An Evaluation of Ported Telephone Numbers in the 2013 California Health Interview Survey

Greg Norman[1] and Ismael Flores Cervantes[1]
[1]Westat, 1600 Research Boulevard, Rockville, MD 20850

Abstract

In telephone surveys, ported numbers are those telephone numbers originally assigned as a wireline that are subsequently converted to wireless service. These ported numbers are sampled as a part of a landline sample and identified as wireless during the post-sampling processing. Ported telephone numbers, as of 2013, represent a small but growing proportion in landline samples, and there is little research on their characteristics. In some survey operations, ported numbers are excluded from data collection, while in others they are dialed as a part of a cell phone sample. In the California Health Interview Survey (CHIS), a representative dual-frame random digit dialed (RDD) telephone survey of the California non-institutionalized population, ported telephone numbers have been sampled and dialed as a part of the cell phone sample since 2007. In the 2013 CHIS, sampling and subsequent processing identified over 10,000 ported telephone numbers. This paper analyzes the results of dialing these ported telephone numbers and compares their respondents' characteristics to those from both the landline and cell phone samples to assess potential bias if the ported numbers are excluded in data collection.

Keywords: Ported telephone number, RDD, landline, cell phone, health survey, CHIS

1. Introduction

In this paper, we use data from the 2013 California Health Interview Survey (CHIS) to profile those respondents that ported their telephone number from a landline to a cell phone. Although small in number, these households represent a growing proportion of numbers found in landline frames in the United States. This paper presents an analysis of these respondents and provides recommendations for handling ported numbers in future surveys.

We begin in Section 2 by defining ported numbers, we talk about their history, describe how they are identified, and discuss their impact on telephone surveys. In Section 3, we provide the details of the sample design of CHIS. This includes a discussion of how the sampling methodology has changed over time, how respondents are selected, and how the analysis weights are created. Section 4 shows the results of our analysis of the ported numbers in CHIS. We provide estimates of their prevalence, their screener eligibility rates when compared to the random digit dialed (RDD) cell sample, and compare the respondents to their RDD cell and landline counterparts. Lastly, a discussion of the results is found in Section 5.

2. Ported Telephone Numbers

A ported telephone number is one that has been switched from one telecommunication provider to another or from one type of service to another. [1] The ability to port telephone numbers resulted from the implementation of a requirement known as Local Number Portability (LNP). In the United States, LNP was mandated by the Federal Communications Commission (FCC) in 1996. This allowed customers to switch their service providers and maintain their telephone number. In 2003, the FCC further clarified LNP requirements setting guidelines for "type of service" portability, allowing for the porting of telephone numbers from a landline to a cell phone.

2.1 Implications for Survey Research

Ported number identification in telephone survey research in the United States is needed to comply with the 1991 Telephone Consumer Protection Act (TCPA), which restricts the use of an "autodialer" for contacting cellular telephone numbers. In order to comply with this regulation, survey research organizations rely on companies that offer sampling services for telephone surveys to identify telephone numbers that have been ported from a landline to a cell phone as a part of sample screening. In turn, these sampling services companies obtain this information directly from the LNP administrator, NeuStar. Once identified, all ported numbers must be removed from any landline samples that use an autodialer.

The prevalence of ported cellular telephone numbers in RDD telephone numbers and the ability of sample screening to identify them was initially examined by Link, Town, and Mokdad (2007). Besides this, there has been little research on ported numbers. Furthermore, there is currently no standard practice or consensus on how ported numbers should be handled in telephone surveys. In some surveys ported numbers are excluded from dialing, while in others, such as the CHIS, ported numbers are dialed using cell phone dialing protocols.

3. The California Health Interview Survey

The CHIS is a RDD telephone survey of California's noninstitutionalized population first administered in 2001. It is the largest health survey ever conducted in any state and one of the largest health surveys in the United States. The CHIS is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Funding for CHIS comes from multiple sources, including Federal Government agencies and private foundations. The first five cycles of CHIS were conducted biennially from 2001 to 2009. In 2011, CHIS became a continuous survey with data collected over a two-year period with yearly milestones. Westat conducted the sampling and data collection for the first seven two-year cycles of CHIS.

