

Figure 1 Geodemographic composition of adults living in CPO households (Source: Blumberg and Luck 2012).

ARTICLES

Practical Guidelines for Dual-Frame RDD Survey METHODOLOGY (Now That the Dust is Settling)

Mansour Fahimi¹

¹ GfK Custom Research, LLC

Keywords: and design effect., optimal sample allocation, cell phone only, dual-frame rdd

<https://doi.org/10.29115/SP-2014-0011>

Survey Practice

Vol. 7, Issue 2, 2014

Practical Guidelines for Dual-Frame RDD Survey METHODOLOGY (Now That the Dust is Settling)

Overview

For nearly two decades, the traditional sampling methodology of list-assisted landline random digit dialing (RDD) has served as the survey research workhorse for population-based studies. In recent years, however, virtually all RDD surveys have come to rely on dual-frame techniques in an attempt to improve coverage. This change is primarily due to the growing number of households that are abandoning their landline phones and relying exclusively on cellular services (Fahimi and Kulp 2009). Figure 1 shows the geodemographic composition of adults living in households without landline services, highlighting the potential coverage bias that can result should such individuals be excluded from sample surveys. Consequently, the dual-frame RDD (DFRDD) technique has become the standard practice whereby samples of landline telephone numbers are supplemented with cellular numbers to produce probability-based samples of all households, including the so-called cell phone only (CPO) households.

While including cellular numbers has offered an effective remedy for improving coverage of the traditional landline samples, current practice of DFRDD methodology is subject to technical and operational inconsistencies. On the technical side, most survey researchers rely on ad-hoc assumptions to determine the mixture of landline and cellular numbers for their samples. This

inconsistency, which is mostly due to unavailability of current counts of CPO households, has implications for both sample selection as well as subsequent methods used to weight the resulting survey data.

On the operational front, some surveys continue to rely on complicated data collection protocols that sap the available budget without producing any notable gains. For example, a number of surveys screen out dual-use respondents (those reachable by both cellular and landline phones) using ad-hoc thresholds for answering landline calls. In this paper, we introduce a simple strategy for DFRDD surveys that ameliorates the existing inconsistencies for sample selection and weighting applications, as well as eliminates the need for the costly practice of screening out dual-use responders.

Nature of the Problem

With n_i and N_i representing the sample and population sizes, S_i and C_i denoting the variability of any key outcome measure and data collection cost for the i -th stratum, respectively, the following standard formula is often used to provide guidelines for allocation of the total sample to each stratum (Cochran 1977). Here, these parameters are indexed simply by c for CPO and l for households with at least one landline. Accordingly, for any DFRDD survey, an optimal allocation of the total sample size (n) to the two strata (n_c and n_l) should be in reverse relation to the cost and direct relation to the size and variability in each stratum. Furthermore, this allocation reduces to simpler forms depending on the homogeneity across the two sampling strata of variability ($S_c \rightarrow S_l$) and/or cost ($C_c \rightarrow C_l$) measures.

$$\begin{aligned} n_c &= n \times \frac{N_c S_c C_c^{-0.5}}{N_c S_c C_c^{-0.5} + N_l S_l C_l^{-0.5}} \\ &\rightarrow \left\{ n \times \frac{N_c C_c^{-0.5}}{N_c C_c^{-0.5} + N_l C_l^{-0.5}}, S_c \right\} \\ &\rightarrow \left\{ S_l \frac{N_c S_c}{N_c S_c + N_l S_l}, C_c \right\} \\ &\rightarrow C_l \frac{N_c}{N_c + N_l}, \{C_c \rightarrow C_l S_c \rightarrow S_l\} \end{aligned}$$

However, the input parameters required for the above allocation schemes pose two separate problems. First, current estimates for the number of CPO and landline households at different levels of geography have not been available until recently. Second it is not currently possible to select RDD samples that are dedicated to such households separately. Numbers from the landline frame can reach households that are reachable via cell phone as well, just as numbers from the cellular frame can also reach landline households. The only point of differentiation is that landline-only (LLO) households are only covered by the landline frame, whereas CPO households are only covered by the cellular frame.

