September 2016

Telephone Appends

White Paper

Prepared by

Rachel Harter
Joe McMichael
Derick Brown
Ashley Amaya
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

Trent Buskirk
David Malarek
Marketing Systems Group

RTI INTERNATIONAL

Telephone Appends

White Paper

September 2016

Prepared by

Rachel Harter
Joe McMichael
Derick Brown
Ashley Amaya
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

Trent Buskirk
David Malarek
Marketing Systems Group

CONTENTS

TABLES

Number **Page**

1. INTRODUCTION

Some mixed-mode studies include telephone contacts as well as mail or in-person contacts. An address-based sampling (ABS) study with telephone as one of the contact modes requires that telephone numbers be available for the sample addresses. Thus, a telephone number is an auxiliary variable sometimes appended to an ABS frame or sample (AAPOR 2016). The usefulness of the telephone numbers depends on both the percentage of addresses for which telephone numbers can be appended (append rate or match rate) and the accuracy of the telephone numbers associated with addresses (McMichael & Roe, 2012). Both the append rate and accuracy rate vary considerably by vendor, geography, and address type (Harter & McMichael, 2012).

Telephone numbers cannot be matched to all addresses, and we have no reason to believe that the availability or lack of telephone matches is random. Households that legitimately have no phone number available to append may be vacant, seasonal, or have residents for shorter periods of time. For example, consider an off-campus student apartment that has high turnover with many cell telephone numbers over a short period of time.

Sometimes errors occur in the process of matching the address list to the telephone source. Some vendors provide an indicator of match quality. Marketing Systems Group (MSG), for example, has three match levels for landline numbers: exact match, inexact match, and no match. And a vendor may provide multiple phone numbers for an address with different match levels.

The quality of appended phone numbers may be unknown for other reasons, as well. For example, appended phone numbers may not be current; a household may disconnect the phone service or move and take its telephone number to the new address.

Because of nonmatches and inaccuracies, random digit dialing (RDD) surveys are a better choice than ABS with phone appends for data collection that is primarily phone. Even so, telephone numbers with ABS samples may be useful for some mixed-contact or mixed-mode studies.

The append rate can vary considerably by vendor, source data, geography, and matching algorithm. If you are planning a study with phone as a supplemental mode, it helps to know how many telephone numbers you are likely to get.

For illustrative purposes, RTI partnered with MSG to estimate append rates from MSG's sources. Some basic findings are described below, and the interested reader may wish to explore append rates for custom subsets of addresses in the ABS Atlas exploration page for phone appends.

2. LANDLINE APPENDS

RTI selected a very large systematic national sample of 12 million residential addresses and asked MSG to indicate via a flag variable whether a landline telephone number was available for each sample address. (To keep costs low, RTI did not purchase the actual phone numbers.) Nationally, using only MSG's internal sources, the append rate was 42.6%. This rate can be considered a lower bound of what is possible; the national append rate may have been up to 30% higher if we had also used outside sources available to MSG.

The states with the highest append rates were Massachusetts at 52.3% and Rhode Island and Pennsylvania at 52.2%. The state with the lowest append rate was Alaska, at 19.0%. For city-style addresses, the landline append rate was 43.1%. Addresses flagged as vacant had an append rate of only 7.7%. Addresses flagged as seasonal (but not educational) had an append rate of 39.4%, and addresses flagged as educational had a paltry append rate of 3.6%. Addresses in high-rise buildings (2.7 million of our 12 million sampled addresses) had an append rate of 19.8%, while non-high-rise addresses had an append rate of 49.2%. Drop-point addresses had a landline append rate of 51.7%, although the unit-level accuracy of phone numbers appended to drop points is questionable. (If drop-point addresses are expanded to include a record for each unit, the same phone number may be appended to all units at a drop point, distorting both the match rate and the accuracy rate.)

It is possible for a vendor to have multiple landline telephone numbers for an address. In this study, MSG's flag variable for a landline append indicates the number of matched telephone numbers up to a maximum of three, and the most common value of the append flag indicated three or more matches. Having multiple distinct numbers for an address is not necessarily a good thing, as it may indicate a high degree of turnover in the address's occupancy. On the other hand, it is good to have multiple numbers to try if the first number is disconnected or associated with the wrong address. If multiple sources provide the same number, the accuracy is likely to be better.