CHIS collects extensive information on public health, health status, prevalence of chronic conditions, health-related behaviors, health insurance coverage, and access to health care services. Data from CHIS supports the production of estimates for the state, many counties, and for groups of the smallest counties in California. The survey also supports the study of the characteristics for the major racial and ethnic groups and a number of

[1] See more information at https://www.npac.com/number-portability.

smaller ethnic groups within the state. Adults, parents or guardians of children, and adolescents within California households are eligible for sampling.

3.1 Sampling and Weighting Methodology in CHIS

Initially, CHIS was designed as a list-assisted RDD telephone survey of landline households with oversamples from surname lists of ethnic groups of analytic importance. Since its first administration in 2001, CHIS has seen the rapid growth in the number of households that only have cell phones. Early research on this cell-only population found important differences in demographic and health characteristics between the cell-only and landline respondents (Tucker, Brick, & Meekins, 2007). As a result, a RDD sample of cell phone users was piloted in CHIS 2005, the results of which were described in Brick, Edwards, and Lee (2007). Since 2007, CHIS has fully incorporated cellular telephones into its sample design. Additional details on the sample design of the CHIS landline component can be found in Flores Cervantes, Jones, Alvarez Rojas, Brick, Kurata, and Grant (2006) and for cell component in Brick, Flores Cervantes, Kali, Norman, and Grant (2012).

3.2 Sampling Frames

Because the landline and cell samples represent two different populations – households with landlines and individuals with cell phones – samples of telephone numbers were selected from the two distinct frames (landline and cell) and processed separately. This is shown in Figure 1 and described in the sections that follow.

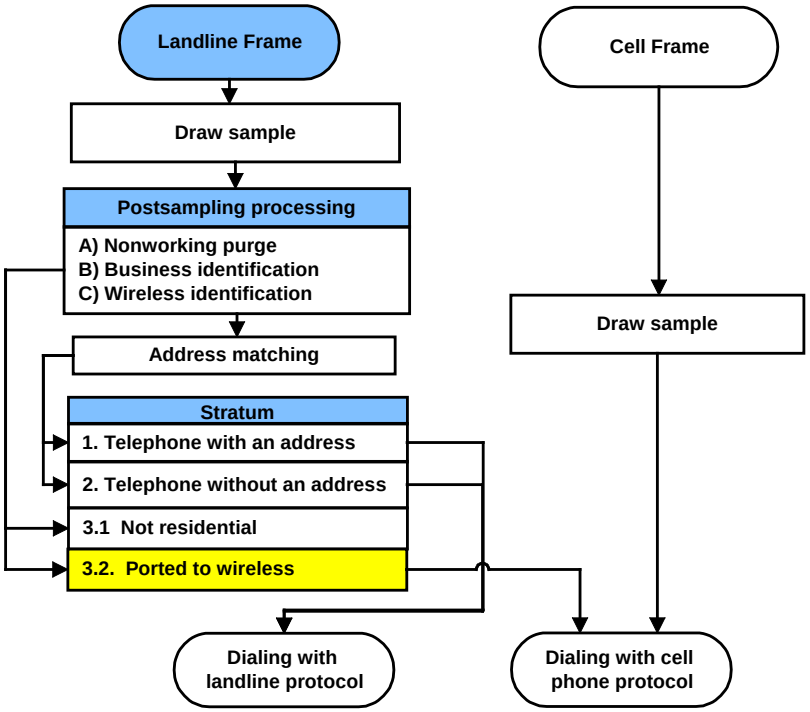

Figure 1: Dual-frame sampling process in CHIS

3.2.1 Landline Sampling and Dialing

In CHIS, a stratified sample of telephone numbers was selected from a landline frame of telephone numbers assigned to area codes in California. Then, a two-step process was applied to the sample. In the first stage, the sampled numbers were screened to identify (a) disconnected or nonworking numbers, (b) known businesses, or (c) numbers ported to wireless service. The elimination of numbers in the first two groups a priori increases the efficiency and decreases the cost in landline samples because these numbers can be removed from dialing. The numbers in the last group were removed from the landline sample as they cannot be dialed using an autodialer.

