Buskirk, Ph.D.[1] and Jonathan Best[2]
[1]The Nielsen Company,
8084 Watson Road Ste 226, Saint Louis, MO 63119
[2]Princeton Survey Research Associates International,
600 Alexander Road, Suite 2-2
Princeton, NJ 08540

Abstract
The continued rise in cell phone penetration creates a real potential for undercoverage bias in many RDD sample surveys. To respond to such threats researchers have begun implementing dual frame RDD sampling strategies. In this paper we present a method for constructing first-stage sampling weights derived under an overlapping, dual frame design (e.g. cell and landline RDD numbers) based on probability 101 fundamentals. Because these two frames potentially overlap at the user level, selection probabilities must be adjusted for multiplicity of selection. Our method resembles weighting strategies consistent with a "single frame" approach and does not require estimation of a compositing factor traditionally used in the "dual frame" approach. Estimators employing resulting sampling weights are effectively Horvitz-Thompson estimators that in some cases can be approximated using a slightly simplified Hansen-Hurwitz type estimator. We use our proposed method to construct base sampling weights for both a national and state level dual frame RDD samples of landline and cell phone numbers. Using national and state-level benchmark data we also present bias estimates for a battery of health related outcomes.

Key Words: Dual Frame Sampling, Cell Phone Sampling, RDD health surveys, Single Frame Estimators, Sampling Weights

1. Background

The era of RDD sampling of only landline telephones for health surveys has nearly come to an end as we know it. In fact, the most recent American Association of Public Research Cell Phone Task Force (2010) recommends that researchers interested in using survey samples now augment regular RDD landline samples with samples taken from the cell phone frame. This recommendation derives in part because of the rapidly changing telephone landscape within the United States – with approximately 30% of U.S. households being cell phone only (CPO) (Blumberg and Luke 2011). The recommendation also hinges upon the fact that the distribution of cell phone only status varies substantially across various levels of demographic variables such as age, race, sex, household structure, poverty status, geographic region and household ownership status. In particular, Blumberg and Luke (2009) estimate that approximately half of renters are CPO compared to less than 20% among homeowners or purchasers. Similarly, roughly 53.5% of adults aged 25-29 reside in wireless only households compared to less than 20% among those ages 45-64 years old. The rise in the prevalence of adults who live in CPO households has also been linked with the potential for noncoverage bias for health risk related outcomes including HIV testing, binge drinking and financial barriers to medical care for surveys that exclude cell phone only adults aged 18-29 (Blumberg and Luke, 2009). These trends continue among all adults when considering outcomes such as health insurance with nearly 29% of wireless only adults under 65 years of age being without health insurance compared to 13% among comparable aged adults living in landline households. In summary, ignoring the cell phone only households for health related surveys related to health behavior and risk factors across the age distribution threatens the overall validity of inferences drawn from RDD samples of landline households in ways that appear to be no longer negligible or ignorable.

To respond to threats imposed by noncoverage bias, health survey researchers have begun implementing and exploring dual frame RDD sampling strategies in order to cover cell phone households. In particular, at the national level, the BRFSS has begun experimenting and testing the use of dual frame RDD sampling designs to incorporate both landline and cell phones into their samples (Hu et al, 2011) and at the state level the California Health Interview Survey has also explored the use of dual frame sampling with cell phone screeners (Brick, Edwards and Lee, 2007). While there is no clear best practice yet, the two types of dual frame designs that appear to have had the most traction in practice according to the AAPOR task force include: Dual Frame Sampling with Screening Designs, in which numbers identified as cell only from the cell phone frame are screened in, and Dual Frame Sampling without Screening in which numbers are selected and called from each of the two frames without any phone service screening or exclusions. Two other approaches that have also been reported are RDD sampling from only the Cell Phone Frame and Address Based Sampling Designs. We note that issues surrounding the selection, fielding and other survey protocols for cell phone sampling designs continue to emerge in the research literature (see Lavrakas et al., 2007 and AAPOR, 2010) for a discussion of some of the more salient aspects related to cell phone sampling and surveying. Specific studies comparing feasibility and implementation of cell phone sampling to landline sampling have also been reported in the literature including Zuwalcak, 2009, Brick, Brick, et al. 2007, Steeh, Buskirk and Callegaro, 2007 and Kennedy, 2011. We also note that for sampling plans related to cell phones, the use of alternate disposition codes for computation of response rates, for examples, has also been proposed (Callegaro, Steeh, Buskirk, et al., 2007).

1.1 More Details about Dual Frame Sampling Designs

Dual frame sampling designs involve sampling from two different frames that when put together typically increase the coverage of the target population. The two frames can be disjoint or could have overlap depending on the design or frames themselves. In the case of telephone sampling for health surveys a dual frame approach would sample from a landline telephone number frame as well as a cell phone number frame. Members of the household/adult population that have both a cell phone and a landline phone (i.e. so called "dual users") would be considered part of the overlap as depicted in Figure 1. A general overview of dual frame surveys with and example related to RDD surveys is given in Buskirk (2008). Theoretical aspects of dual frame designs and inference have been provided by Lohr and Rao (2000) for dual frame sampling designs and by Hartley (1962) and Lohr and Rao (2006) for sampling designs involving multiple frames. Applications of weighting methods for multiple frame health surveys have been described by Metcalf and Scott (2009). While the dual frame approach has not yet become ubiquitous among RDD surveys, its use (at various levels) continues to increase. Hu and colleagues (2011) report recent expansions of the BRFSS RDD landline data collection to include cell phone only households via screening samples selected from the cell phone frame. Battaglia et al. (2010) describe a city-wide dual frame with screening design used for the New York Community Health Survey. A dual frame RDD design has also been described for use with the National Immunization Survey (Srinath et al., 2004).