3. CELL PHONE APPENDS

As Blumberg and Luke (2015) have shown, an increasing number of households have no landline telephone. As of the latter part of 2014, 45.4% of households had cell phone service only, 42.7% had both cell and landline service, while 8.4% had only landline service (the rest either having no phone service or landline with unknown cell service). Furthermore, they showed that demographic subgroups have different rates residing in cell-only households; therefore, a telephone survey that excludes cell-only households can introduce coverage bias. For ABS studies that use telephone as a supplemental mode, having cell phone appends as well as landline appends seems desirable.

MSG currently has a source that matches cell phone numbers to addresses, and we put it to the test. To keep costs low, RTI selected a stratified systematic subsample of approximately 120,000 addresses from the landline sample, where the strata were the 93 state or substate areas shown in Blumberg et al. (2011).

Nationally, 31.5% of this weighted sample of addresses had at least one cell phone append, 15.7% had only cell numbers available, 15.8% had both cell and landline numbers, and 41.9% had no numbers available to append. Delaware had the highest cell append rate (43.3%), and Alaska had the lowest cell append rate (9.1%). Among addresses flagged as vacant, 10.5% had a cell append. Addresses flagged as seasonal (noneducational) had a 19.2% cell append rate, and educational addresses had a 39.5% cell append rate, which is contrary to the landline append rate pattern, but not unexpected given the demographics of cell-only and landline-only households (Blumberg & Luke, 2015). High-rise addresses had a 19.9% cell append rate. Drop points had a cell append rate of 49.6%, based on only 1,809 drop points in the smaller sample.

4. ACCURACY

Remember that in ABS studies, the address is the sampling unit. When contacting an ABS address by telephone, it is essential to confirm whether you have reached the correct address. Amaya, Skalland, and Wooten (2010) found that a higher append rate from including lower quality matches resulted in a greater percentage of inaccurate numbers; that is, they found a tradeoff between the match rate and match quality.

McMichael and Roe (2012) came to a similar conclusion in their study of landline and cell numbers. Multiple vendors appended numbers of any match quality to a national sample of 45,000 addresses; 71% of the sample addresses had a telephone append initially, landline or cell or both. After discounting for disconnects and improperly matched addresses, the survey found that effectively 42% of the landline appends and 3% of the cell appends were accurate.

It was cost-prohibitive for us to test the accuracy of the available phone appends in our study. For this research, MSG provided append flags for only the landline numbers that met their highest match standards (exact match).

On the other hand, MSG provided a record for each cell append, and the number of available appends for a single address could be quite large. For each cell telephone number associated with an address, MSG provided a phone score or "confidence" variable whose values are described in Table 1 below. (Only the first four phone score values were included among the available appends.) We were unable to test the accuracy of cell phone appends by phone score, but testing the accuracy would be a useful next step. MSG's own internal testing has indicated that nearly half of the cell phone appends with a score of 4 are no longer working.

Table 1. MSG Phone Score for Cell Phone Appends

Score	MSG's Description	Unweighted Proportion of Cell Appends in Sample
1	Public DA (Daily Validation)	0.8%
2	Private Very High Confidence (Near Term Telco plus Transactional Verification)	55.3%
3	Private High Confidence (Moderate Term Telco plus Transactional Verification)	5.2%
4	Private Medium Confidence (no recent Validation and No Disconnect Transaction)	38.7%
5	Private Low Confidence (Disconnects and No Longer In Service)	0% (excluded from this exercise)

5. SUBNATIONAL ESTIMATES

For subnational studies, Appendix A provides the estimated append rates for landline and cell numbers by state based on the smaller sample provided to MSG. County-level estimates were not produced because the sample counts were often too small at that level.

6. ABS ATLAS REPRESENTATIONS

The ABS Atlas exploratory tool is intended to allow users to investigate append rates and other ABS features for U.S. geographies. The page with both cell and landline append figures in the ABS Atlas is based on the smaller sample sent to MSG. The ABS Atlas page for landline alone is based on the much larger sample and is expected to be correspondingly more reliable, if less complete because of the lack of cell phone information. Minor discrepancies in comparable figures between the two pages are to be expected because of sampling variability.

No standard errors are shown for the totals and percentages in the ABS Atlas interactive charts and tables. When the large landline sample is used, the sampling error is likely overwhelmed by nonsampling error. Sampling error is more of an issue for the cell sample, and standard errors are included in Appendix A. Standard errors were excluded from the ABS Atlas to keep the development of the interactive tool simple.

7. DISCLAIMERS

Keep in mind that append rates and accuracy rates are in constant flux, and will likely vary by vendor.

RTI International cannot endorse any vendor, even though MSG participated in this demonstration research, which is greatly appreciated.