The remaining landline numbers underwent a process of address matching. Here we appended a mailing address to the sampled telephone number if it was available. Advance letters in RDD surveys have been shown to increase response rates for cases with mailable addresses (Groves, 2006), so those telephone numbers with a valid address were sent an advance or prenotification letter that introduced and described the goals of the survey. Telephone numbers without a matching address were not mailed a prenotification letter. In the last step, both sets of numbers (mailable and non-mailable) were dialed using landline protocols.

3.2.2 Cell Phone Sampling and Dialing

Similar to the sample selection of the landline sample, a stratified sample of telephone numbers was selected from a cell frame telephone numbers in California. However, in contrast to the landline sample, there was no post-sampling processing for the cell phone sample.

The final CHIS cell phone sample included the sample selected from the cell frame and the ported telephone numbers identified in the processing of the landline sample. This combined set of telephone numbers was then dialed using cell phone protocols.

3.3 Eligibility and Adult Sample Selection

In CHIS, household eligibility was determined during the screener interview. If the telephone number was dialed using landline protocols, the household was eligible if the telephone number was used for non-business purposes, there was at least one adult in the household, and that the respondent lived in California. After confirming household eligibility, we sampled one adult among all adults that live in the household.

In contrast, a different eligibility criterion was used for the cell phone sample. In the cell phone screener interview, in addition to the requirements used in the landline sample, additional questions were asked to further determine the eligibility of individuals. First, it was confirmed that the dialed telephone number was for a cell phone. If so, the screener respondent was asked if the telephone was shared by any other adults who lived in the same household. We then sampled one adult from among all adults in the household who shared that cell phone.

3.4 Weighting

Due to their different sampling frames, sample selection methods, and post-sampling procedures, parallel weighting adjustments were performed for the landline and cell samples. This is shown in Figure 2.

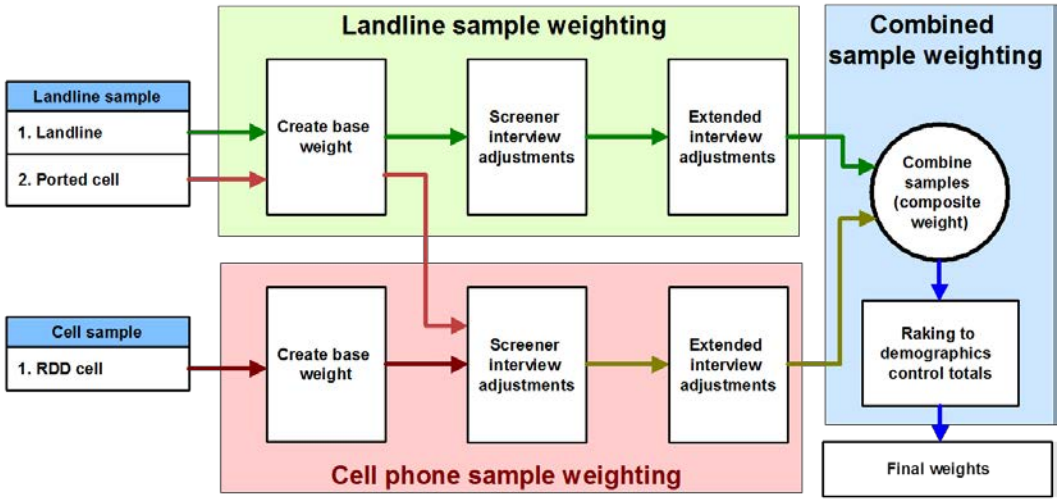

Figure 2: Weighting process in CHIS

The first step in weighting was to create a base weight for each sample. In this step, the ported telephone numbers receive a base weight that corresponds to their selection probabilities from the landline sampling frame. In the next step, the base weights were sequentially and independently adjusted for subsampling and nonresponse in the screener and extended interviews. Because they were dialed using cell protocols, the weighting adjustments for the ported numbers followed the same path as the cell sample. After separately adjusting the weights of the samples for nonresponse, a composite weight was created. This weight accounted for those respondents that had both a landline and a cell phone. In doing this, the resulting weights represent all California residents who had either type of telephone service. The last step of weighting was raking the composited weights to demographic control totals to create the final weights that represent the population of California.