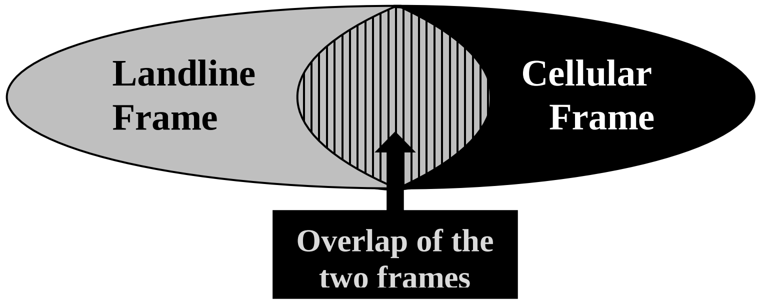


Figure 1: Venn Diagram illustrating the dual-frame nature of current telephone samples: landline and cellular number frames and their overlap. Note- technically the frames only overlap for ported numbers in the "sampling population," rather the overlap occurs in the associated target population (i.e. people who have been identified by their cell/landline phone number).

1.2 Survey Weighting for Dual Frame RDD Surveys

Weighting methods appropriate for landline RDD survey sampling need to be modified for dual frame designs that incorporate sampling from a cell phone frame. According to the most recent AAPOR Cell Phone Task Force (2010) there is no consensus among the survey research community regarding a best practice for weighting dual frame RDD samples. However, there is general agreement that the computation of final weights would include several steps including: (a) computation of a base weight consistent with the sampling design within each frame; (b) nonresponse adjustments applied within each frame and (c) post-stratification adjustments to known totals. As with any survey, post-stratification of the base weights to population based control totals (e.g. demographic variables, phone usage, etc.) is generally recommended not only for nonresponse, but for external validity of the estimates that would be derived from the study. The sampling population (i.e. two frames of telephone numbers) and the target population (i.e. adults in the U.S.) are not necessarily the same in RDD telephone surveys since many of telephone numbers are not associated adults in the U.S., for example. Post-stratification to control totals that are pertinent for the target population then provide calibration from the sampled population to the target population.

Weighting approaches for dual frame RDD surveys that have appeared in the recent literature have offered methods for base weight construction and post-stratification adjustment. In particular, much of the work regarding base weight construction for dual frame RDD surveys stems from the work presented by Hartley 1962. Brick and colleagues (2007) have eloquently presented an exposition of the standard "composite" estimator applied to dual frame RDD surveys (i.e. $y_{comp} = y_a + y_b + y_\lambda$) where y_a is the weighted (including nonresponse adjustments) estimate from those respondents who are landline only, y_b is the weighted (including nonresponse adjustments) estimate of the population total of the outcome of interest derived from cell phone only respondents. Lastly, $y_\lambda = \lambda y'_{ab} + (1-\lambda) y''_{ab}$ where y'_{ab} is the nonresponse-adjusted weighted estimate of the population total from dual users selected from the landline frame and y''_{ab} is the nonresponse-adjusted weighted estimate of the population total using dual users selected from the cell phone frame. The λ value is called the compositing factor. Current recommendations are to use $\lambda = .5$ (simple compositing) unless an optimal value is desired based on cost/variance of outcome across the two frames. Another method suggests setting the composite factor in direct proportion to the effective sample size of dual users from each of the two frames relative to the sum total of dual users (effective size) from both frames known as the Hartley class of estimators as explored in (Frankel et al. (2007) and Metcalf and Scott (2009). A third method that has been mentioned in the AAPOR task force report involves computing the composite factor as a function of dual user response rates from each of the two samples as well as estimates of land and cell mainly usage status among dual users derived from external sources such as the National Health Interview Survey (early release program: http://www.cdc.gov/nchs/nhis.htm) (Brick, 2011). The main difficulty with implementing this method stems from the need to estimate usage status from external totals that may not be available at the state or local level with acceptable precision. The effective sample size method provides a more sophisticated and possibly more efficient method for computing the base weights beyond simple compositing and may be a viable alternative to the optimal allocation method if variances across the two frames are unknown or may vary from one survey outcome to another (thus creating the need for a new compositing factor for each particular outcome). We note that the estimates of effective sample size will be influenced by the degree of nonresponse adjustments within each of the two samples and the degree to which one can accomplish nonresponse adjustments within each sample may vary due to differences in the amount and nature of auxiliary information that is available from each frame. Moreover, various nonresponse adjustment techniques may impact the variability of the resulting weights which in turn will impact effective sample sizes. It is not yet known how sensitive the effective sample size method is to various types of nonresponse adjustments performed within each frame prior to computing the compositing factor.

Post-stratification methods that use demographic variables and phone service data for (simple) composite base weights has been discussed (Brick et al. 2006) and examined within and across the two frames. Extending the post-stratification to include additional phone usage variables (i.e. cell mostly, landline mostly, etc.) has also been examined (Kennedy, 2011). For national level data, control totals for phone variables are available from the NHIS and for states and larger MSAs (Blumberg, Luke and Davidson, et al. 2009) – however, these estimates are based on small area statistical models and their precision may not be acceptable to deem them control totals for some applications. Currently, we are unaware of a best practice or theoretically acceptable/optimal method for incorporating this error into the weighting adjustment process. The limited precision of some state-level estimates has motivated other work on how to generate local level control totals that are based on linking local survey data to national data via statistical models (Battaglia 2010). An alternative method based on using differing response rate ratios across cell phone only, landline only and the overlap domains of a dual frame RDD has also been proposed (Gutterbock, 2009).