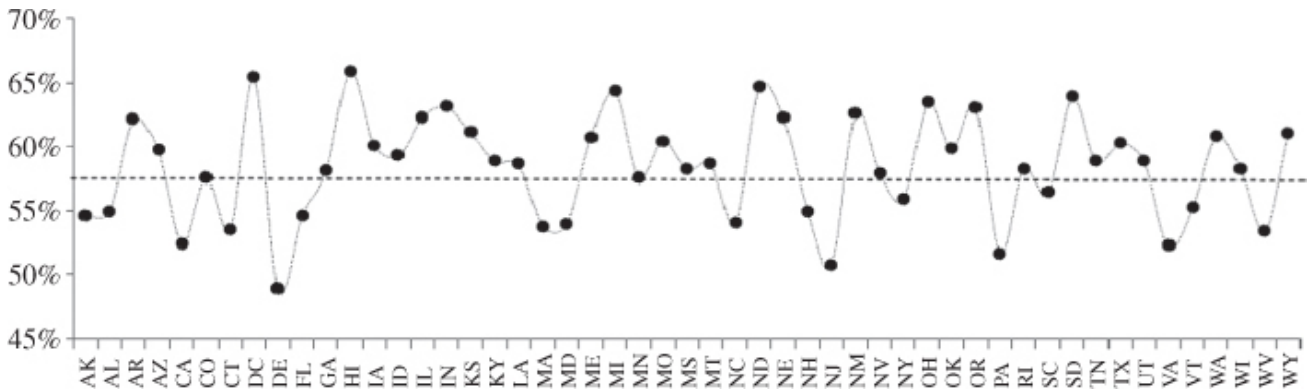


Figure 2 Percent nonproductive landline numbers by state (Source: MSG 2013).

Join this with the notable variability of CPO household rates in different geographic domains (see Figures 4 and 5), and it is no wonder the DFRDD sampling methodology has been applied so inconsistently in recent years. That is, applications of inconsistent sample allocation schemes and weighting adjustments based on varying target totals. Arguably, this can explain much of the variability observed in the 2012 presidential election polling results produced from surveys that have relied on this methodology.

What further complicates the situation is that the working rates for landline and cellular numbers are not similar. This is particularly true of areas with geodemographic compositions atypical of the nation. For instance, Figure 2 shows estimates of the percentage of nonproductive landline numbers – nonworking or nonresidential – for a national RDD.

Chasing the Moving Target

Figure 3 shows how rapidly the numbers of adults living in CPO households has been changing in each state, increasing the urgency for reliable and current estimates for this critical design parameter. Given the impressive pace and geographically diverse rates of this change, available estimates of CPO households that rely on survey data from prior years cannot produce precise enough guidelines for determining the right mixture of landline and cellular numbers. This insufficiency is substantially magnified for surveys that target lower geographic levels (e.g., counties) because of the remarkable variability in the number of CPO households within states.

In the next section, we will discuss the methodology developed by Marketing Systems Group (MSG) for producing current estimates of CPO households for all counties in the United States (see Figure 4). These estimates, which are expected to be updated quarterly, will provide the missing guidelines for determining the proper mixture of landline and cellular numbers for DFRDD samples. Moreover, these county estimates can be rolled up to higher levels of aggregation and provide the needed control totals for proper construction of survey weights.

