# A Comparative Evaluation of Traditional Listing vs. Address-Based Sampling Frames: Matching with Field Investigation of Discrepancies

Jill M. Montaquila[1,2], Valerie Hsu[1], J. Michael Brick[1,2]
Ned English[3], Colm O'Muircheartaigh[3]
[1] Westat, 1600 Research Blvd., Rockville, MD 20850
[2] The Joint Program in Survey Methodology, University of Maryland/University of Michigan, 1218 LeFrak Hall, College Park, MD 20742
[3] NORC, University of Chicago, 55 E. Monroe, Chicago, IL 60603

## Abstract

Increasingly, survey practitioners are considering the use of address-based sampling frames for household surveys. These frames are being studied as alternatives to random digit dial (RDD) frames, to facilitate the use of mixed-mode methods, and to supplement or replace traditional listing of addresses. The recent consideration of these frames has been made possible by the availability of files based on the United States Postal Service (USPS) Delivery Sequence File (DSF) and changes in addressing, particularly those resulting from the development of Enhanced-911 systems.

Previous studies have, collectively, yielded an abundance of information about the utility of address-based sampling frames, and it has been critical to continue evaluating these frames in light of the ongoing changes in addressing. This paper extends the earlier results by evaluating the USPS-based address frames in seven primary sampling units in the National Children's Study. Our evaluation is both more current and it includes comparisons of addresses listed through traditional listing to those on two different USPS-based frames. In a few select areas, this comparison is enhanced by a field effort that goes back to the sampled areas to better understand and characterize discrepancies or nonmatches. This evaluation will help establish the relative merits of traditional listing and USPS-based address frames, and whether particular kinds of places are more likely to be undercovered by each method.

**Key Words:** Delivery Sequence File, coverage

## 1. Introduction

Increasingly, survey practitioners are considering the use of address-based sampling frames for household surveys. These frames are being studied as alternatives to random digit dial (RDD) frames, in light of recent declines in RDD survey response rates (Curtin, Presser, and Singer, 2005; Battaglia et al. 2008; and Holbrook, Krosnick, and Pfent, 2008) and coverage rates (Blumberg and Luke 2009). Additionally, address-based samples can be used to facilitate the use of mixed-mode methods (Link et al. 2008). With the ability to match telephone numbers to addresses, primary contact with sampled cases in an address-based sample may be in-person, by mail, or by telephone (for the subset of

matched cases), and other modes (such as web and interactive voice response, or IVR) may be used for follow-up. A third use of address-based frames is to supplement or replace traditional listing of addresses (O'Muircheartaigh, Eckman, and Weiss 2003; O'Muircheartaigh et al. 2006; Iannacchione, Staab, and Redden 2003).

The recent consideration of these frames has been made possible by the availability of files based on the United States Postal Service (USPS) Delivery Sequence File (DSF) and changes in addressing, particularly those resulting from the development of Enhanced-911 systems. The DSF-based files are available through vendors who maintain these lists of addresses and often offer other services or enhancements.

Previous studies have, collectively, yielded an abundance of information about the utility of address-based sampling frames, and it has been critical to continue evaluating these frames in light of the ongoing changes in addressing. This paper extends the earlier results by evaluating the USPS-based address frames in seven primary sampling units (PSUs) in the National Children's Study (NCS). Our evaluation is both more current and it includes comparisons of addresses listed through traditional listing to those on address-based frames.

## 2. Summary of Previous Research on the Utility of Address-based Sampling Frames

Research has been undertaken to evaluate the use of address-based sampling frames in a variety of contexts. An early study by Iannachhione, Staab, and Redden (2003) was a coverage evaluation of USPS-based lists in Dallas, Texas. In this study, the authors used the USPS-based lists as a sampling frame and evaluated coverage using the half-open interval procedure (Kish 1965). They reported an estimated undercoverage rate of 1.9 percent, and also found evidence that in Dallas, the majority of households with P.O. box addresses also receive mail at their street addresses, and that the occupancy rate was consistent with rates generally found for listed housing units in studies of metropolitan areas. Although this study was restricted to a single metropolitan area, its results indicated that there may be potential for use of USPS-based address lists as sampling frames.

