The Best of Both Worlds: A Sampling Frame Based on Address-Based Sampling and Field Enumeration

Vincent Iannacchione[1], Katherine Morton[2], Joseph McMichael[2]
Bonnie Shook-Sa[2], Jamie Ridenhour[2], Stephanie Stolzenberg[2],
David Bergeron[2], James Chromy[2], and Arthur Hughes[3]

[1]RTI International, 701 13th ST NW, Suite 750, Washington, DC 20005
[2]RTI International, 3040 East Cornwallis RD, Research Triangle Park, NC 27709
[3]Substance Abuse and Mental Health Services Administration, Rockville, MD 20857

Abstract

Cost savings are the primary advantage of using address-based sampling (ABS) over field enumeration (FE) for in-person surveys of the civilian, non-institutionalized population. These cost savings are tempered by research which indicates that FE provides more complete coverage than ABS, especially in rural areas. We developed and piloted a hybrid sampling frame for the National Survey on Drug Use and Health (NSDUH) that uses ABS supplemented with a frame-linking procedure in area segments where we anticipate adequate ABS coverage and FE in segments where we anticipate poor ABS coverage. The objective of the hybrid frame is to lower costs without sacrificing coverage levels of the current NSDUH sampling frame which is based solely on FE. We report on the trade-offs between coverage and cost savings as area segments are shifted from FE to ABS.

Key Words: in-person surveys, frame-linking procedures, frame coverage, survey costs

1. Introduction

We developed and field tested a hybrid sampling frame for the National Survey on Drug Use and Health (NSDUH) that is designed to produce significant cost savings without sacrificing the coverage of the current sampling frame that is based exclusively on field enumeration (FE) supplemented with the half-open interval (HOI) frame-linking procedure (Kish 1965).

The current NSDUH sample design is implemented with a five-step process:

1. Stratify each State into a specified number of regions.
2. Select census tracts within regions.
3. Select segments (which are a collection of adjacent census blocks) within sampled census tracts.
4. Select dwelling units (DUs) within sampled segments.
5. Select 0, 1, or 2 persons within sampled DUs.

Steps 1, 2, and 3 are completed using data readily available electronically from the U.S. Census Bureau and other sources. However, before step 4 can be completed, field staff

(called "listers") physically travel to the sampled segments to field enumerate all DUs in each sampled segment. Segments for the NSDUH comprise one or more adjacent census blocks that in combination meet or exceed the minimum requirement of 100 DUs in rural areas or 150 DUs in urban areas. The enumeration data is used by statisticians to select a sample of DUs. Field interviewers visit sampled DUs in person where a screening is done to determine if 0, 1, or 2 persons will be selected for the NSDUH interview. Before the field interviewer leaves each selected DU, units that may not have been documented by the lister are identified using the HOI.

The hybrid frame relies on an address-based sampling (ABS) list that consists of a predetermined list of locatable city-style addresses based on the Computerized Delivery Sequence File (CDSF) maintained by the U.S. Postal Service. A locatable city-style mailing address has a street name and number, a unit number if appropriate, a city, state, and ZIP Code. The ABS list is then supplemented with the Check for Housing Units Missed (CHUM) frame-linking procedure to identify and record DUs that are not included on the ABS list (McMichael et al. 2008). The ABS-based frame is not expected to completely replace the current FE-based frame. FE and the HOI procedure will be retained in segments where ABS coverage is expected to be low. Research from prior studies, e.g., Iannacchione et al. (2007) or O'Muircheartaigh et al. (2007), indicate that rural areas are more likely to have low ABS coverage than urban areas.

Operationally, the hybrid sampling frame would be implemented as follows:

1. Select the sample of segments.
2. Obtain estimates of the number of locatable city-style mailing addresses and the number of DUs in each sampled segment.
3. Estimate the expected ABS coverage rate for each sampled segment. The current method for estimating this rate is the ratio of the number of city-style mailing addresses on the ABS list in a segment to the estimated number of DUs in the segment.
4. Compare the expected ABS coverage rate to a predetermined coverage threshold. The ABS list supplemented with the CHUM procedure will serve as the frame for segments with coverage rates at or above the threshold. All other segments will be scheduled for FE supplemented with the HOI procedure.