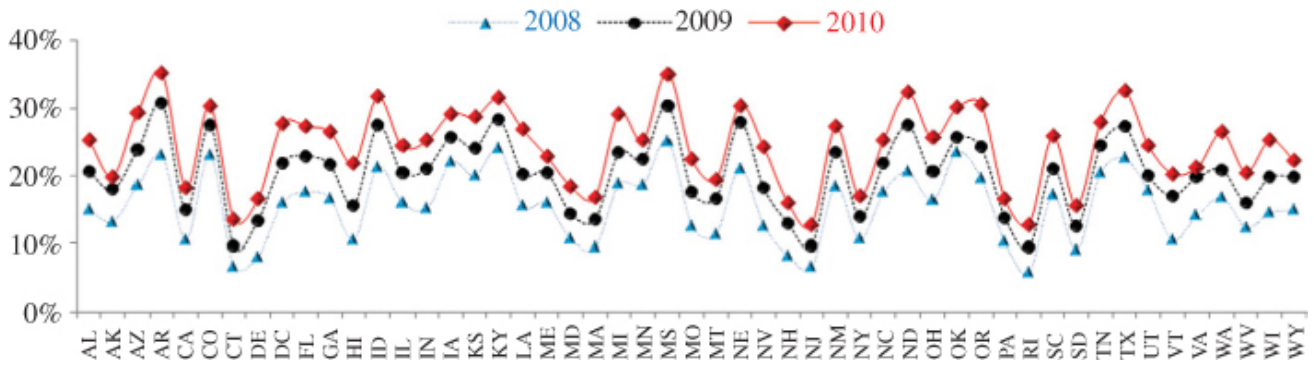


Figure 3 Percent adults living in CPO households by year and state (Source: Blumberg and Luck 2012).

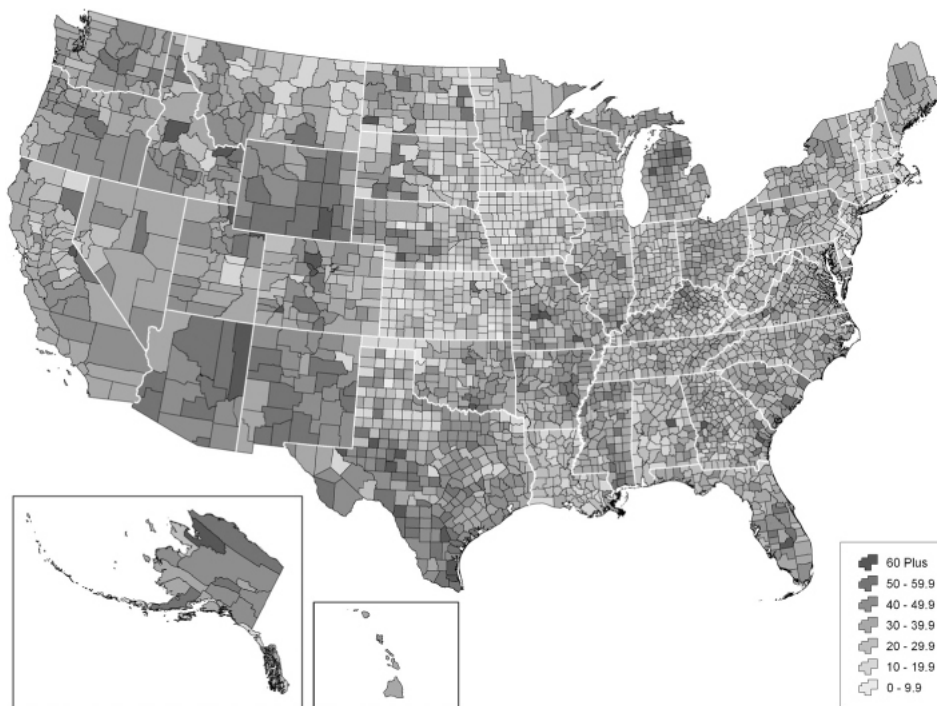


Figure 4 April 2013 estimates of percent CPO households by county (Source: MSG 2013).

Estimation of County-Level Counts of CPO Households

Developing survey-based estimate of any population parameter for all counties in the United States is a cost-prohibitive proposition. Even a modestly reliable survey can require close to one million completed interviews to produce estimates of CPO households counts for each of the 3,143 counties. Moreover, because of the rapidly changing nature of this measure, it would be necessary to repeat such a survey on regular basis. Clearly, this idea is at best academic and practically impossible to implement.