O'Muircheartaigh, Eckman, and Weiss (2003) compared traditional listing and "enhanced field listing." With enhanced listing, the listers were given the USPS addresses (geocoded so that they could appear in geographic sequence), with instructions to make corrections and note addresses that were missing from the USPS-based lists. The result, the enhanced listings, were then compared to traditional listings (generated by a separate group of listers) and to the original USPS-based address list (before enhancement). The authors observed problems with using the USPS-based lists in rural areas due to the high prevalence at that time (2001) of rural route addresses and P.O. boxes. In non-rural areas, however, they noted the superiority of enhanced listing to traditional listing, and also concluded that this study demonstrated potential for the use of USPS-based lists (without enhancement) in place of traditional listing.

O'Muircheartaigh et al. (2006) used field verification to arrive at a "best" address frame. As the basis for their evaluation, they used a set of area segments that had been traditionally listed, and obtained USPS-based address lists for these segments. The combination of these two lists served as the basis for the field verification effort. The

authors concluded that, overall, the USPS-based list is superior to traditional listing. They also offered a set of criteria for identifying segments where traditional listing is likely to be superior: those with irregular street patterns that are more susceptible to geocoding errors; those with counts of addresses from the USPS-based list that are substantially below the decennial census counts; and those with high rates of growth in population.

In contrast to the aforementioned evaluations that examined the potential use of USPS-based address lists in place of traditional listing in multi-stage area probability samples, Link et al. (2008) considered USPS-based lists as alternative to RDD frames. Their evaluation, conducted in six states, encompassed several aspects, and they concluded that "While the DSF appears to be an effective frame for conducting address-based sampling of the general population, its true potential may be in facilitating mixed-mode surveys." (Link et al. 2008, p. 26.) The authors also commented on the need for further evaluation of address-based sampling.

## 3. Evaluation

### 3.1 Methods

The NCS, sponsored by a consortium of agencies from the U.S. Department of Health and Human Services and the U.S. Environmental Protection Agency, is designed to examine the effects of environmental influences on the health and development of more than 100,000 children across the U.S., following them from before birth until age 21. The sample design for NCS is a multi-stage area probability household sample (Montaquila, Brick, and Curtin forthcoming). The first stage of sampling is the selection of PSUs, which are single counties or groups of contiguous counties. The second sampling stage for large and densely populated PSUs is the selection of geographic areas within the PSU. The third stage for large urban PSUs and the second stage for small PSUs is the selection of clusters of contiguous census blocks within the geographic areas or PSUs, called "segments" (with a segment typically comprising approximately 500 to 1,200 households). Within the sampled segments, household enumeration is attempted in all dwelling units* (DUs) and essentially all births that occur during the enrollment period are eligible for the Study.

One of the early steps in the household-based data collection effort for the NCS is the preparation of lists of all residential addresses in each sampled segment. This evaluation is based on listing conducted for a Pilot Study in seven PSUs:

- DC: Duplin County, NC
- BYPL: Brookings County, SD; Yellow Medicine, Pipestone, and Lincoln Counties, MN
- WC: Waukesha County, WI
- MC: Montgomery County, PA
- SLC: Salt Lake County, UT
- OC: Orange County, CA
- QC: Queens County, NY

* In very densely populated areas the sample segment may contain too many households and may be subsampled ("chunked") to achieve the appropriate sample size.

In area probability samples like the NCS, the compilation of residential addresses has historically been done using traditional listing. Trained listers canvass the sampled segment, locate segment boundaries, and then compile a hard-copy list of residential addresses as they move in a systematic manner through the segment. For these seven NCS PSUs, the process of listing in the sampled segments was conducted by different organizations in each PSU using the traditional paper-and-pencil approach. Listings were prepared in 2008 for the seven PSUs prior to the launch of the data collection in 2009.

Since all DUs in the sampled segments are eligible, and meeting the birth targets for NCS depends on obtaining complete coverage, it is of critical importance to ensure good coverage of the address lists. Following the listing operation, quality control checks were done on the listed addresses and, as part of this process, the listed addresses were compared to commercially available address lists originating from the USPS.