A key aspect of the hybrid frame is the determination of a threshold of ABS coverage below which FE will be used. To this end, we assume *a priori* that FE and the HOI will be more efficient (from a cost as well as a statistical standpoint) than ABS and the CHUM in segments where the ABS coverage is less than 20 percent. Consequently, our evaluation of potential coverage thresholds presupposes that FE will be used in sampled segments with an expected ABS coverage rate below 20 percent.

We designed the field study to address three research questions:

1. What coverage threshold is needed to ensure that the hybrid frame provides comparable coverage of the current NSDUH sampling frame?
2. What coverage threshold is needed to produce comparable prevalence estimates between persons covered by the hybrid frame and those covered by the current frame?
3. What are the cost savings associated with the hybrid frame?

This paper describes the design, methods, and results of the field study that was conducted in April, May, and June 2009.

2. Methods

2.1 Sample Selection

The sample for the field study comprised 3,878 screened and eligible[1] sampled DUs in a subsample of 200 NSDUH segments. Segments in Alaska and Hawaii were excluded from the sampling frame. A total of 1,725 interviews were obtained from these sampled DUs in the first quarter of 2009. The use of segments already fielded for NSDUH offered two advantages:

1. The virtually complete coverage of the current NSDUH frame could be used to evaluate the coverage of the hybrid frame.

2. Additional interviews were not needed to compare prevalence rates associated with the hybrid frame with those of the NSDUH frame.

The field study took advantage of segments in which enumeration, screening, and interviewing had already occurred. This enabled us to simulate what would have happened if the ABS list had been utilized in segments above the estimated ABS coverage threshold, both in terms of field staff identifying DUs not on the ABS list and estimating prevalence rates if the hybrid sampling frame had been used.

Prior to selecting the sample for the field study, we stratified the NSDUH segments based on the expected coverage of city-style mailing addresses. The coverage estimate used for stratification was based on the ratio of the number of city-style mailing addresses to the 2007 Claritas housing unit projection. Stratification enabled us to estimate the practical implications of various coverage thresholds. We also defined a separate stratum for segments with a high percentage of group quarters (GQs) because college dormitories represent an important segment of the NSDUH target population and previous research indicates that the ABS list has limited coverage of college dormitories (Dohrmann et al. 2006).

Table 1 summarizes the sample stratification scheme. Based on our preliminary analysis, we expected that the mailing address coverage threshold would be set somewhere between 30 and 70 percent. We expected that all segments with less than 20 percent mailing address coverage would be field enumerated. Similarly, we assumed that all segments with greater than 80 percent mailing address coverage would be able to use the ABS list supplemented with the CHUM procedure. Therefore, we defined four sampling strata around these coverage thresholds and whether the segment had a high percentage (25 percent or more) of GQs. Stratum 1 consisted of all segments with less than 20 percent mailing address coverage, regardless of the GQ percentage. Stratum 2 included segments with estimated mailing address coverage between 20 and 80 percent and a low percentage of GQs. These segments were oversampled to obtain a reliable estimate of the resources needed for the CHUM procedure in segments with low expected coverage. If, for example, we found that the level of effort needed to implement the CHUM procedure

[1] An eligible DU for the NSDUH is either a housing unit (HU) for a single household or a non-institutional group quarters (GQ) where at least one civilian aged 12 years or older resides for the majority of a calendar quarter.

was excessive for segments with 60 percent ABS coverage, we would recommend that this be used as a threshold for the hybrid frame. Stratum 3 includes segments with greater than 80 percent mailing address coverage and a low percentage of GQs. The final stratum, Stratum 4, consists of all segments with 25 percent or more of the listed units being GQs and at least 20 percent mailing address coverage.