Another less widely applied method for computing base weights for the dual frame RDD design is the so called “single frame” approach using fixed weights to ensure design unbiasedness (Mecatti, 2007). This estimator applies a base weight separetly computed for each of the three domains (i.e. cell only, dual users, landline only) depicted in Figure 1. The single-frame family of estimators are appealing in that (1) they do not require the computation of a compositing factor and (2) they are fairly stratighforward to compute.

For example, assuming simple random sampling from each frame the base weight for numbers selected from the cell frame who are cell only using the single frame approach would be $f_{Cell}^{-1} = N_{Cell} / n_{Cell}$ and likewise, $f^{-1}{}_{Land} = N_{Land} / n_{Land}$ for landline only numbers selected from the landline frame . Sampled numbers tracked to dual users, *regardless of the frame from which they were sampled initially*, receive the base weight $\left(f_{Cell} + f_{Land} \right)^{-1}$ (Mecatti, 2007 and Metcalf and Scott, 2009). Using this approach yields the Hansen-Hurwitz type estimator as described for dual frame surveys by Metcalf and Scott (2009). These single frame estimators ignore the potential of a sampled number to come from both frames.

In this paper we will compute a single frame estimator that is based directly on first principles of probability – namely the addition rule of probability: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. We refer to weights derived using this approach basic probability method weights (or SF-BP weights, for simplicity). The intersection or overlap component (OLC) comes from considering the overlap of the two frames and estimators using these weights are akin to Horvitz-Thompson. Estimators based on weights derived by ignoring the overlap component (i.e. No OLC) as described by Metcatti (2007) and Metcalf and Scott (2009) are essentially Hansen-Hurwitz type estimators and design effects can be derived for the later by incorporating the frame as a “super stratum” in the analysis. In this paper we will focus on deriving estimates based on basic probability weights that do and do not ignore the overlap component. The single frame basic probability method (SF-BP) was first described by Best (2010) for the dual frame RDD scenario and is in part based on earlier discussions relating to various forms of single frame estimators (Bankier, 1986). We believe this method will be appealing because it does not rely on a possibly subjective choice of estimation for the compositing factor, nor does it require external benchmarks for computing the base weights. Furthermore, this method (and other single frame estimation methods) does not require separating (and processing) the “dual users” explicitly into their respective frames and combining respective estimates by means of the compositing factor. Moreover, application of post-stratification methods to SF-BP base weights proceeds as if a single sampling frame was used for the design. In the next section we will provide the SF-BP method base weight computation formula using terms explicitly relevant for RDD dual frame surveys. We formalize the base weight construction in the next section and describe both national and state-level dual frame surveys to which the SF-BP weighting method will be applied. We will also discuss extensions of this method to stratified designs. In the third section we apply the SF-BP method to a national dual frame RDD survey and to a state-wide stratified dual frame survey. We will conclude the paper with a discussion of the strengths and limitations of this method.

2. Materials and Methods

In this section we describe the SF-BP method for computing dual frame RDD base weights and then apply this methodology to two dual frame RDD surveys. Final weights were derived from these base weights by applying a post-stratification adjustment via raking (i.e. iterative proportional fitting) these base weights to demographic and phone service (i.e. cell phone only, landline only, dual user) population totals where appropriate using a SAS 9.2 Sample Balancing Macro (Izrael et al., 2009). Nonresponse adjustments were computed for the state survey prior to computing the post-stratification adjustments. The base and final weights are summarized by their implied design effects approximated by 1 plus the squared coefficient of variation of the respective weights. We also estimated bias and relative bias for several survey outcome estimates derived using the final post-stratified SF-BP weights using benchmark totals from the Current Population Survey and the Behavioral Risk Factor Surveillance Survey. For the state specific example, state-level BRFSS data were used to compute appropriate benchmarks.

2.1 Data Sources

2.1.1 National Data Source

The national level survey was a weekly dual frame RDD omnibus survey (Omni) conducted in July 2011 by Princeton Survey Research Associates International (PSRAI). More specifically, the PSRAI July 2011 Omnibus Week 1 survey obtained telephone interviews with a nationally representative sample of 1,109 adults living in the continental United States. Telephone interviews were conducted by landline (776) and cell phone (333, including 126 without a landline phone) and all interviews were conducted in English from July 7-10, 2011. A combination of landline and cellular random digit dial (RDD) samples was used to represent all adults in the continental United States who have access to either a landline or cellular telephone. Both samples were provided by Survey Sampling International, LLC (SSI). As many as three attempts were made to contact every sampled telephone number. For the cellular sample, interviews were conducted with the person who answered the phone. Interviewers verified that the person was an adult and in a safe place before administering the survey. The response rate was 9.3 percent for the landline sample and 8.3 percent for the cell sample (AAPOR RR3). Base weights were post-stratified to balance sample demographics to national population parameters for sex, age, education, race, Hispanic origin, region (U.S. Census definitions), number of adults in household, and telephone usage.

2.1.2 State Level Data Source

The Missouri State Level Health Literacy Survey (HL-MO) sponsored by the Missouri Foundation for Health served as the state level data source. Specifically, the HL-MO was a dual frame RDD survey of Missouri residents 18 years and older conducted between March and June of 2010. The sampling design used strata based on Health Foundation Regions (i.e. collectations of contiguous counties) and an “At risk Low Literacy Zone” (i.e. a collection of 4 counties whose estimated literacy levels were “below average”). Within each of the 12 strata a dual frame RDD sample was taken using both landline and cell phone frames with county/fips code used as the defining point for identifying area code and prefix combinations for the sample selection per stratum. The overall survey had 3358 respondents and realized an overall response rate of 20% (AAPOR 3). Using response rates that were aggregated across each of the two frames, the SF-BP base weights were directly adjusted by the reciprocal of the *stratum* AAPOR 3 response rate. Furthermore, the nonresponse adjusted base weights were further calibrated to match the marginal state totals for 6 control variables known to be associated with health literacy levels including: age, sex, race, education, urbanicity and county. We note that cell phone service (i.e. landline only, cell only and dual users) was not reliably avaialble for the state of MO at the time of publication so we did not use these variables in the post-stratification adjustment. An iterative proportional fitting and trimming algorithm was employed to perform the calibration and for the purposes of this illustration no weight trimming was performed (Izrael et al., 2009). More complete details about the sampling design and the weigting procedures are described elsewhere (Buskirk, 2010).