Lists of residential addresses in the ZIP codes associated with the sampled segments were obtained from a vendor. These addresses were geocoded to determine the associated geographic identifiders, in particular, census tract and block. Addresses that geocoded to within sampled segments were retained; all non-geocodable addresses and addresses that geocoded to a location outside the segments were dropped. Note that non-city-style addresses (e.g., Post Office boxes and rural route addresses) were treated as non-geocodable. To provide a general idea of the coverage of the USPS-based address lists, we compared counts of the numbers of residential units obtained from the USPS-based address lists to counts of housing units obtained from listers through the traditional listing approach and from the decennial census.

Table 1 shows ratios that can be used to examine aspects of the quality of the USPS-based frames. The first two columns contain data at the segment level, while the third column presents ratios at the PSU level as an indicator of growth. The numbers of residential USPS addresses that geocoded into the sampled segments compared favorably to the listing counts and census counts for the most part. The comparisons between the counts did reveal possible coverage insufficiency in the USPS-based lists of retained addresses, especially in rural regions. The USPS-based lists generally provided good coverage in urban areas, but poorer coverage in high-growth areas with new housing under construction in several developments.

We attempted to match each paper-listed address with an address on the USPS address list using an automated matching program followed by manual matching. The retained USPS-based and paper-listed addresses were merged together by all address fields except for city and ZIP code (i.e., street number, street name, street suffix, pre- and post-direction, unit designator, unit number, and 2-letter state abbreviation), as listers might not have been able to acquire accurate city and ZIP code information while in the field. Any unmatched addresses resulting from the automated matching were investigated manually to resolve minor discrepancies, such as differences in spelling or typos (e.g., “Weatherby Rd.” vs. “Wetherby Rd.”), differences in street type (e.g., “Oak St.” vs. “Oak Ln.”), and “No number” addresses (e.g., matching a “no number” address listed between 123 Main St. and 127 Main St. with a “125 Main St.” listing on the USPS-based list).

**Table 1:** Comparison of Address Counts from the USPS DSF to Counts of Traditionally Listed Addresses and Counts of Housing Units (HUs) from Census 2000

| *PSU* | *Across segments: range of #USPS addresses that geocoded into the sampled segment/# traditional listed addresses* | *Across segments: range of # USPS addresses that geocoded into the sampled segment / # housing units from Census 2000* | *PSU level: # HUs from ACS 2007† / # HUs from Census 2000* |
|---|---|---|---|
| DC | 0.35 – 0.91 | 0.39 – 0.90 | 1.04* |
| BYPL | 0.26 – 1.30 | 0.26 – 1.07 | 1.06‡ |
| WC | 0.75 – 1.09 | 0.80 – 2.15 | 1.11 |
| MC | 0.71 – 1.09 | 0.83 – 1.11 | 1.05 |
| SLC | 0.75 – 1.02 | 0.92 – 8.26 | 1.13 |
| OC | 0.72 – 1.01 | 0.75 – 5.22 | 1.06 |
| QC | 0.80 – 1.27 | 0.80 – 1.21 | 1.02 |

†2007 American Community Survey 1-Year estimates of housing units
*2005-2007 American Community Survey 3-Year estimate of housing units was used as the numerator
‡Census Bureau estimates of housing units as of July 1, 2007 was used as the numerator

Each USPS-based address without a match was investigated using Google Earth, the annotated segment maps, and the listing sheets to determine whether (1) it was equivalent to a paper-listed address; (2) it was outside the segment but inadvertently retained due to geocoding error; or (3) it was missed by the listers. In some cases, we were able to positively match many addresses that did not match through the automated matching by inferring from the comments that the lister recorded. For example, we were able to match a paper-listed address that had no street number and a comment that said "between 102 First Street and 106 First Street" to a USPS listing that had an address at 104 First Street.

In some cases, we were able to correct keying or listing errors through the USPS comparison, thereby identifying additional matches. These errors were confirmed by reviewing other house numbers on the street in a series as well as searching on the USPS website for the address in question to see whether the website returns an error message that the particular address is non-deliverable. In some rural segments with many dwelling units that had no house numbers, the USPS-based listings were not as useful for the quality control of the paper-listed addresses as they were in urban sites. Google Earth as well as the segment map was used to confirm that the USPS listings included addresses that were not within the segment boundaries due to geocoding errors.