Table 1: Field Study Sample Segments

Stratum	Expected ABS Coverage	Group Quarter Percentage	Frame Count	Sample Count	Eligible Dwelling Units
1	< 20%	Any	48	13	167
2	20-80%	< 25%	178	90	1,670
3	> 80%	< 25%	643	90	1,923
4	≥ 20%	≥ 25%	7	7	118
Total			876	200	3,878

We conducted a sample optimization to determine the number of segments to be selected from each stratum. A slight departure from the optimal allocation was made to ensure adequate sample sizes in Strata 1, 2, and 4. The recommended sample distribution was 10, 90, 90, and 10 to Strata 1 through 4, respectively. Because there were fewer than 10 segments in Stratum 4 on the sample frame, the sample distribution was modified to 13, 90, 90, and 7. Table 2 displays the final sample distribution and the number of eligible DUs associated with the sample in each stratum.

Prior to selecting the field sample, the segment frame was sorted by urban/rural status and expected mailing address coverage. Tables 2 and 3 display the sample distribution by DU type (household unit or GQ) and segment type (urban or rural). The 200 sample segments contained 3,878 eligible DUs and 1,725 NSDUH interviews. The majority of the sample segments are urban.[2] Thus, almost 75 percent of the eligible DUs are in urban areas. The field sample included 149 eligible GQs, resulting in 162 interviews from persons residing in GQs.

Table 2: Distribution of Eligible Sampled Dwelling Units by Urban/Rural Status

Segment Type	Households		Group Quarters		Total	
	Count	Percent	Count	Percent	Count	Percent
Urban	2,744	73.6	147	98.7	2,891	74.5
Rural	985	26.4	2	1.3	987	25.5
Total	**3,729**	**100.0**	**149**	**100.0**	**3,878**	**100.0**

[2] For a segment to be classified as rural, all of the census blocks in the segment have to be rural. If one or more of the segment's blocks are urban, the segment is also urban.

Table 3: Distribution of Interview Respondents by Urban/Rural Status

Segment Type	Households		Group Quarters		Total	
	Count	Percent	Count	Percent	Count	Percent
Urban	1,160	74.2	162	100.0	1,322	76.6
Rural	403	25.8	0	0.0	403	23.4
Total	**1,563**	**100.0**	**162**	**100.0**	**1,725**	**100.0**

We prepared segment-, DU-, and person-level weights to generalize the results of the field study to the NSDUH population. The segment-level weight was defined as the inverse of the segment's probability of selection for the field study. The DU-level weight was computed as the segment-level weight times the NSDUH household weight. Finally, the person-level weight was computed as the segment weight times the NSDUH analysis weight[3]. For the hybrid frame, which is defined by a 20 percent coverage threshold, the person-level weights also were multiplied by a poststratification adjustment factor. The segment-level weight is used in the analysis of cost, the DU weight is used in the coverage analysis, and the person-level weight is used in the outcomes analysis. Table 4 displays weight distributions by stratum.

Table 4: Weight Distributions by Stratum

Stratum	Segment Weights		Eligible DU Weights		Respondent Weights	
	Sum	Percent	Sum	Percent	Sum	Percent
1	384	5.5	1,727,857	1.8	4,479,090	2.0
2	1,424	20.3	18,696,177	19.4	37,931,511	17.2
3	5,144	73.4	75,541,036	78.2	176,506,759	80.3
4	56	0.8	594,483	0.6	987,813	0.5
Total	**7,008**	**100.0**	**96,559,553**	**100.0**	**219,905,173**	**100.0**

DU = dwelling unit.

In addition to the weight variables, variance estimation strata and replicate variables were created to appropriately adjust for the sample design when computing variances of survey estimates. Within strata, sample segments were paired to form variance estimation strata. Because strata 1 and 4 contain an odd number of segments, two of the variance estimation strata have three replicates rather than two.