However, by relying on various public and commercial databases this gap has been bridged to produce current estimates for the number of CPO households for different geographic domains (Fahimi and Kulp 2009). These estimates are

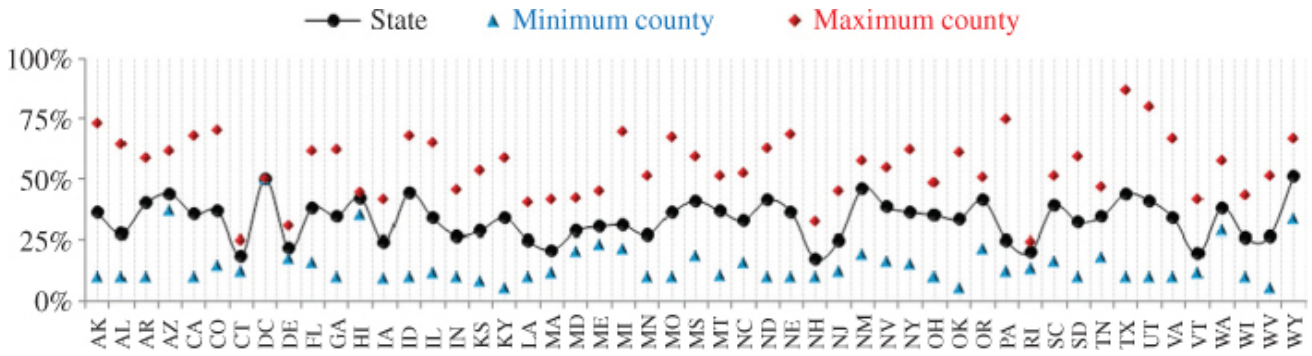


Figure 5 April 2013 estimates of percent CPO households by state and county (Source: MSG 2013).

not survey-based and can be produced several times throughout each year for domains as small as counties, using an intuitively simple triangulation process as follows:

1. Starting with the number of occupied housing units – adjusted for seasonal, vacant, and non-telephone households – number of telephone households is produced for each county. (Telephone households can be reached by at least one telephone, regardless of type.)
2. Starting with the number of landline assignments – adjusted for the prevalence of multiline households in different location types in each state – number of landline households is produced for each county.
3. Subsequently, number of CPO households in each county is produced as the difference between the above two estimates. (CPO households = telephone households – landline households.)

Figure 5 shows the state-level estimates for the rate of CPO households along with those for the counties with the lowest and highest CPO rates in each state. Once again, the huge variability in CPO rates at the state and county levels signifies the importance of this crucial parameter when designing and weighing DFRDD surveys.

Reducing the Cost of Data Collection

In addition to eliminating the inconsistencies currently exercised when designing and weighting DFRDD surveys, current estimates for the number of CPO households also eliminates the need to screen out respondents who use both landline and cellular services. This important cost saving step also removes the awkward practice of excusing otherwise cooperating respondents.

Including all dual-use respondents is particularly beneficial for surveys, such as the BRFSS that use specific thresholds of landline usage as screening criteria. Obviously, there cannot be a scientific base for establishing such a threshold for a measure as fluid as the abandon rates for landline usage. By not eliminating

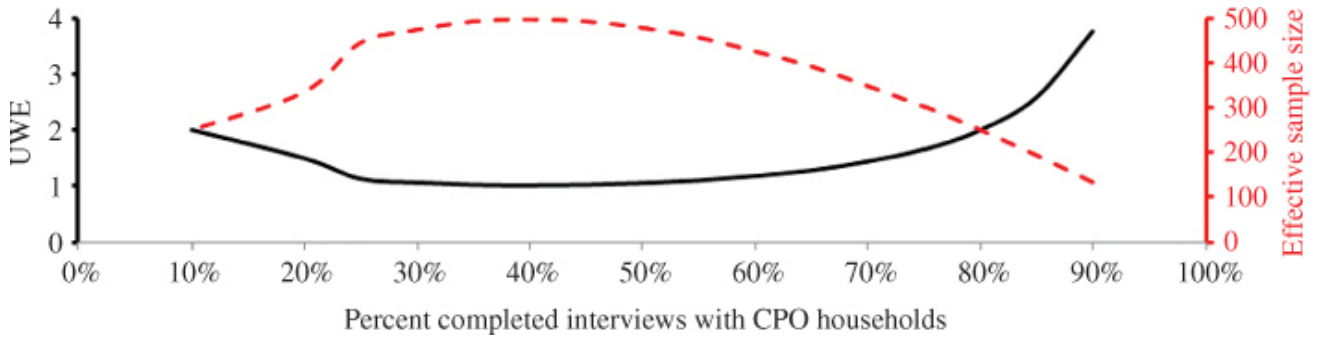


Figure 6 Reduction in effective sample size as a function of sample allocation.

such respondents, data collection costs for cellular respondents will no longer be several times that of their landline counterparts. This cost-saving can justify a deservedly larger number of interviews with cellular respondents in DFRDD surveys.