2.1.2 External Data Sources

Distributions for the control totals used for post-stratification adjustments for the OMNI survey came from a special analysis of the Census Bureau's 2010 Annual Social and Economic Supplement (ASEC) that included all households in the continental United States. The telephone usage parameter came from an analysis of the July-December 2010 NHIS. Similarly, control total values used in the post-stratification adjustments for the HL-MO survey came from the 2000 Decennial Census, the 2009 American Community Survey (ACS) and the 2006-2008 3-year aggregated ACS.

We compared estimates derived from the OMNI and the HL-MO to comparable estimates derived from several external, nationally based surveys including the National Behavioral Risk Factor Surveillance Survey (BRFSS) (2010) and the Missouri State Behavioral Risk Factor Surveillance Survey (BRFSS-MO, 2010), the 2009 National Household Survey on Drug Use and Health (NHSDUH), and the 2010 National Health Interview Survey (NHIS).

2.2 The SF-BP Dual Frame Weighting Approach

The SF-BP sampling weights are constructed using two steps including: base weight computation and nonresponse/noncoverage adjustments. The second step could be accomplished as a single step or could consist of two parts including a nonresponse adjustment first, followed by an overall post-stratification adjustment for under/non-coverage. We describe both of these steps in detail below.

2.2.1 SF-BP Weight Construction, Step 1: Base Weight Computation

Viewing figure 1 as a Venn diagram and considering the event of being included in the final sample (InS) as the union of two events – being included in the sample from the landline frame as the first event (InS_{LL}) and being included in the sample from the cell phone frame as the second (InS_{CP}) we see that from the basic axiom of probability, the probability of being in the overall sample for any given phone number is:

$$P(InS) = P(InS_{LL} \cup InS_{CP}) = P(InS_{LL}) + P(InS_{CP}) - P(InS_{LL} \cap InS_{CP}) \quad (1)$$

Assuming that sampling from the two frames is done independently we quickly see that the probability of being included in a dual frame RDD as depicted in Figure 1 becomes:

$$P(InS) = P(InS_{LL}) + P(InS_{CP}) - P(InS_{LL}) \times P(InS_{CP}) \quad (2)$$

The BP-method applies the equation in (2) directly to compute the base-weights noting that the inclusion probabilities from within each of the frames is dependent upon several factors including: (a) frame size (i.e. population size) (b) sample size within each frame, (c) within household selections of adults sampled from the landline frame (d) the multiplicity of phones connected to an adult/household (in both frames). We note that the SF-BP method could be applied either at the household level or at the person level. Household selection within the cell phone frame would depend on assumptions related to the degree to which the cell phone is shared within the household and within adult selection may require additional efforts to obtain completed information due to the personal/portable nature of cell phones. Person level applications require knowledge of the person selection methods that are used to select a respondent within the household. In this paper we will focus on the person level application rather than household level applications for this weighting method. Furthermore, we assume that a simple selection of a single adult occurs randomly from among all eligible adults within the house for households contacted in the landline frame; moreover, we will assume that there is no additional adult selection that occurs in the cell phone frame. Finally, we assume that samples are taken independently from each frame.

There are four variables that determine the probability that an adult is sampled from the landline sample frame: the size of the landline sample; the size of the landline sample frame; the number of landline telephones used to receive calls in the household and the number of adults in the household. The three variables that determine the probability that someone is sampled by cell phone are the size of the cell sample, the size of the cell sample frame and whether or not the adult has a cell phone. The notation for each of these quantities is provided for reference in Table 1.

Table 1: Notation/variables for computation of base sampling weights using the BPM in the context of an RDD dual frame sampling design.

Landline Frame	Cell Frame
$\mathbf{U_{LL}}$= the size of the landline sample frame	$\mathbf{U_{CP}}$ = the size of the cell sample frame
$\mathbf{S_{LL}}$ = the amount of landline sample released	$\mathbf{S_{CP}}$= the amount of cell sampled released
LL= the number of landline telephones in the household that are used to receive calls	**CP**= The number of cell phones owned by the respondent (or more simply whether or not the adult has a cell phone).
AD= the number of adults in the household (assumed to be at least 1)	

Substituting the variables listed in Table 1 into the formula (2) we see that the SF-BP base weight (considering the overlap of the two frames (OLC) is computed as:

$$\mathrm{bw}_{\{\mathrm{SF-BP\ (OLC)}\}} = \big(\mathrm{P}(\mathrm{in\ S})\big)^{-1} = \left(\left(\frac{\mathrm{S_{LL}}}{\mathrm{U_{LL}}} \times \frac{\mathrm{LL}}{\mathrm{AD}}\right) + \left(\frac{\mathrm{S_{CP}}}{\mathrm{U_{CP}}} \times \mathrm{CP}\right) - \left[\left(\frac{\mathrm{S_{LL}}}{\mathrm{U_{LL}}} \times \left(\frac{\mathrm{LL}}{\mathrm{AD}}\right) \times \frac{\mathrm{S_{CP}}}{\mathrm{U_{CP}}} \times \mathrm{CP}\right)\right]\right)^{-1} \quad (3)$$