2.2 Matching Mailing Addresses to NSDUH DUs

To develop a hybrid frame of DUs, we attempted to match the street name and number, city, state, and ZIP code of eligible sampled DUs obtained from the NSDUH screening to a list of mailing addresses purchased from a commercial vendor. The initial matching of eligible, sampled DU addresses to the mailing list yielded 2,776 matches (71.6 percent). The weighted initial match rate was 78.5 percent. Among the 200 sampled segments, 41

[3] The NSDUH analysis weight includes all weight components, including nonresponse and poststratification adjustments at the household and person levels.

had a 100 percent match rate, and 11 had a 0 percent match rate. One segment did not contain any eligible sampled DUs. Sampled DUs whose mailing address did not initially match to the ABS list were followed up with a telephone or field check to verify or correct the mailing address of the DU. In the 187 segments with a 20 percent or higher expected ABS coverage rate, corrected addresses were obtained for 440 sampled DUs, which increased the ABS match rate to 86.4 percent.

The expected ABS coverage was calculated for each sampled segment as the ratio of the estimated number of city-style mailing addresses to the estimated number of DUs. Using this method, only 15.4 percent of the NSDUH segments and 10.2 percent of the eligible sampled DUs would be subject to FE with a 50 percent ABS coverage threshold. We calculated a correlation of 0.66 between the expected coverage rate and the observed rate (defined as the number of matched addresses divided by the number of sampled DUs times 100).

2.3 Implementation of the CHUM Frame-Linking Procedure

The CHUM procedure is designed to be used in segments where ABS coverage is adequate but not complete. As Figure 1 demonstrates, the CHUM procedure uses two components to supplement the coverage of an ABS frame. The CHUM1 component enables FIs to establish a *path of travel* from the sampled DU to the "next DU." This methodology requires segments with discernible boundaries so that DUs can be identified as being within a sampled segment. Facing the located sampled DU, the FI travels clockwise around the block, without crossing a street to find the next DU. Street crossings are avoided to ensure that each path of travel is non-overlapping.

After the address of the next DU is found, it is checked against the ABS frame to determine whether it was missed. This check requires access to the entire ABS frame within each selected segment. If the address of the next DU is not on the ABS frame, the DU is sampled and assigned a probability of selection that is linked to the probability of selecting the sampled DU. If multiple missed DUs are picked up by the CHUM1, then subsampling may be used to control the sample size. Otherwise, a missed DU will be included in the sample and assigned the same design weight as the sampled DU. This step is repeated until either the address of a DU on the ABS frame is found or the block is circumnavigated.

Because CHUM1 is restricted to blocks associated with a sampled address, DUs in blocks with no addresses on the ABS frame will be missed. CHUM2 mitigates this source of undercoverage by adding the "missed blocks" and their associated DUs to the frame. For segments that contain multiple missed blocks, efficient sampling methods can be used to select a subsample of missed DUs.

For the field study, the CHUM procedure was applied to the 505 nonmatching sampled DU addresses to estimate the gain in coverage afforded by this portion of the hybrid frame methodology. However, it was necessary to modify the CHUM procedure so that the coverage of NSDUH sampled DUs missed by the ABS frame could be evaluated. To evaluate the CHUM procedure, an ABS address was selected in the vicinity of the missed NSDUH DU and then treated as a sampled DU for purposes of implementing the CHUM procedure. If the missed NSDUH DU was picked up by CHUM, it was considered covered by the hybrid frame.

DUs that are not included on the ABS list and are in segments that are above the expected coverage threshold rely on the CHUM procedures for coverage. This study utilized two processes (CHUM1 and CHUM2) to make sure that DUs not on the ABS had a chance to

be included in NSDUH. Prior to fieldwork, we examined NSDUH listing maps for each sampled DU that did not match to the ABS list either initially or after the phone verification process. We determined whether each sampled DU would be covered by the Phase I CHUM procedure (CHUM1) or whether it would rely on the Phase II CHUM procedure (CHUM2) for coverage.