Putting Theory to Practice

Let us assume a DFRDD survey is to be conducted to secure 500 interviews in a county with an estimated CPO household rate of 40 percent. Consequently, the cellular RDD component of the sample must be large enough so that about 40 percent of survey respondents would be from CPO households as well. However, even without screening out dual-users, the cost of data collection is still slightly higher for cellular respondents. This is in part due to FCC regulations that prohibit the use of autodialers when calling cellular numbers, and the fact that no effective prescreening services are currently available for removing nonproductive cellular numbers.

On the one hand, cost-saving considerations suggest reducing the size of the cellular sample component. However, any departure from proportional allocation will have unequal weighting effect (UWE) implications that will lead to inflation of the error margins for survey estimates. In order to gauge precision loss due to application of weights needed to offset disproportionate sample allocations, this inflation is often approximated during the design phase by the following formula.

$$\delta = 1 + \frac{1}{\bar{W}^2} \sum_i \frac{(W_i - \bar{W})^2}{n-1}$$

The UWE changes as various shares of completed interviews are allocated to CPO households (Figure 6). The black solid line represents the proportion of CPO households in the sample, while the red line dashed represents the reactionary change in effective sample size – hence highlighting the need for striking a practical balance when designing the sample.

Accordingly, the largest effective sample size (500) is achieved when exactly 40 percent of completed interviews are secured with CPO households. That is, no loss to precision results with an UWE of unity when the target number of

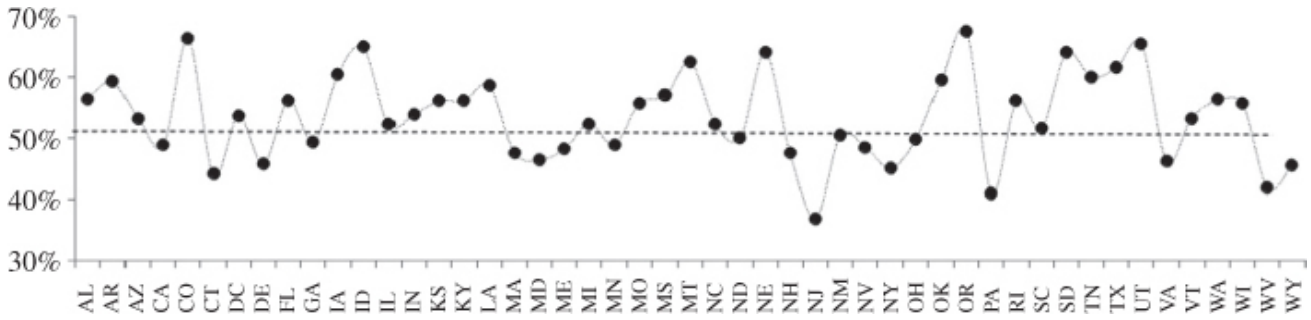


Figure 7 Percent cellular calls reaching CPO households by state (Source: SSRS 2013).

CPO interviews are achieved. In contrast, UWE doubles when only 10 percent of interviews are completed with CPO households, resulting in an effective sample size that is only half of the starting sample. Alternatively, when 25 percent of interviews are with CPO households, the UWE is only slightly over 1 and the effective sample size is 446. Arguably, this reduction in the effective sample size is fairly tolerable as compared to more aggressive allocation schemes that aim to further reduce data collection costs yet result in progressively smaller effective sample sizes.