We can also compute base weights that essentially ignore the overlap in the two sampling frames by dropping the subtraction terms in formulae (1) and (2). Modifying formula (3), the SF-BP base weights formed by ignoring the overlap (NO OLC) become:

$$\mathrm{bw}_{\{\mathrm{SF-BP\ (NO\ OLC)}\}} = \big(\mathrm{P}(\mathrm{in\ S})\big)^{-1} \approx \left(\left(\frac{\mathrm{S_{LL}}}{\mathrm{U_{LL}}} \times \frac{\mathrm{LL}}{\mathrm{AD}}\right) + \left(\frac{\mathrm{S_{CP}}}{\mathrm{U_{CP}}} \times \mathrm{CP}\right)\right)^{-1} \quad (4)$$

Because the inclusion probabilities are larger when the frame overlap is ignored relative to when it isn't, the SF-BP NO OLC base weights will generally be smaller than the SF-BP OLC base weights. For the sake of streamlining computations the number of CP is generally capped at 3, the number of adults in the households at 4 and the number of landlines at 2. We note that in our experience that the percentage of the sample effected by these caps is generally less than 4-5% as was the case in both our National and State examples.

2.2.2 SF-BP Weight Construction, Step 2: Nonresponse/Poststratification Adjustments using SF-BP base weights

Depending on the particular dual frame RDD survey the base weight would not generally be the final weight. In particular, depending on the level and type of survey being deployed, various post-stratification and nonresponse adjustments may be used to modify the base weights to account for undercoverage, nonresponse or a combination of the two (see Kalton and Flores-Cervantes, 2003, for example). Whatever method of adjustment is used will rely on the base weights computed as we have described. In the examples that follow we will describe how the base weights were modified using nonresponse and post-stratification adjustments at the State and National levels. Because the extent of adjustments will vary from one survey to another (e.g. some localities may not have benchmark data to use for post-stratification adjustment) and because the final unequal weighting effect and design effects are in part a function of these adjustments (Liu et al., 2002) we will report summary statistics for both the raw base weights along with the final post-adjusted base weights (computed without any trimming or truncation).

3. Results

3.1 Evaluating Sampling Weights Derived Using the SF-BP Method

We note that for both the National and State examples, stratification was employed to either increase coverage of underrepresented groups (first example) or to improve precision of estimates based on auxiliary data known to be related to key study outcomes (second example), so the design effects are likely to be greater than one – regardless of the weighting technique applied to account for the dual frame nature of the sampling design. A non-adjusted design effect was also approximated by the unequal weighting effect (Liu et al., 2002) defined as: $UWE = 1 + \left(\frac{n-1}{n}\right)CV^2$ where the CV is the coefficient of variation for the sampling weights. Absolute Relative Biases (ARB) were also computed using survey estimates based upon the SF-BP weights and various national/state estimates derived using reference data sources assumed to have minimal error (e.g. served as parameter values in evaluating the relative biases). Where multiple data sources were available for a given variable of interest, the ARB was computed as the average of the respective ARB computed from each external source. The design effects for the Hansen-Hurwitz estimators (based on SF-BP weights that ignore the overlap in the two frames) were also computed using SUDAAN version 10.2.

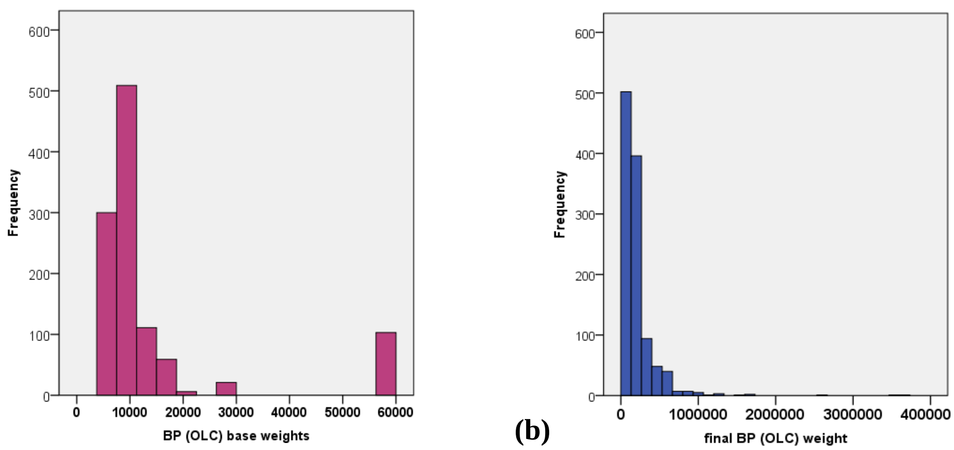
3.1.1. SF-BP weights and estimates derived from the OMNI National Survey

Based upon the 1109 responses obtained from the OMNI Survey, SF-BP base weights with and without the overlap of frames considered were computed and basic summary statistics are reported in Table 2. We note that the OLC baseweights were just slightly larger than those computed without the overlap and many were identical up to three decimal places in large part due to the small sampling fractions in both the landline and cell phone number frames. The unequal weighting effects were identical up to four decimal places for the base weights with and without the overlap and the unequal weighting effect for the final weights with and without the overlap were identical up to five decimal places. We note that the differences in the final weights with and without the overlap varied from -23 to just under 14 compared to 0 to .25 for the base weights indicating some effects due to the post-stratification adjustment in the opposite direction than was expected with the base weights. The overall distribution of both the base and final weights (OLC) are given in Figure 2. We note that approximately 9.3% of the 1109 respondents for the OMNI survey were single CPO householders with only one cell phone. Thus, their base weights depended only on the size of the cell phone sample relative to the size of the cell phone frame which produced the large base weights to the right of histogram (a) in Figure 2.