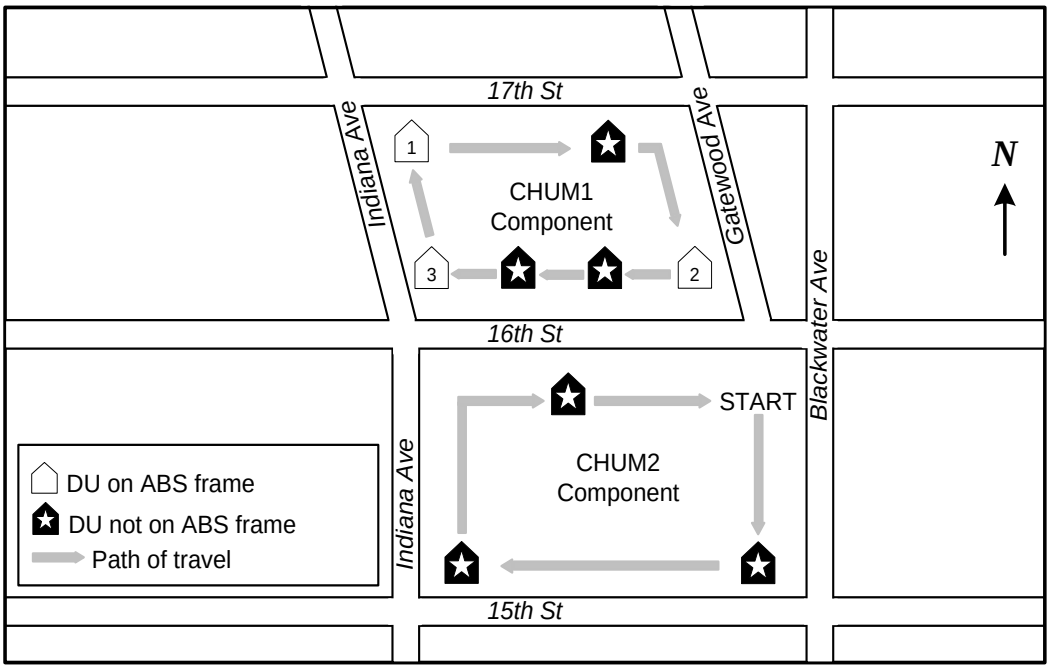

Figure 1: Examples of CHUM Components

CHUM1 Component: Path of travel moves clockwise from the selected DU. If DU 1 is selected, one missed DU would be added to the ABS frame. If DU 2 is selected, two missed DUs would be added. No missed DUs would be added if DU 3 is selected. **CHUM2 Component**: First, a block or area is randomly selected within a segment. Then, FIs are provided a segment map that enables the entire block to be field enumerated. In this case, three missed DUs are added to the frame. Notice that the three missed DUs have no chance for inclusion solely using CHUM1.

In the CHUM1 procedure, FIs start at an address that is contained on the ABS list (the starting DU) and follow a predetermined path of travel (usually moving to the left of the starting DU) until they either reach an address that is contained on the ABS list or they return to the starting DU. This allows coverage of DUs that are not on the ABS list but are in the same block as addresses on the ABS list.

In the CHUM2 procedure, FIs perform the CHUM procedure from a predetermined start point in a randomly selected area rather than a starting DU. The FIs follow the same path of travel that they do for the CHUM1 procedure, stopping when they either list an address that matches to the ABS list or they return to the start point without finding a match. This procedure enables coverage of DUs that are in blocks where none of the DUs are on the ABS list.

The hybrid frame, utilizing a combination of FE, the ABS list, and the CHUM1 and CHUM2 procedures, theoretically provides 100 percent coverage of the target population.

However, as with any frame supplementation method, a loss of coverage occurs when FIs neglect to follow the correct protocols when implementing the CHUM procedures. In an attempt to incorporate the loss of coverage because of the human error inherent in performing the CHUM procedures, we based coverage of NSDUH DUs relying on the CHUM procedures on the field results rather than the theoretical coverage that we determined from the listing maps. We expect this coverage estimate to represent a "worst-case scenario" for the coverage that the CHUM procedures can provide because FIs received less training than they typically would, were not monitored in the field, were performing multiple tasks while in the field, and were encouraged to perform all fieldwork in one trip to the segment.