The opinions expressed are those of the authors.

8. REFERENCES

American Association for Public Opinion Research (2016). *Address-based Sampling*. Report prepared for AAPOR Council by the Task Force on Address-based Sampling. Available from: http://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx.

Amaya, A., Skalland, B., & Wooten, K. (2010). What's in a match? *Survey Practice* 3(6).

Blumberg S. J., & Luke, J. V. (2015). Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2014. National Center for Health Statistics. Available from: http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201605.pdf

Blumberg S. J., Luke, J. V., Ganesh, N., Davern, M. E., Boudreax, M. H., & Soderberg, K. (2011). Wireless substitution: State-level estimates from the National Health Interview Survey, January 2007–June 2010. National health statistics reports; no 39. National Center for Health Statistics. Available from: http://www.cdc.gov/nchs/data/nhsr/nhsr039.pdf

Harter, R., & McMichael, J. P. (2012). Scope and coverage of landline and cell phone numbers appended to address frames. In *JSM Proceedings, Survey Research Methods Section*, pp. 3651–3665. Available from: https://www.amstat.org/sections/srms/proceedings/y2012/files/304339_72951.pdf

McMichael, J., & Roe, D. (2012). ABS and cell phones: Appending both cell phone and landline phone numbers to an address based sampling frame. Presented at the Joint Statistical Meetings, San Diego. Available from: http://www.rti.org/pubs/jsm2012mcmichael-abscell.pdf

Appendix A: Estimated Append Rates for Landline and Cell Numbers, by State

Both samples for this study were selected from an ABS frame based on the U.S. Postal Services' January 2016 Computerized Delivery Sequence File without the No-Stat file. Post Office boxes that were not an address-only way to get mail were excluded from the frame. Drop-point addresses were expanded to include a record for each drop unit.

The systematic sample of 12 million addresses was selected from the January 2016 file after sorting the frame by geographic stratum, county, ZIP code, carrier route, delivery sequence order, random number. The sample of approximately 120,000, allocated equally among the geographic strata, was selected systematically after sorting the same way within strata. Flags were appended in February 2016.

Landline rate estimates based on the sample of 12 million addresses are the same weighted or unweighted because all sample weights are the same. For rates estimated from the sample of 120,000 addresses, weights are necessary for geographic areas that cross boundaries of the 93 geographic strata. In Appendix Table 1, the weighted sample of 120,000 addresses was used for both landline and cell estimates.

Appendix Table 1. Landline and Cell Appends

	Append Percent (Standard Error)			
State	**Landline or Cell**	**Landline Only**	**Cell Only**	**Landline and Cell**
Total U.S.	58.1 (0.2)	26.7 (0.2)	15.7 (0.2)	15.8 (0.2)
Alaska	25.3 (1.2)	16.2 (1.0)	6.5 (0.7)	2.6 (0.4)
Alabama	62.3 (1.2)	29.8 (1.1)	16.0 (0.9)	16.4 (0.9)
Arkansas	51.6 (1.4)	25.8 (1.2)	15.1 (1.0)	10.6 (0.9)
Arizona	52.0 (1.0)	18.4 (0.8)	20.3 (0.8)	13.4 (0.7)
California	56.4 (0.8)	21.6 (0.7)	19.9 (0.7)	14.9 (0.6)
Colorado	60.1 (1.0)	22.8 (0.9)	22.1 (0.8)	15.3 (0.7)
Connecticut	65.0 (1.3)	30.2 (1.3)	13.2 (0.9)	21.6 (1.1)
District of Columbia	45.0 (1.4)	20.7 (1.1)	13.5 (1.0)	10.8 (0.9)
Delaware	71.6 (1.3)	28.3 (1.3)	20.1 (1.1)	23.2 (1.2)
Florida	58.2 (1.1)	26.5 (1.0)	15.9 (0.8)	15.8 (0.8)
Georgia	58.5 (1.1)	30.2 (1.1)	13.3 (0.8)	15.0 (0.8)
Hawaii	45.8 (1.4)	15.4 (1.0)	19.0 (1.1)	11.4 (0.9)
Iowa	66.9 (1.3)	27.4 (1.2)	18.9 (1.1)	20.5 (1.1)
Idaho	51.8 (1.4)	19.8 (1.1)	20.0 (1.1)	11.9 (0.9)

Appendix Table 1. Landline and Cell Appends (continued)

