

ARTICLES

Billing Zip Codes in Cellular Telephone Sampling

David Dutwin¹

¹ Social Science Research Solutions

Keywords: telephone, sampling, cellphone

<https://doi.org/10.29115/SP-2014-0019>

Survey Practice

Vol. 7, Issue 4, 2014

Major sampling companies have recently begun to offer the appending of billing zip codes to cellular telephone samples. This study considers the utility of the billing zip flag by investigating first what percent of cell phone telephone numbers receive the flag, and then, through the use of a large scale national study, by measuring the percent of cell owners whose billing zip actually matches the self-reported zip of their household. Differences by geography, demographics, and characteristics of the zip codes themselves are analyzed to assess the degree of bias inherent in utilizing only sample which has a billing zip flag, and then of respondents who actually qualify for a study by reporting that they in fact live in the zip code(s) targeted by such sample. Finally, this paper considers the distribution of respondents who have a billing zip flag but do not live in their billing zip, and considers the increase in coverage that can be attained by casting a wider net to nearby zip codes, outside of the study target geography.

Dual-frame (landline and cell phone) telephone interviewing has become “business as usual” for survey research. Yet there continues to be challenges with dual-frame surveys. Notably, cell phone interviews remain more expensive to attain (Guterbock, Peytchev, and Rexrodee 2013). More significantly, while landlines continue to offer a fairly refined ability to target households geographically and demographically – thanks to our understanding of where telephone exchanges fall geographically and the ability to append demographic data to landline telephone numbers, – cell phones, in contrast, are nearly impossible to target to narrow geographic areas and have almost no available demographic appending available. Researchers conducting local studies, generally defined as areas smaller than a state, have not been able to effectively select cellular telephone numbers that will attain a high incidence of reaching persons in the target population. Even researchers of state and national studies have found it challenging to match the efficacy of landline stratifications in cellular frames. For example, researchers interested in interviewing African Americans might stratify landline telephone exchanges in order to sequester and oversample exchanges of high African American incidence. Within the cell phone frame, exchange stratification is ineffective. Most often, researchers have been forced to stratify by cellular rate center (Crawford 2012; Dutwin, Malarek, and Fahimi 2012; Skalland, Khare, and Furlow 2012; Wolter et al. 2011). While relatively effective, many major cities have only a single rate center (Dutwin and Malarek 2014). As such, cell phone samples are limited in their ability to oversample specific neighborhoods. In short, attempts to “mirror” landline stratification designs onto cell phones are approximate at best, and typically far less efficacious.

Making matters worse, on average, only 50 percent of adults who own cell phones in the United States live in their rate centers, though another 25 percent live within 15 miles of their home rate center boundary (Dutwin and Malarek 2014). Second, rate centers vary tremendously in size and shape and may not be suitable for certain studies. Most problematic is the fact that the largest rate centers are located in metropolitan areas, such that the centers in cities like Minneapolis, Houston, Miami, Atlanta, and others span across multiple counties and reach far into suburbia. Researchers interested in studies of the city limits or the central county of these and many other locations will find rate center to be too large for effective sample selection.

But there is a potential solution for these challenges. In mid-summer 2012, major sampling companies began to offer billing zip codes for cellular telephone numbers. Obviously, the ability to target cell phones to the level of zip code offers the potential of solving the local sample selection problem entirely. But does it work? That is the focus of the present investigation.

Data and Methods

To explore questions of efficacy, incidence, and coverage of billing zip sample selection, a large sample of cell phone owners from the SSRS omnibus survey were merged with billing zip data. The SSRS omnibus has been running consecutively for 26 years as a national, weekly, and (since 2009) dual-frame bilingual telephone survey. Each weekly wave consists of 1000 interviews, of which 500 are completed with respondents on their cell phones, and a minimum of 35 interviews are completed in Spanish. Data for this research was drawn from all cell phone interviews spanning from January 2011 through May 2012, for a total of 12,229 cell phone interviews. AAPOR RR3 for the cell phone frame was 8 percent.

By design, omnibus surveys, like many political polls and other opinion research, are short-field studies and by definition, attain low response rates. To mitigate this concern, billing zip was appended to a final wave of the omnibus in March 2013. Then, this sample was “rolled-over” to the next two subsequent waves of the omnibus, allowing up to 12 call attempts of all active sample, and refusal conversions made of all initial refusals (and placed at least one week after the initial refusal, per Triplett 2002). Additionally, the CATI system was set in the second wave to ring eight times, in order to allow respondents as much time as possible to answer their phones, and as well, to ensure that we were able to trigger every voice mail system, thereby minimizing the number of “no answer” dispositions and maximizing the number of “answering machine” dispositions. Overall response rate for this final wave was 17 percent.

These data afford us with the ability to investigate the following research questions:

- RQ1: What percent of U.S. cell phone owners have a billing zip on file?