Table 2: Some summary statistics pertaining to the four versions of sampling weights computed for the Omni Survey: BP base weights with and without overlap component (OLC) and the final, post-stratified adjusted weights based on base weights with and without the overlap component.

OMNI Study Sampling Weights	Mean	Standard Deviation	Median	IQR	Minimum	Maximum	Unequal Weighting Effect (UWE)
BP Base Weights (OLC)	13847.605	15228.949	9048.057	7281.487	4205.061	59619.460	2.208367
BP Base Weights (NO OLC)	13847.502	15228.979	9047.928	7281.320	4204.894	59619.460	2.208389
Final Weights (OLC)	205395.859	245495.840	141464.298	132949.198	26166.734	3655353.753	2.427293
Final Weights (NO OLC)	205395.859	245495.893	141463.089	132949.137	26166.404	3655340.222	2.427294
BP Base Weight Difference [$]	0.103	0.068	0.129	0.053	0.000	0.250	N/A
BP Final Weight Difference [$]	0.000	1.279	-0.011	0.265	-23.038	13.530	N/A

[$] Weight (OLC) – Weight (NO OLC)



(a) (b)

Figure 2: Distributions of OMNI base weights considering the overlap in sampling frames (a) and final, post-stratified adjusted base weights (b).

Turning attention to the key outcomes, the OMNI Survey measured lifetime cigarette use for adults 18 and older and 20 and older along with employment status, parental/guardian status and the age and number of adults in the home. Estimates derived using SF-BP final weights with and without the overlap are presented in Table 3 along with corresponding estimates obtained from comparable questions from the 2010 BRFSS, 2010 HNIS and the 2009 NHSDUH. We note that the 2009 NHSDUH data set was the most recent publically available data set and differences between the other estimates and the one derived using the NHSDUH data may be due to time differences. Generally speaking, as expected, the estimates between OLC final weights and NO OLC final weights were minimal and in most cases presented only in the fourth or larger decimal place. As a result, the ARB estimates using SF-BP OLC and NO OLC final weights were identical up to four decimal places and only the ARB based on Horvitz-Thompson estimators which used the SF-BP OLC weights are given in Table 3. In most cases, larger ARB estimates were driven in large part from discrepancies in the NSDUH parameter estimates relative to those obtained from the other national surveys. While some of these ARB values are possibly unacceptably large even though no weight truncation was performed, factors associated with nonresponse in the OMNI survey may go beyond the simple factors used in the post-stratification adjustments. We note that there was no separate adjustment made for nonresponse in this study compared to what might have been done in the national studies, for example.

Table 3: Percentage and means for selected Omni survey variables computed using both Horvitz-Thompson estimates (i.e. SF-BP final weights with OLC) and Hansen-Hurwitz estimates (i.e. BP final weights with NO OLC). Also provided are comparable estimates from larger National U.S. Government surveys along with estimates of the Absolute Relative Biases.

Variable	Omni Horvitz-Thompson Estimate (OLC)	Omni Hansen-Hurwitz Estimate (No OLC) and 95% CI	BRFSS (2010)	NHIS (2010)	NSDUH (2009)	Average Absolute Relative Bias (Horvitz-Thompson)
Smoked >= 100 Cigarettes (Lifetime)	48.69376%	48.69375% (44.07, 53.32)	42.2150%	41.0576%	55.8770%	14.5797%
Smoked >= 100 Cigarettes (Lifetime) (Age 20 years or older) #	49.43308%	49.43307% (44.98, 53.88)	43.0854%	41.9241%	56.6700%	14.2237%
Currently Employed	57.02364%	57.02363% (52.55, 61.50)	57.0547%	59.5985%	60.4950%	3.5525%
Parent /Guardian of Minor ##	36.286309%	36.286306% (31.31, 41.26)	42.8339%	-NC-	38.8367%	12.5364%
Average Age of Adults	46.40508	46.40508 (44.59, 48.22)	47.3235	46.2333	-NC-	1.1747%
Average Number of Adults (Age 18 years or older) Living in Household	2.208205	2.208206 (2.11, 2.30)	2.3138	-NC-	2.303	4.5374%

Lifetime Cigarette Use (100 or more) was only available for Adults Aged 20 or older in the NHANES public release data file

For some National Surveys this estimate was derived from questions relating to having related children living within the home

-NC- indicates the estimate was either not available or not readily computable/comparable from the Public Release Survey Data Files

3.1.2. SF-BP weights and estimates derived from the HL-MO State Survey

Using the 3358 respondents from the HL-MO survey we computed estimates for the proportion of MO adults who were ever told by a physician that they had diabetes, the proportion of MO adults who have some type of health coverage based on the SF-BP base weights (with and without the overlap considered) that have been adjusted for nonresponse and undercoverage. ARB estimates were computed and based upon estimates derived from using the 2010 BRFSS data from the state of MO. We note that the 2010 BRFSS did not incorporate cell phones into the design so our estimate of relative bias may be impacted by differences in diabetes status and health insurance among cell phone only adults residing in MO and those who are not cell phone only. As a further illustration, we recomputed estimates based upon the subdomain defined as "have a landline" that were available from either frame and examine two dichotomous survey outcomes: "Have you ever been told by a physician that you have diabetes?" The final estimates and ARB values for the key outcomes are displayed in Figure 4. Compared to the OMNI Survey, the differences in the estimates based on OLC and No OLC weights were a bit larger and this can be traced back, in part, to relatively larger sampling fractions from each of the respective sampling frames. We note that the design effects estimated by using SF-BP No OLC weights seem reasonable and consistent with the fact that the sampling design incorporated stratification with unequal probabilities of selection across the strata.

Table 4: HL-MO Survey estimates derived using the dual frame base weight method along with comparable estimates derived from two external state-specific surveys.