2.4 Cost Estimates

The costs associated with the FE portion of the hybrid frame included segment editing, field supplies, shipping of segment kits, lister training, lister labor, mileage reimbursement, miscellaneous expenses incurred in the field, field supervisor support, survey specialist support, keying of listing data, and implementation of the half-open interval (HOI) procedure during screening and interviewing. The per segment cost of supplies, training, editing, shipping, field supervisor labor, survey specialist labor, keying, and implementation of the HOI were assumed to be the same whether a segment was urban or rural. The cost difference between urban and rural segments was in lister labor hours and mileage. Miscellaneous expenses are difficult to predict prior to the lister going to the field, so the average miscellaneous expenses per segment from the 2009 NSDUH FE effort were used to determine the per segment average.

Most estimates related to the ABS portion of the hybrid frame, particularly those for CHUM-related activities, come from RTI's experience implementing an ABS sample design for the 2008 American National Election Study (ANES) (Lupia et al. 2009). The 2008 ANES was a nationally representative sample of 5,032 addresses and 297 segments that corresponded to census block groups. Because of the size difference between NSDUH and ANES, estimates needed to be substantially inflated and do not take into account economies of scale obtained by mass production. We tried to be conservative (erring on the high side) when scaling labor projections up to the NSDUH. The per segment cost components associated with implementing the ABS and CHUM portions of the hybrid frame included the purchase of the address lists, programming and field costs related to the CHUM, and other miscellaneous expenses such as segment map production.

3. Answers to Research Questions

Research question 1: ***What coverage threshold is needed to ensure that the hybrid frame provides comparable coverage of the current NSDUH sampling frame?***

Table 5 presents the estimated DU coverage for each component of the hybrid frame. Because we assume the FE frame component has 100 percent coverage of the NSDUH sampling frame, coverage estimates increase as the coverage threshold increases because more of the segments require FE. Sampled DUs in segments determined not to need FE will rely on ABS and the CHUM.

Table 5: Estimated Coverage of NSDUH Dwelling Units of the Hybrid Frame

ABS Coverage Threshold = 20%

	ABS	CHUM1	CHUM2	FE	Overall
Urban	93.9%	2.5%	0.8%	1.8%	**98.9%**
Rural	71.8%	13.4%	5.3%	1.8%	**92.3%**
Overall	89.3%	4.8%	1.7%	1.8%	**97.5%**

ABS Coverage Threshold = 50%

	ABS	CHUM1	CHUM2	FE	Overall
Urban	87.3%	2.1%	0.7%	8.8%	**99.0%**
Rural	63.9%	10.8%	3.2%	15.4%	**93.4%**
Overall	82.5%	3.9%	1.2%	10.2%	**97.8%**

ABS Coverage Threshold = 80%

	ABS	CHUM1	CHUM2	FE	Overall
Urban	70.1%	1.4%	0.4%	27.2%	**99.1%**
Rural	34.9%	7.6%	2.9%	50.2%	**95.6%**
Overall	62.7%	2.7%	0.9%	32.0%	**98.4%**

ABS = address-based sampling; FE = field enumeration.

CHUM1 coverage estimates reflect how accurately the method was implemented because perfect execution would yield total coverage. This estimate should be considered a lower bound for how well the CHUM covers DUs because the field study did not adequately represent how the CHUM procedure could be implemented in the field and therefore negatively affects the coverage estimates of the CHUM. Reasons for the negative bias include the following:

- CHUM performance was not monitored for data quality during data collection. Typically, methods are used to measure the quality of the CHUM implementation in an attempt to identify and then rectify mistakes in the field.
- Field staff received less training than they would receive for an operational implementation.
- Field staff were asked to perform multiple tasks using multiple forms. This led to some confusion on what their task was regarding the CHUM.

- Field staff had limited field time to perform the CHUM procedure. Typically, the CHUM could be performed on return visits to the segment while the FI waited for clarification from sampling support staff.

CHUM2 coverage estimates also reflect how accurately the method was implemented because perfect execution would yield total coverage. For similar reasons described above, the CHUM2 is also an underestimate.