Table 1 Frequencies of billing zip flag (n=12,299).

Category	Percent of total
No billing zip	40.0
Billing zip	60.0
Lives in billing zip	31.2
Does not live in billing zip (mismatches)	28.9
Distance from billing zip=less than 3.9 miles*	10.0
Distance from billing zip=3.9 to 6.8 miles*	10.0
Distance from billing zip=6.8 to 10.1 miles*	10.0
Distance from billing zip=10.1 to 15.6 miles*	10.0
Distance from billing zip=15.6 to 22.7 miles*	10.0
Distance from billing zip=22.7 to 42.2 miles*	10.0
Distance from billing zip=42.2 to 91.2 miles*	10.0
Distance from billing zip=91.2 to 228.7 miles*	10.0
Distance from billing zip=228.7 to 723.9 miles*	10.0
Distance from billing zip=723.9 miles or more*	10.0

*Percent of mismatches.

- RQ2: What percent of U.S. cell phone owners that have a billing zip on file actually live in their billing zip code?
- RQ3: Are there geographic or demographic differences between respondents who do not have a billing zip, who do have one but do not live in the billing zip code, and who do in fact live in their billing zip code?
- RQ4: How far away from their billing zip are respondents who have a billing zip flag but who do not live in their billing zip code?

Results

Table 1 provides the overall frequency of the bill zip flag for the 2012 omnibus data. Exactly 60 percent of respondents were found to have a billing zip on file. The 2013 “high-effort” data also put this estimate at exactly 60 percent. Furthermore, 58.7 percent of all confirmed households, including refusals, residential answering machines, callbacks, and completed interviews, have a billing zip code. Of respondents with a billing zip, 52 percent (31 percent of all respondents) live in their billing zip code. Table 1 also shows the decile distribution of respondents in the data who do not live in their billing zip. Approximately 30 percent of respondents who do not live in their billing zip nevertheless have the centroid of their current zip code that is within ten miles from the centroid of their billing zip. For about half of respondents who do not live in their billing zip, this metric is 20 miles.

Given that 40 percent of respondents do not have a billing zip flag at all and that a little under half of those with a billing zip flag do not live in their billing zip, it is of considerable interest whether limiting sample only to records that contain a billing zip will yield biased survey estimates. Table 2 explores

this question. With over 12,000 cases, nearly every difference is statistically significant. However, with a lens toward substantive differences, there is a mixed range of results. There is a modest trend by population density, such that fewer respondents who live in their billing zip code report living in rural areas compared to those who do not live in their billing zip (16.5 percent vs. 20.0 percent). But largely, the differences on density across billing zip subsamples and in comparison to the full sample are relatively insignificant, and this is also the case on region, on the number of persons within the household, and on gender, with somewhat more significant effects on employment status.

But from there the differences become more pronounced. While 34 percent of the total sample reports renting their home, the same is true for 49 percent of those without a billing zip flag, and there is a considerable difference between respondents who live in their billing zip (35 percent) and those who do not live in their billing zip (49 percent). Similar substantive differences between those who live in their billing zip and those who do not are found on cell phone only status (49 percent vs. 57 percent); earning \$25,000 or under (22 percent vs. 30 percent); being single (34 percent vs. 43 percent); being age 18–29 (26 percent vs. 35 percent); being Caucasian (74 percent vs. 64 percent); and being registered to vote (80 percent vs. 68 percent).

There are meaningful differences between the total sample compared to respondents with and without a billing zip flag. On income, for example, there are considerable differences at each end of the scale. While 40 percent of total respondents report earning \$50,000 per year or more, this is only true for 31 percent of those without a billing zip flag and is the case for 46 percent of those with a billing zip flag. As might be expected, there are subsequently meaningful differences in employment status (64 percent total; 68 percent with billing zip flag; 60 percent without billing zip flag). And again, a surely related pattern is found on education, such that while 31 percent of all respondents report the attainment of a college degree, this is the case for 36 percent of respondents with a billing zip flag and only 23 percent those without a billing zip flag. There are substantive differences on ethnicity as well. Compared to 64 percent of the full sample, only 55 percent of those without a billing zip are Caucasian while 70 percent of those with a billing zip are Caucasian. Finally, there is a 14 percent gap in those with and without a billing zip on being registered to vote.

With some understanding then of the degree to which sample will have not only a billing zip flag but that respondents with a billing zip flag will actually live in their billing zip code, researchers should be interested in the degree to which respondents live in nearby zip codes compared to zip codes in other counties or states. Table 3 provides a measure of the degree to which respondents live in a zip code “next door” to their billing zip, or the zip code beyond that, or further out. These data were computed by taking the square root of the square miles of each respondent’s billing zip and multiplying by one plus half to approximate persons who might on average live one zip code

Table 2 Differences in geography and demographics by billing zip status (n=12,299).