State	Append Percent (Standard Error)			
	Landline or Cell	Landline Only	Cell Only	Landline and Cell
Illinois	62.5 (0.9)	20.1 (0.8)	21.7 (0.8)	20.7 (0.8)
Indiana	66.7 (1.0)	25.5 (1.0)	23.5 (0.9)	17.7 (0.8)
Kansas	62.1 (1.1)	25.8 (1.0)	21.3 (0.9)	15.0 (0.8)
Kentucky	55.7 (1.4)	30.8 (1.3)	11.9 (0.9)	13.1 (0.9)
Louisiana	59.9 (1.4)	29.5 (1.3)	14.2 (1.0)	16.2 (1.0)
Massachusetts	61.3 (1.2)	33.3 (1.2)	10.0 (0.7)	18.0 (1.0)
Maryland	62.7 (1.0)	31.6 (1.0)	12.2 (0.7)	18.9 (0.8)
Maine	55.0 (1.4)	39.3 (1.4)	5.3 (0.6)	10.4 (0.8)
Michigan	60.2 (1.1)	30.1 (1.1)	12.6 (0.8)	17.4 (0.9)
Minnesota	67.3 (0.9)	29.2 (0.9)	20.6 (0.8)	17.5 (0.8)
Missouri	55.7 (1.1)	29.0 (1.0)	12.1 (0.7)	14.7 (0.8)
Mississippi	50.2 (1.4)	25.1 (1.2)	11.9 (0.9)	13.2 (0.9)
Montana	51.9 (1.4)	27.0 (1.2)	13.9 (1.0)	11.0 (0.9)
North Carolina	59.8 (1.4)	27.0 (1.2)	16.6 (1.0)	16.3 (1.0)
North Dakota	51.0 (1.4)	27.3 (1.2)	12.2 (0.9)	11.6 (0.9)
Nebraska	58.4 (1.4)	34.0 (1.3)	13.2 (0.9)	11.2 (0.9)
New Hampshire	60.9 (1.4)	42.2 (1.4)	8.7 (0.8)	10.0 (0.8)
New Jersey	60.9 (1.2)	29.4 (1.2)	9.8 (0.7)	21.7 (1.1)
New Mexico	46.4 (1.0)	21.4 (0.8)	15.3 (0.8)	9.7 (0.6)
Nevada	53.1 (1.1)	18.9 (0.8)	20.1 (0.9)	14.1 (0.8)
New York	57.6 (1.0)	29.0 (0.9)	12.1 (0.6)	16.5 (0.7)
Ohio	57.4 (1.1)	28.1 (1.0)	13.8 (0.8)	15.5 (0.8)
Oklahoma	51.3 (1.4)	24.4 (1.2)	13.2 (0.9)	13.7 (1.0)
Oregon	48.8 (1.4)	22.6 (1.2)	15.5 (1.0)	10.7 (0.9)
Pennsylvania	65.2 (1.0)	33.5 (1.0)	12.7 (0.7)	19.0 (0.8)
Rhode Island	64.8 (1.3)	29.4 (1.3)	11.6 (0.9)	23.8 (1.2)
South Carolina	54.6 (1.4)	29.0 (1.3)	12.6 (0.9)	12.9 (0.9)
South Dakota	65.6 (1.3)	28.1 (1.3)	19.6 (1.1)	17.8 (1.1)
Tennessee	56.4 (1.1)	29.3 (1.0)	13.0 (0.7)	14.1 (0.7)
Texas	53.4 (0.9)	21.1 (0.8)	17.3 (0.7)	14.9 (0.7)
Utah	56.1 (1.4)	19.8 (1.1)	19.5 (1.1)	16.8 (1.0)
Virginia	61.8 (1.4)	33.7 (1.3)	12.8 (0.9)	15.3 (1.0)
Vermont	58.0 (1.4)	43.7 (1.4)	6.8 (0.7)	7.4 (0.7)

Appendix Table 1. Landline and Cell Appends (continued)

State	Append Percent (Standard Error)			
	Landline or Cell	**Landline Only**	**Cell Only**	**Landline and Cell**
Washington	51.6 (0.8)	25.6 (0.7)	15.6 (0.6)	10.4 (0.5)
Wisconsin	66.3 (1.1)	29.8 (1.1)	19.7 (0.9)	16.9 (0.9)
West Virginia	44.2 (1.4)	31.9 (1.3)	5.6 (0.6)	6.7 (0.7)
Wyoming	49.8 (1.4)	20.9 (1.1)	14.0 (1.0)	15.0 (1.0)

Based upon the sample of approximately 120,000 addresses for both a landline and cell phone appends.