As a simple rule of thumb, the share of completed interviews with CPO households can be the actual percentage of CPO households in the geography of inference divided by the cost ratio of one cellular interview to that of a landline. For example, assuming for the above example this cost ratio is 1.5-to-1, then about $133 = 500 \times 0.4 / 1.5$ of all completed interviews should be with CPO households. That is, instead of 40 percent a compromised 27 percent $= 133 / 500$ of the completed interviews will be with CPO households. Of note, as the cost difference between cellular and landline interviews approaches zero, then the percentage of interviews with CPO households should approach the target percentage for the corresponding geography to reduce undue UWE.

Another important consideration when determining the sample mixture for DFRDD surveys has to do with the fact that, on average, only about one-half of all cellular interviews end up with CPO households. Consequently, when aiming for a fixed number of CPO interviews, twice as many cellular interviews have to be completed to account for cellular respondents who will not emerge as CPO. Hence, for the above example, this means securing about $266 = 133 \times 2$ cellular interviews. Figure 7 shows the percentage of cellular calls expected to reach CPO households by state, which can vary by about 30 percentage points.

Finally, the calculation of the needed number of landline and cellular sample numbers must also include other important yield rates, such as nonworking, ineligibility, and refusal rates. Table 1 provides a simple prototype for sample size calculations for a DFRDD survey that aims to complete 500 interviews,

Table 1 Sample size calculation prototype for completing 500 interviews via DFRDD.

Category	Cellular		Landline (1+listed frame)		Total
	Count	Rate	Count	Rate	
Starting Sample	4,860		4,130		8,990
Working numbers	3,159	65%	1,446	35%	4,605
Residential/personal	2,369	75%	1,229	85%	3,598
Eligible	1,776	75%	1,167	95%	2,944
Respondents	266	15%	234	20%	500
CPO Respondents	133	50%	0	0%	133

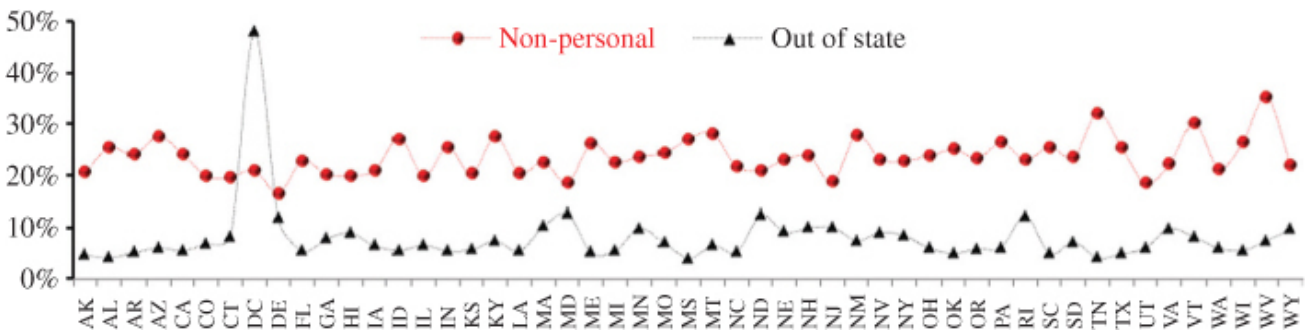


Figure 8 Cellular calls received on personal devices or out of state by state (Source: BRFSS).

of which about 27 percent are expected to be with CPO households. The assumed yield rates here are often used for national surveys for which most contacted households qualify.

Clearly, depending on specific features of a survey and its budgetary constraints, the assumed rates in the above table stand to be modified accordingly. Moreover, there are other yield issues that may have to be considered beyond what is reflected here. For instance, Figure 8 shows the rates at which cellular calls are received on nonpersonal devices and those that reach individuals residing outside of the state their cellular number is associated with. The latter can be a more pronounced issue for smaller geographic domains, since a progressively larger number of cellular calls can reach individuals outside of their expected locations. Last a portion of cellular calls are received by individuals <18 years of age, which currently averages at about 8.5 percent at the national level but varies state-to-state.