Metric	Total respondents	No billing zip	Billing zip	Do not live in billing zip	Live in billing zip
Population density					
Densest quintile	23.4%	24.5%	22.6%	21.3%	23.7%
2nd quintile	20.1%	19.7%	20.3%	19.6%	21.0%
3rd quintile	19.1%	17.6%	20.0%	19.4%	20.4%
4th quintile	18.6%	18.0%	18.9%	19.7%	18.3%
Least dense quintile	18.9%	20.1%	18.1%	20.0%	16.5%
Region					
Northeast	17.2%	17.1%	17.3%	15.9%	18.5%
Midwest	23.8%	23.0%	24.4%	24.2%	24.5%
South	38.5%	41.1%	36.8%	37.7%	36.0%
West	20.4%	18.8%	21.5%	22.2%	20.9%
Home ownership					
Rented	44.8%	49.2%	41.8%	49.3%	34.9%
Owned	55.2%	50.8%	58.2%	50.7%	65.1%
Number of persons					
1	17.8%	18.3%	18.1%	19.7%	16.6%
2	31.0%	29.9%	31.7%	31.8%	31.7%
3	19.5%	18.9%	20.0%	19.7%	20.2%
4+	31.3%	32.8%	30.2%	28.8%	31.5%
Cell phone only					
No	47.3%	47.3%	47.3%	43.0%	51.2%
Yes	52.7%	52.7%	52.7%	57.0%	48.8%
Income					
Under \$25,000	31.3%	39.6%	25.9%	30.4%	22.0%
\$25k-\$50k	28.5%	29.5%	27.8%	27.9%	27.8%
\$50,000 or more	40.2%	30.9%	46.2%	41.7%	50.2%
Gender					
Female	42.5%	42.2%	42.7%	41.8%	43.5%
Male	57.5%	57.8%	57.3%	58.2%	56.5%
Marital status					
Single	39.1%	40.2%	38.3%	42.9%	34.2%
Married	41.6%	38.6%	43.7%	39.2%	47.7%
Widow/div./sep.	19.3%	21.2%	18.0%	17.9%	18.1%
Employment					
Employed	64.4%	59.5%	67.6%	66.6%	68.6%
Unemployed	9.0%	10.9%	7.7%	8.3%	7.2%
Other	26.7%	29.7%	24.7%	25.1%	24.2%
Age					
18-29	30.8%	31.4%	30.4%	34.9%	26.3%
30-44	23.6%	23.0%	24.0%	23.6%	24.3%
45-64	33.5%	32.7%	34.0%	30.5%	37.2%
65 and older	12.1%	12.9%	11.6%	11.0%	12.1%
Education					
No H.S. diploma	11.8%	16.8%	8.4%	9.8%	7.2%
H.S. diploma	30.6%	35.5%	27.3%	28.1%	26.6%

Metric	Total respondents	No billing zip	Billing zip	Do not live in billing zip	Live in billing zip
Some college	27.2%	24.4%	29.1%	27.6%	30.4%
4-year degree	20.0%	15.9%	22.7%	22.3%	23.1%
Graduate degree	10.5%	7.4%	12.5%	12.2%	12.7%
Ethnicity					
White	63.8%	54.8%	69.8%	65.4%	73.7%
Black	13.0%	16.6%	10.6%	11.6%	9.7%
Hispanic	16.1%	21.3%	12.6%	14.8%	10.6%
Other	7.1%	7.2%	7.0%	8.2%	6.0%
Registered to vote					
No	29.6%	34.8%	26.1%	32.3%	20.5%
Yes	70.4%	65.2%	73.9%	67.7%	79.5%

Table 3 Approximate number of zip codes distant from billing zip for respondents whose billing zip code does not match their self-reported zip code (n=3,025).

Estimated number of zip codes distant	Percent	Cumulative percent
1	18.6%	
2	11.7%	30.3%
3-5	20.1%	50.4%
6-10	12.4%	62.8%
11-50	15.8%	78.6%
51-100	6.2%	84.8%
101-250	6.1%	90.9%
251 or more	9.1%	100.0%

away from their own; two plus half for two zip codes away, and so on. In other words, a respondent who lives in a zip code of nine square miles will live in a zip code that is (very) roughly three linear miles from north to south and three linear miles from east to west. Assuming the respondent lives in the middle of his or her zip code, he or she would have to travel half the linear distance (1.5 miles) to make it to the border of the zip code, and therefore to encompass all of the boundaries of the next zip code would mean that person would have to travel not just the 1.5 miles but another 3 miles as well, assuming that the most likely probability is that the neighboring zip is approximately the same size. While certainly there are many potential sources of error in this measure, there is reason to believe that much of the error will be random in aggregate, and in any event, the purpose of this exercise is to simply gain a general sense of how many persons not living in their billing zip live in an adjacent zip, or one beyond the adjacent zip, or further away.