How closely the estimated ABS coverage threshold comes to predicting the actual ABS coverage will have enormous influence on both cost and coverage. If we are able to accurately predict ABS coverage, low coverage segments would always be field enumerated, and fewer segments with marginal ABS coverage would rely on the CHUM. The ability to accurately predict ABS coverage is directly related to higher coverage and lower costs. To illustrate this potential, let us assume we could accurately predict the ABS coverage. If actual ABS coverage of a segment falls below 80 percent, we would use FE; otherwise, we would use ABS and CHUM. Under these assumptions and using data from this study, only 19 percent of NSDUH segments would require FE compared to 37 percent using the current method. More impressively, both rural and urban segments would achieve over 99 percent coverage using the hybrid frame methodology. Additional research is being conducted on determining the best way to predict ABS coverage *a priori* (McMichael et al. 2010).

Research question 2: What coverage threshold is needed to produce comparable prevalence estimates between persons covered by the hybrid frame and those covered by the current frame?

A total of 1,699 of the 1,725 persons were covered by the hybrid frame at the 20 percent ABS coverage threshold. The 26 persons not covered were more heavily concentrated in rural areas and were more likely to be older and non-Hispanic white than the remainder of the NSDUH frame. At ABS coverage thresholds of 50 and 80 percent, only 23 and 14 persons were not covered by the hybrid frame, respectively.

We compared prevalence estimates for a number of key NSDUH outcomes based on the current FE frame with those based on the hybrid frame. None of the differences were substantively different at the 20 percent ABS coverage threshold. Because the estimates based on the NSDUH frame and the estimates based on the hybrid frame share a large portion of their cases, these comparisons have the statistical power to declare very small differences in the overall prevalence estimates statistically significant (e.g., 0.0002). For more information on the differences in prevalence estimates, see Morton et al. (2010).

Research question 3: What are the cost savings associated with the hybrid frame?

The cost estimates presented in Table 6 assume that the cost of selecting the sample of 7,200 segments is the same whether the current NSDUH frame is used or the hybrid frame is used. Because Alaska and Hawaii have atypical listing costs, only the 7,008 segments that fall in the continental United States are considered in the cost comparisons. Note that the cost estimates include materials development and production, field labor and expenses, editing, and indirect fees.

Table 6: Estimated Annual Cost Savings[1] of the Hybrid Frame

ABS Coverage Threshold[2]	FE Segments		ABS Segments		Estimated Cost Savings[3]
	#	%	#	%	
20%	384	5.5%	6,624	94.5%	62.3%
30%	592	8.4%	6,416	91.6%	60.2%
40%	789	11.3%	6,219	88.7%	58.3%
50%	1,076	15.4%	5,932	84.6%	55.4%
60%	1,353	19.3%	5,655	80.7%	52.7%
70%	1,632	23.3%	5,376	76.7%	50.0%
80%	2,565	36.6%	4,443	63.4%	40.1%
FE Only	7,008	100.0%	0	0.0%	0.0%

ABS = address-based sampling; FE = field enumeration.

[1] Cost savings include costs incurred after the selection of the segment and before the start of field interviewing. Costs include the half-open interval (HOI) for FE segments and the Check for Housing Units Missed (CHUM) for ABS segments. Excluded are the costs associated with the 192 NSDUH segments in Alaska and Hawaii.

[2] Ratio of the number of city-style mailing addresses in a segment to the Census Bureau/Claritas dwelling unit (DU) estimate.

[3] Compared with doing FE in all 7,008 segments.

4. Considerations

Choice of an ABS Coverage Threshold. The scope of the field study was limited to national estimates of coverage, cost, and study outcomes. If the hybrid frame becomes operational and a coverage threshold decided upon, the threshold will need to be re-evaluated periodically to ensure that full coverage is retained with maximal cost savings. It should be expected that states will have different fractions of segments handled with ABS no matter what threshold is adopted.

Segment Size. Field enumeration (FE) restricts the current NSDUH frame to one or more census blocks. The hybrid frame would enable segments above the ABS coverage threshold to be as large as census block groups or even census tracts. The current segment size would be retained for segments below the threshold. Large segments will increase the accuracy of the Claritas dwelling unit (DU) estimates and decrease the frequency of geocoding error. Minimizing geocoding error will decrease reliance on the CHUM procedures by increasing ABS coverage and will decrease inefficiencies caused by over-coverage (Shook-Sa et al. 2010). Also, large segments may enable a reduction in the number of segments needed to support the NSDUH design objectives at the national level but could have an impact on the precision of estimates in small geographic areas if they are made too large.