Targeting Cellular RDD Samples

Constructing cellular RDD frames, particularly for substate geographic domains, is subject to both operational and definitional challenges because cellular numbers are assigned to mobile devices that may be located anywhere across the globe. This is in spite of the fact that cellular numbers in the United

States are assigned in blocks of 1,000 numbers (1-K Blocks) to wireless service providers that serve subscribers in specific locations. Given the existing sampling imprecisions, geographic targeting for cellular RDD can be accomplished in several ways.

1. Wire centers consist of physical structures to house telecommunication equipment, such as landline and cellular switches, used for routing and connecting calls. As such, the location of cellular switches can be used to identify the set of 1-K blocks of cellular numbers that should be included in the sampling frame for a geographic area of interest. However, this methodology can result in sizable coverage problems.
2. Rate centers delineate the local call boundaries set by service providers for billing purposes. Using rate centers to construct the cellular RDD sampling frame can reduce many of the coverage difficulties associated with wire-center methodology. Moreover, relying on various *Telcordia* databases it is possible to identify specific subsets of cellular 1-K blocks that are assigned to different wireless service providers covering the area of interest.
3. Billing zip codes associated with cellular accounts are available for about one-third of all active cellular numbers. While this subset is not random and the matching process can only take place after a cellular RDD sample has been selected, it is possible to more accurately identify the potential areas where cellular subscribers reside.

Also, other processes are currently being investigated that would allow identification of nonproductive cellular numbers. With such processes, it will be possible to identify active numbers before the selected sample is released to a telephone center for data collection.

Weighting Considerations

Virtually all survey data must be weighted in order to reduce the bias caused by coverage issues and differential nonresponse. Moreover, weights will be necessary if, by design, the sample includes disproportionate allocations. This is certainly the case for DFRDD surveys because the mixture of cellular and landline telephone numbers is often based on cost considerations rather than their true proportions in the geography of inference.

While the weighting process for DFRDD surveys is fundamentally similar to that for other probability-based sample surveys, there are a few steps unique to this type of survey. In most surveys, base weights are computed as the reciprocal of selection probabilities. Next, these starting weights are adjusted

for differential nonresponse and then poststratified to known survey population figures. For DFRDD surveys, typically, these steps are carried out as follows:

1. Base weights need to reflect any differential selection probabilities, both at the level of primary sampling units as well as any subsequent subselections that may take place down the sampling path. For DFRDD, base weights are computed as the ratios of available telephone numbers to those selected in each stratum, separately for the landline and cellular strata. The resulting weights are then adjusted to compensate for the following:
 - a. Subsampling can occur in several ways, such as oversampling of listed landline numbers to increase residential hit rates. This step must compensate for any employed oversampling for each stratum.
 - b. Multiple landline households are rapidly disappearing; however, an adjustment can be made to compensate for the resulting frame multiplicity by dividing the starting base weights by the number of landline numbers serving the household. If applied, this adjustment is typically capped to a factor of about 2.
 - c. Within household subsampling occurs when a subset of eligible householders is selected at random within each household. Even though this selection is not always carried out in a completely random fashion, base weights should be multiplied by the number of eligible individuals in each household to reflect the imposed subsampling.
 - d. Trimming of extreme weights will be less of an issue at the end should such weights be reduced during the early steps of weighting. Because of this and the fact that counts of landline and cellular telephone numbers do not inaccurately reflect universe counts, it is advisable to apply some form of smoothing of the base weights at this stage.
2. Nonresponse adjustment is a weighting refinement that requires information about both respondents as well as nonrespondents. While this adjustment can effectively compensate for some of the observed differential nonresponse, for DFRDD surveys this step is typically excluded because very few RDD frame data exist for both respondents and nonrespondents.
3. Poststratification is often the final stage of weight adjustment, whereby adjusted base weights are calibrated with respect to a set of geodemographic characteristics so that the final survey weights aggregate to the corresponding reported totals. This step is typically

carried out simultaneously with respect to several characteristics using an iterative process commonly referred to as raking. For DFRDD, the raking dimensions typically include the following:

- a. Geodemographic characteristics such as counts of eligible population in different areas indexed by gender, age, race/ethnicity, and education.
- b. Telephone status is the critical dimension that can correct for any misalignment that has occurred, by design or differential nonresponse, depending on whether the respondent lives in a CPO household or otherwise. As mentioned earlier, this is where current CPO estimates play a critical role in developing proper survey weights.