We find that just under a fifth (18.6 percent) of all respondents who have a billing zip flag but who do not live in their billing zip most likely live in an adjacent zip code. Another 11.7 percent live two zip codes away, and 20.1

percent live three to five zip codes away. In other words, three out of ten (30.3 percent) live two zip codes away and half (50.4 percent) live within five zip codes. Nearly two thirds (62.8 percent) live within a 10 zip code radius.

Discussion

The overarching point of these analyses is to provide researchers with information by which they can design local studies utilizing bill zip flags. This much is clear:

1. Only 60 percent of cellular telephone numbers will yield a billing zip code as of the publication of this research article. Therefore, the baseline, best case coverage for any study solely using a sample that contains billing zips, is only 60 percent.
2. There are significant differences between respondents whose cellular telephone numbers are associated with a billing zip flag and those whose numbers do not yield a billing zip flag. Respondents without a billing zip flag are more likely to reside at the lower end of socioeconomic status as evidenced by lower income, lower rates of home ownership, greater rates of unemployment, and a significantly lower level of educational attainment in comparison to respondents with a billing zip flag. As well, respondents without a billing zip flag are far more likely to be African American or Hispanic.
3. A study of a single zip code, on average, will only attain 31 percent coverage, given that only 60 percent will have a billing zip flag and then only 52 percent of those with a billing zip flag will reside in their billing zip code.
4. The gap between total cellphone owners and those who not only have a billing zip flag but live in their billing zip is even more substantial with regard to socio-economic status. Therefore, a study of a single zip code not only attains just 31 percent coverage, on average, but those who are able to be interviewed are off by 10 percent or more, compared to all cell owners, on home ownership, income, race (percent Caucasian), and reporting that they are registered to vote. Such a sample will also be modestly older, undereducated, employed, and married.
5. Researchers can increase the coverage of their studies by sampling adjacent zip codes. Giving oneself a “five zip code pad” to the selected zip codes can potentially increase coverage from 31 percent to 45 percent.

Of course, the application of billing zip will not be limited to studies of single zip codes, which are exceedingly rare in survey research. But when is billing zip a useful tool? Coverage issues aside, research exclusively using billing zip should probably be limited to studies whose target geographies are smaller than

a rate center and have significant budgetary constraints, given the relatively low coverage afforded by the sample. That said, billing zip provides a tremendous opportunity to oversample very small geographic areas, again, those smaller than the rate centers in which they are embedded, or in studies that use rate center, area code, or some other larger geography as their primary method of sample selection and billing zip for oversampling. For example, many state health surveys are interested in understanding specific at-risk populations for certain health disparities. A statewide study could utilize billing zip to oversample, for example, Native Americans on reservations or zip codes where persons earning under 100 percent of the federal poverty level are most concentrated. With proper weighting techniques, researchers can eliminate the bias of such oversampling, and depending on the aggressiveness of the probability of selection of this sample, do so without an undue increase in variance. Billing zip flags look to be a highly enticing tool for researchers looking to save costs in highly localized studies, but given the relatively low coverage of persons whose contain a billing zip flag and in fact live in their billing zip code, one must utilize such sample with caution and take steps to specifically adjust for the use of such sample, lest they attain estimates that are biased when compared to the overall target population.

REFERENCES

- Crawford, M. 2012. "Cell Phones in Smaller Geographies: Are You Reaching the Right People?" In *Presented at the American Association for Public Opinion Research Conference*. Orlando, FL.
- Dutwin, D., and D. Malarek. 2014. "The Use of Recent Activity Flags to Improve Cellular Telephone Efficiency." *Survey Practice* 7 (1): 1–10.
- Dutwin, D., D. Malarek, and M. Fahimi. 2012. "Sampling Local Cell Phone Samples by Rate Center: Efficacy, Coverage and Incidence."
- Guterbock, T., A. Peytchev, and D.L. Rexrodee. 2013. "Cell-Phone Costs Revisited: Understanding Cost and Productivity Ratios in Dual-Frame Telephone Surveys." In *Presented at the American Association for Public Opinion Research Conference*. Boston, MA.
- Skalland, B., M. Khare, and C. Furlow. 2012. "Geographical Accuracy of Cell Phone Samples and the Effect on Telephone Survey Bias, Variance and Cost." In *Presented at the American Association for Public Opinion Research Conference*. Orlando, FL.
- Triplett, T. 2002. "What Is Gained from Additional Call Attempts and Refusal Conversion and What Are the Cost Implications?" *Urban Institute*.
- Wolter, K.M., R. Montgomery, X. Tao, S. Greby, and E. Kennedy. 2011. "Accuracy of Geographic Stratification in a Cell-Phone Survey." In *Presented at the American Association for Public Opinion Research Conference*. Phoenix, AZ.