As a result of the manual comparison, a few blocks/apartment complexes that were missed completely by the listers were identified. Listers were sent out to relist the block(s) in question. Two segments (92 additional addresses) in BYPL, one segment (12 additional addresses) in WC, one segment (42 additional addresses) in OC, and two segments (70 additional addresses) in MC required partial relisting.

## 3.2 Results

The result of the matching described in section 3.1 was an "augmented traditional listing list"; this list contained the addresses listed by the traditional listers, less any addresses identified through the quality control process to have been listed in error, with address corrections and updates as described above.

Table 2 presents the results of matching. We calculated the match rate as the ratio of the number of addresses on both lists to the number of address on the augmented traditional listing list. In urban PSUs, only a small percentage of the paper-listed addresses could not be matched to a USPS address. The high match rate in urban regions is important in verifying both that the listers successfully captured most of the potential residential addresses, and that the USPS list contained a low percentage of non-existent or incomplete addresses, or addresses outside the segment. The low match rates in DC and BYPL are partly attributable to the fact that many trailers with no unit numbers in could not be matched to addresses on the USPS list. It is also worth noting that substantial variation in match rates at the segment level was observed across sites regardless of urbanicity (e.g., match rates ranging from 21% to 92% in BYPL, and from 72% to 100% in OC).

**Table 2:** PSU-Level Matching Results

| *PSU* | *Urbanicity (%)* | *Match rate (%)* | *% of matches obtained through manual matching* | *% of USPS-based addresses with nonmatches* |
|---|---|---|---|---|
| DC | 14 | 50 | 17 | 23 |
| BYPL | 44 | 54 | 25 | 13 |
| WC | 88 | 91 | 11 | 5 |
| MC | 97 | 86 | 13 | 6 |
| SLC | 99 | 92 | 6 | 3 |
| OC | 100 | 96 | 6 | 1 |
| QC | 100 | 94 | 34 | 2 |

In general, the percentage of matches obtained through the manual matching process was much higher in rural PSUs than in highly established urban PSUs. However, among the seven PSUs, QC had the highest percentage (34%) of matches obtained through manual matching due to the fact that listers recorded many addresses in a different format than those on the USPS lists (e.g., “123 45th Rd. Apt 1” vs. “123 45th Rd. Unit 1”). This evaluation also demonstrated that the percentage of USPS addresses with nonmatches was much higher in rural PSUs than in urban PSUs.

Table 3 contains statistics on a few aspects of the addresses that affect their utility in an address-based sampling frame, particularly the ability to geocode the addresses and the presence of multiple-drop addresses. As shown in Table 3, nongeocodable rates (i.e., the proportion of addresses to which a census tract and block could not be associated) were much lower in urban PSUs as compared to rural PSUs. Nongeocodables include non-city-style addresses (e.g., Post Office boxes and rural route addresses) as well as city-style addresses that could not be placed in a specific block using the information available in the geocoding databases. In some instances, the mail carrier delivers the mail at an address for subsequent distribution. A drop is a single delivery point or receptacle that services multiple residences. Examples of drop sites include a box on a wall for duplexes or gated communities where mail for all homes is delivered to a gatehouse. Only one record is kept on the USPS address lists for each address even if it represented multiple drops, but the number of drops to an address is provided on the commercially available address lists originating from the USPS. The percentage of USPS-based addresses that are multi-drops was highest in QC (10%) and negligible in other PSUs. For 4% of the paper-listed addresses in QC, the listers listed more secondary units for a particular

address than was on the USPS lists. It is possible that many single family homes were converted into multiple dwelling units and the USPS-based listings failed to capture those.

**Table 3:** Nongeocodables and Multi-Drops

| *PSU* | *Urbanicity (%)* | *% of USPS-based addresses that are nongeocodable* | *% of USPS-based addresses that are multi-drops* |
|---|---|---|---|
| DC | 14 | 18 | 0.10 |
| BYPL | 44 | 25 | 0.05 |
| WC | 88 | 5 | 0.14 |
| MC | 97 | 4 | 0.26 |
| SLC | 99 | 7 | 0.03 |
| OC | 100 | 2 | 0.03 |
| QC | 100 | <1 | 10.42 |