Concluding Remarks

Prior to the 1960s, most sample surveys were conducted either by mail or face-to-face because a nonignorable percentage of households did not have access to telephones. Starting in 1990s, however, telephone became a prominent tool for survey administrations as the prevalence of nontelephone households diminished to single percentages (Brick et al. 1995). Ironically, in recent years exactly the same old criticism has been leveled against landline telephone samples due to coverage error associated with the exclusion of households without landline telephones. Thornberry and Massey (1988) observed that the percentage of households without telephone had declined from about 20 percent in 1963 to about 10 percent in the early 1970s, and then to <5 percent in mid 1990s. Yet in 2008, Blumberg and Luck 2012 indicated that the opposite trend was reoccurring as the percentage of households without landlines was rapidly increasing – currently estimated at more than 35 percent.

In addition to reducing the cost of the traditional methods of data collection, several studies began showing that data quality from telephone surveys rival that from more expensive options via mail or face-to-face modes (Groves and Kahn 1979; Hochstim 1967). It is because of these supporting arguments that in spite of its temporary hiatus during the turn of the century, RDD methodology has come back to reclaim its status as an effective method of data collection. Of course, this reemergence has been made possible through a series of fundamental refinements, such as supplementation of the traditional method of landline RDD with cellular numbers.

The resulting method of DFRDD, however, has been experiencing various growing pains, as earlier practices had to rely on cumbersome and expensive detours that spanned across sampling, data collection, and weighting phases. Up to this day, sample selection for DFRDD surveys is subject to inconsistent blending of landline and cellular telephone numbers – an inconsistency that

carries through the weighting steps as well. On the other hand, there are still surveys that continue to rely on costly options of screening out otherwise cooperating respondents based on ad-hoc criterion.

It is safe to argue that many of the earlier and existing inconsistencies with which the method of DFRDD has grappled are due to inadequate understanding of the illusive CPO subpopulation. Other than the occasional survey-based estimates about the size of this group, which were subject to nonignorable sampling errors and only available for larger geographic domains, researchers had to rely on improvised assumptions when designing and weighting DFRDD surveys. However, with the availability of quarterly updated counts of CPO households, it is now possible to design and weight such surveys based on reliable population parameters.

Additionally, data collection protocols can now avoid abandoning a good number of respondents whom interviewers have to work hard to secure their corporations, simply because they are reachable by both cellular and landline telephones. And in some instance because they receive 20 percent of their calls on landlines as compared to 21 percent! Here too, availability of current estimates for the number of CPO households at virtually all levels of aggregation makes it possible to abandon such practices and instead divert the resulting resources to securing larger sample sizes or implementing more robust options for refusal conversions.

REFERENCES

- Blumberg, S.J., and J.V. Luck. 2012. "Wireless Substitution: State-Level Estimates from the National Health Interview Survey, 2010-2011." In *National Health Statistics Reports, CDC*. Vol. 61. <http://www.cdc.gov/nchs/data/nhsr/nhsr070.pdf>.
- Brick, J.M., J. Waksberg, D. Kulpand, and A. Starer. 1995. "Bias in List-Assisted Telephone Samples." *Public Opinion Quarterly* 59: 218–35.
- Cochran, W.G. 1977. *Sampling Techniques*. 3rd ed. New York: Wiley.
- Fahimi, M., and D. Kulp. 2009. "Topology of the Landline Telephone Sampling Frame." *Survey Practice* 2 (9).
- Groves, R.M., and R.L. Kahn. 1979. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Hochstim, J. 1967. "A Critical Comparison of Three Strategies of Collecting Data from Households." *Journal of the American Statistical Association* 62: 497–989.
- Thornberry, O., and J. Massey. 1988. "Trends in United States Telephone Coverage across Time and Subgroups." In *Telephone Survey Methodology*, edited by R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls, and J. Waksberg. New York: Wiley.