Sampling Frame Updates. Currently, 7,200 segments are selected for FE during the initial survey year and then 3,600 segments are selected in each of the following four

years so that each FE segment gets used twice. The economy of the hybrid frame would enable yearly frame updates.

5. Next Steps

1. The gains in coverage afforded by the Check for Housing Units Missed (CHUM) procedure are essential to the veracity of the hybrid frame. Like the half-open interval (HOI) procedure, the CHUM offers virtually complete coverage of DUs if implemented correctly. For the field study, it was necessary to retrofit the CHUM procedure so that the coverage of sampled dwelling units DUs not on the ABS frame could be evaluated. This atypical application of the CHUM confounds any conclusions regarding its effectiveness. Therefore, a separate study is underway to determine the protocol and training requirements needed to ensure that the CHUM is a viable part of the hybrid sampling frame.

2. An examination of the relationship between initial and observed segment coverage is needed. For example, it may be that the initial coverage tracks differently in urban segments than in rural ones. Additional research is underway to improve the accuracy of ABS coverage predictions before sample selection (McMichael et al. 2010).

3. The coverage of the hybrid frame can be increased with the inclusion of two supplemental lists: (1) a listing of college dormitories, which constitute the vast majority of GQs; and (2) a listing of HUs that are located in areas without home mail delivery. More research is needed to determine the accuracy, availability, and cost of these lists.

Acknowledgement

This project is funded by the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract no. 283-2004-00022 and Project no. 0209009. The views expressed in this paper do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government

References

Dohrmann, S., D. Han, and L. Mohadjer. 2006. Residential address lists vs. traditional listing: Enumerating households and group quarters. In *JSM Proceedings,* Survey Research Methods Section, Alexandria, VA: American Statistical Association. 2959-2964.

Kish, L. (1965). *Survey Sampling* (p. 56). New York: John Wiley & Sons.

Iannacchione, V., K. Morton, J. McMichael, D. Cunningham, J. Cajka, and J. Chromy. 2007. Comparing the coverage of a household sampling frame based on mailing addresses to a frame based on field enumeration. In *JSM Proceedings,* Survey

Research Methods Section, Alexandria, VA: American Statistical Association. 3323-3332.

Lupia, A., J. Krosnick, P. Luevano, M. DeBell, and D. Donakowski. 2009. "User's Guide to the ANES 2008 Time Series Study. Ann Arbor, MI and Palo Alto, CA: the University of Michigan and Stanford University." Available at: http://www.electionstudies.org/studypages/2008prepost/2008prepost_UsersGuide.pdf

McMichael, J.P., J.L. Ridenhour, and B.E. Shook-Sa. 2008. A robust procedure to supplement the coverage of address-based sampling frames for household surveys. In *JSM Proceedings*, Survey Research Methods Section, Alexandria, VA: American Statistical Association. 4329-4335.

McMichael, J.P., J.L. Ridenhour, B.E. Shook-Sa, K.B. Morton, and V.G. Iannacchione. 2010. Predicting the coverage of address-based sampling frames prior to sample selection. In *JSM Proceedings*, Survey Research Methods Section, Alexandria, VA: American Statistical Association.

Morton, K.B., J.P. McMichael, J.L. Ridenhour, and J. Bose. 2010. Address-based sampling and the National Survey of Drug Use and Health: evaluating the effects of coverage bias. In *JSM Proceedings*, Survey Research Methods Section, Alexandria, VA: American Statistical Association.

O'Muircheartaigh, C., E. English, and S. Eckman. 2007. Predicting the relative quality of alternative sampling frames. In *JSM Proceedings*, Survey Research Methods Section, Alexandria, VA: American Statistical Association. 3239-3248.

Shook-Sa, B.E., J.P. McMichael, J.L. Ridenhour, and V.G. Iannacchione. 2010. The implications of geocoding error on address-based sampling. In *JSM Proceedings*, Survey Research Methods Section, Alexandria, VA: American Statistical Association.