In addition to the evaluation described above, a field research study was conducted in WC to further evaluate alternative listing strategies for the NCS (English, O'Muircheartaigh, and Eckman forthcoming). The evaluation included comparisons of addresses listed through traditional listing to those on two different USPS-based frames (obtained from two different vendors), followed by fieldwork. An electronic comparison of the three sets of addresses was supplemented by inspection of nonmatches to correct for minor formatting and classification errors. Once the comparison was made, field staff were sent into the sampled segments to check the remaining discrepancies (i.e., addresses that did not appear on all three lists). The validation exercise did not involve interaction with potential respondents. Field staff took pictures of dwelling units that were not on all three lists and collected observational data about these addresses. Field staff also paid a visit to a set of blocks (specifically designated based on their match percentages) to collect data about the blocks using a special block face questionnaire.

The field research conducted in WC was designed to help to establish whether there are particular places where a discrepancy is more likely and in which direction. As a result of the comparison of the three frames and the field investigation, a 'best' address list was created. Compared to the augmented address list resulting from Westat's QC effort and additions from the first vendor, the number of addresses on the 'best' address list increased by about 1 percent as a result of using the second vendor and additions from the field effort, and was decreased by about 4 percent as a result of the field effort.

## 4. Discussion

The evaluation described here is a further step in the ongoing effort within the survey community to assess the feasibility of using USPS-based lists as sampling frames. In general, the findings reported here are consistent with the findings from previous studies: At the PSU level, coverage of the USPS-based address lists is generally good in urban areas but may be inadequate in some rural areas, and coverage is poorer in high-population-growth areas. However, it is worth noting that there is considerable variation in the coverage of the address lists within PSUs. Thus, it is possible (and, in fact, anticipated) that in many cases, the USPS-based lists would be adequate for some segments within a PSU but would be considerably inferior to traditional listing for other segments within the same PSU.

We have described a few issues with the USPS-based address lists that should be addressed when using these lists as sampling frames: multi-drop addresses and geocoding errors. Additionally, as with traditional listing, coverage enhancement procedures (such as the half-open interval procedure) may be used to improve the coverage of these lists.

As noted, this evaluation was conducted as part of the NCS Pilot Test. Once enumeration of households in the NCS Pilot Test has been completed, enumeration results and study data can be used to further evaluate these address lists—in particular, to examine how addresses and households that appear on one list but not the other (and households that were missed altogether by both the traditional listing process and the USPS-based lists) differ from those that appear on both lists.

## References

Battaglia, M.P., Khare, M., Frankel, M.R., Murray, M.C., Buckley, P., and Peritz, S. 2008. Response rates: How have they changed and where are they headed? Chapter 24 in J. M. Lepkowski, C. Tucker, J. M. Brick, et al. (eds.), *Advances In Telephone Survey Methodology,* New York: Wiley.

Curtin, R., Presser, S., and Singer, E. 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69(1), 87-98.

English, N., O'Muircheartaigh, C., and Eckman, S. (forthcoming). Coverage rates and coverage bias in housing unit frames. In *JSM Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Holbrook, A., Krosnick, J., and Pfent, A. 2008. The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. Chapter 23 in J. M. Lepkowski, C. Tucker, J. M. Brick, et al. (eds.), *Advances In Telephone Survey Methodology.* New York: Wiley.

Iannacchione, V.G., Staab, J.M., and Redden, D.T. 2003. Evaluating the use of residential mailing addresses in a metropolitan household survey. *Public Opinion Quarterly*, 67, 202-210.

Link, M.W., Battaglia, M.P., Frankel, M.R., Osborn, L., and Mokdad, A.H. 2008. A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys. *Public Opinion Quarterly,* 72, 6-27.

Montaquila, J.M., Brick, J.M., and Curtin, L.R. (forthcoming). Statistical and practical issues in the design of a national probability sample of births for the National Children's Study. *Statistics in Medicine*.

O'Muircheartaigh, C., Eckman, S., and Weiss, C. 2003. Traditional and enhanced field listing for probability sampling. In *JSM Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association. 2563-2567

O'Muircheartaigh, C., English, N., Eckman, S., Upchurch, H., Garcia, E., and Lepkowski, J. 2006. Validating a sampling revolution: Benchmarking address lists against traditional listing. In *JSM Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association. 4189-4196.