Variable	HL-MO Horvitz-Thompson Estimate (OLC)	HL-MO Hansen-Hurwitz Estimate (No OLC) (95% CI)	HL-MO Hansen-Hurwitz Estimate Design Effect	BRFSS-MO (2010) Survey Estimates	Absolute Relative Bias Estimates Horwitz-Thompson Estimator	Absolute Relative Bias Estimates Hansen-Hurwitz Estimator
Health Care Coverage (any Type)	81.9804%	81.9784% (80.1686, 83.7881)	1.9313	84.9831%	3.6627%	3.6652%
Told by H.C. Provider of Diabetes	11.3597%	11.3599% (10.0092, 12.7107)	1.5795	9.4051%	17.2064%	17.2079%
Told by H.C. Provider of Asthma	14.2899%	14.2902% (12.7815, 15.8618)	1.7527	14.2120%	0.5451%	0.5472%
Excellent or Very Good Health (Self Report)	49.1923%	49.1917% (46.9698, 51.4136)	1.7208	51.6192%	4.9335%	4.9348%
Have a Doctor #	76.7251%	76.7233% (74.7730, 78.6737)	1.8561	83.6693%	9.0508%	9.0533%
Told by H.C. Provider of Cancer	8.1654%	8.1654% (7.0128, 9.3180)	1.5407	9.2942%	13.8242%	13.8242%
Number of Adults in Household (average)	1.9689	1.9688 (1.9267, 2.0110)	2.0037	2.2185	12.6790%	12.6832%
Age (average)	47.0076	47.0074 (46.2413, 47.7735)	1.5980	47.1277	0.2555%	0.2560%

The BRFSS asked about "personal doctor or health care provider"; the HL-MO survey asked about whether respondent had "one person that was considered to be a personal doctor"

DISCUSSION

In this paper we described and applied the SF-BP method for weighting dual frame surveys at both a national and state scale using basic principles of probability. We note that the overlap of these frames does not occur at the sampling unit level (i.e. numbers) but at the respondent level who is identified by these numbers. The strengths of the SF-BP method include easy implementation in the survey software as a Horwitz-Thompson type “single-frame” estimator (OLC) (assuming that duplicate units in the sample are only used once in the estimation) or as a Hansen-Hurwitz type estimator with No-OLC. In our examples we found virtually no differences in estimates from an assortment of variables both at the national and state scale for estimates derived by considering or ignoring the overlap in the two frames. Moreover, the unequal weighting effects and overall absolute relative biases for these variables were nearly identical for base weights that included and ignored the overlap of the landline and cell phone frames. This result is driven in large part by the sampling fractions applied to both the cell and landline frames. We note that one direct advantage of ignoring the overlap in computing the SF-BP sampling weights is that design effects can de derived directly using the frame as a so called super stratum in the design statements of either Sudaan or SAS’s Survey Procs. Another direct advantage of the SF-BP method is that inclusion probabilities are computed directly with respect to joint inclusion across the two frames which eliminates the need to estimate a composite factor that is discussed in more general applications of dual frame estimation (Lohr and Rao, 2000 and 2006; Metcalf and Scott, 2009;) or within the RDD survey context (Brick, Edwards and Lee, 2007; Brick, Cervantes, et al. 2011). Finally, because the inclusion probabilities are connected to phone numbers (sampling unit) which are directly related to persons (i.e. target population), post-stratification adjustments are needed to translate the inclusion probabilities (computed for the population of numbers) to interpretable and meaningful target population values (i.e.

persons). The post-stratification adjustments can be applied to the entire sample using the SF-BP base weights without having to separate the sample first by cell phone status or sampling frame. We also note that because the SF-BP method is cased in basic principles of probability, it is easily scalable to multiple frames.

As we illustrated here some of the absolute relative biases are high and in part may be due to slight measurement differences between the national and state survey examples and the external government data sources and in part due to the rather limited non-response adjustments that were performed in the two survey examples. We might expect that with further nonresponse adjustments, the relative biases may decrease (at the possible expense of increasing the unequal weighting effect). Further work is needed to evaluate and compare the SF-BP approach for weighting dual frames within the Total Survey Error context. We hope this paper can contribute to the ongoing development of best practices for weighting dual frame surveys.

Acknowledgements

We wish to thank Jonaki Bose of SAMHSA for her assistance in obtaining NHSDU estimates of lifetime cigarette use. We also thank the Missouri Foundation for Health for permission to use the data from Health Literacy Missouri for illustration of our methodology. Finally, we thank the owners of PSRAI (Evans Witt, Larry Hugick and Mary McIntosh) for their generosity in including the OMNI Survey questions as part of their weekly survey.

References

AAPOR Cell Phone Task Force. "New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys within the U.S. With Respondents Reached Via Cell Phone Numbers," Lenexa, KS: American Association of Public Opinion Research; 2010. Available at: http://aapor.org/AM/Template.cfm?Section=Cell_Phone_Task_Force&Template=/CM/ContentDisplay.cfm&ContentID=2818 , accessed June 10, 2012.

Bankier, MD (1986) Estimators Based on Several Stratified Samples with Applications to Multiple Frame, *Journal of the American Statistical Association*, Vol. 81(396), 1074-1079.

Battaglia M, Eisenhower D, Immerwahr S, and Konty K. "Dual Frame Weighting of RDD+Cell Phone Interviews At the Local Level" Presented at the 65th Annual Conference of the American Association of Public Opinion Research, Chicago, IL, May 13-16, 2010.

Best, J "First-Stage Weights for Overlapping Dual Frame Telephone Surveys". Presented at the 65th Annual Conference of the American Association of Public Opinion Research, Chicago, IL, May 15, 2010.

Blumberg SJ, Luke JV. (2009) Reevaluating the need for concern regarding noncoverage bias in landline surveys, *Am J Public Health*, 99(10):1806–1810.

Blumberg SJ, Luke JV, Davidson G, et al. Wireless substitution: State-level estimates from the National Health Interview Survey, January-December 2007. National health statistics report; no 14. Hyattsville, MD: National Center for Health Statistics. 2009.

Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2010. National Center for Health Statistics. June 2011. Available from: http://www.cdc.gov/nchs/nhis.htm.

Brick JM, 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*, 70:780–93.

Brick, JM., 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: 23-39.

Brick JM, Edwards WS, 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(5):793-813.

Brick, J.M., Flores-Cervantes, I., Lee, S., and Norman, G. (2011) "Nonsampling Errors in Dual Frame Telephone Surveys." Survey Methodology, 37(1), 1-12.

Buskirk, TD "Dual Frame Surveys" in Encyclodedia of Survey Research Methods, Sage Publications, © 2008.

Buskirk, T. (2010) Health Literacy Missouri Survey Weighting Technical Documentation, Unpublished Technical Report.

Centers for Disease Control and Prevention. 2011, "Behavioral Risk Factor Surveillance System Questionnaire", Retrieved April 10, 2012, (http://www.cdc.gov/brfss/questionnaires/pdf-ques/2011brfss.pdf)

Callegaro, M., Steeh, C., Buskirk, T., Vehovar, V., Kuusela, V., and Piekarski, L. 2007. Fitting Disposition Codes to Mobile Phone Surveys: Experiences From Studies in Finland, Slovenia, and the USA. *Journal of the Royal Statistical Society A*, 170(3), 647-670.

Frankel, MR, Battaglia, MP, Link, MW and Mokdad, AH (2007) Integrating cell phone numbers into Random Digit-Dialed (RDD) landline surveys, *Proceedings of the American Statistical Association Section on Survey Research Methods*, 3793-3800.

Guterbock, T. Estimating Phone Service and Usage Percentages: How to Weight the Data From a Local, Dual-Frame Sample Survey of Cell Phone and Landline Telephone Users in the United States. Paper presented at the 64th annual meeting of the American Association for Public Opinion Research, Hollywood, FL, May, 2009.

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.

Hu SS, Balluz L, Battaglia M, Frankel MR. (2011) Improving Public Health Surveillance Using a Dual-Frame Survey of Landline and Cell Phone Numbers, *American Journal of Epidemiology*, 173:703-711.

Izrael D, Battaglia M, Frankel M. Extreme survey weight adjustment as a component of sample balancing (a.k.a. raking). (Paper 247). In: Proceedings of the SAS Global Forum 2009 Conference. Cary, NC: SAS Institute Inc; 2009. Available at: http://support .sas.com/resources/papers/proceedings09/247-2009.pdf, accessed June 10, 2011.

Kalton G, and Flores-Cervantes I. (2003) Weighting Methods. *Journal of Official Statistics*, 19(2): 81-97.

Kennedy C. "An Evaluation of Popular Weighting Approaches in Dual Frame RDD Surveys". Presented at the 66th Annual Conference of The American Association of Public Opinion Research, Phoenix, AZ, May 12-15, 2011.

Lavrakas, PJ, Shuttles, CD, Steeh, C and Fienberg, H. (2007) The state of surveying cell phone numbers in the United States 2007 and Beyond, Public Opinion Quarterly, Vol. 71(5), 840–854.

Liu, J, Iannacchione, V and Byron, M. Decomposing Design Effects for Stratified Sampling, Proceedings of the American Statistical Association's Section on Survey

Research Methods, 2002. Available at: http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-001069.pdf, accessed on September 20, 2012.
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.
Metcalf, P. and Scott A. (2009) Using multiple frames in health surveys. *Statistics in Medicine*, 28, 1512-1523.
Metcatti, F. (2007) A single frame multiplicity estimator for multiple frame surveys. *Survey Methodology*, Vol. 33 (2): 151-157.
Srinath, KP, Battalia, MP, Khare, M (2004) A Dual Frame Sampling Design for an RDD Survey that Screens for a Rare Population, *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 4424-4429.
Steeh, C., Buskirk, T., and Callegaro, M. (2007) Use of Text Messages in U.S. Mobile Phone Surveys, *Field Methods*, 19:59-75.
Zuwallack, R. (2009) Piloting Data Collection via Cell Phones: Results, Experiences and Lessons Learned, *Field Methods*, 21(4): 388–406.

APPENDIX A: Question Wording for Omnibus Survey Questions Displayed in Table 1

1. Employment: Are you now employed full-time, part-time, or not employed?
2. Smoking: CIGS Have you smoked at least 100 cigarettes in your entire life? [INTERVIEWER NOTE: 5 packs = 100 cigarettes]
3. Adults: How many adults, age 18 and over, currently live in your household INCLUDING YOURSELF?
4. Parent/Guardian: Are you the parent or guardian of any children under 18 years of age?

APPENDIX B: Question Wording for MOHL Survey Questions Displayed in Table 2

1. General Health: In general, would you say your general health is:
2. Regular Doctor: A personal doctor is the health provider who knows you best. Do you have one person you think of as your personal doctor?
3. Regular Doctor: A personal doctor is the health provider who knows you best. Do you have one person you think of as your personal doctor?
4. Diabetes: The next few questions are about your personal health. Have you ever been told by a doctor that you have diabetes?
5. Asthma: Have you ever been told by a doctor that you have asthma?
6. High Blood Pressure: Have you ever been told by a doctor that you have high blood pressure?
7. Health Insurance Coverage: Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare?
8. Adults: How many adults aged 18 or over live in your household, including yourself?