4. Results

In this section we first provide estimates of the growth of ported numbers in California over the last three cycles of CHIS. Next, we compare the eligibility and geographic coverage of ported numbers and how these compare with RDD cell phones. Lastly, we compare estimates of ported respondents to both the RDD cell and the landline respondents on 38 variables (20 of these are health related and 18 are demographic).

Data in our analysis comes primarily from the 2013 CHIS adult interview. We also utilized data from the 2009 and 2011 CHIS adult interviews for the estimates over time. The sample sizes for the analysis using CHIS 2013 adult interviews are 32,007 landline, 6,982 RDD cell, and 770 ported cell. Child and adolescent interviews were excluded from our analysis because these interviews are a subset of the households eligible for the adult interview, and their sample size for ported cell numbers is too small.

4.1 Increase of Ported Numbers over Time

To assess the increase in ported of numbers over time we looked at three measures. The result of this analysis is shown in Table 1 (all estimates have been computed using the final weights from the adult interview).

Table 1: Estimates on ported numbers in CHIS 2009, 2011, and 2013

	2009	*2011*	*2013*
Number of ported numbers	24,014	67,331	135,423
Percentage of ported numbers in the landline frame	0.12%	0.48%	1.45%
Percentage of ported numbers among cell phone respondents	0.32%	0.49%	0.70%

We first examined the growth of ported numbers in California since 2009. As this table indicates, the estimates of the number of adults who had their telephone number ported was under 25,000 in 2009 and increased to over 135,000 by 2013. We next estimated the percentage of residential telephone numbers that were classified as ported in the landline sampling frame. As indicated in the table, the ported numbers have become much more prevalent in landline samples since 2009 when they accounted for just 0.12 percent of eligible households. By 2013, this estimate of the number of ported number was 1.45 percent. Then, we looked at the estimated proportion of cell respondents who were identified as having ported their number. Although the growth rate was less dramatic, the estimates did more than double from 0.32 percent in 2009 to 0.70 percent in 2013.

4.2 Quality, Eligibility, and Geographic Coverage of Ported Numbers

In the second part of the analysis, measures of the sample quality, eligibility and geographic coverage of ported telephone respondents were compared to the same measures for RDD cell phone respondents. The results of this analysis are presented in Table 2.

Table 2: Comparisons of ported cell sample to RDD cell on measures of sample quality, screener eligibility, and geography

	Ported cell	*RDD cell*
Nonworking or nonresidential	11.64	29.64
California resident	91.37	87.58
Eligible adult cell phone	55.69	78.96
Ineligible	44.31	21.04
Not a cell phone	36.88	0.71
Teen phone	7.43	20.33
Lives in sampling strata	90.01	78.97

4.2.1 Sample Quality

The rate of nonworking or nonresidential numbers for ported cell numbers is lower than for the RDD cell numbers. This result is not surprising because the former were actively ported from an active residential number. We would expect the vast majority to remain as valid numbers (and have a higher residency density) because they have been used in the past. In contrast, we see higher levels of invalid numbers from the RDD cell sample due to their being randomly generated and not prescreened.

4.2.2 Screener Eligibility

We also examined the eligibility of the ported numbers during the screener interview. In CHIS, the first eligibility requirement was that respondents be residents of California. As indicated in Table 2, ported cell cases had a slightly higher in-state eligibility rate than the RDD cell sample (i.e., 91.37% compared to 87.58%, respectively).

The second eligibility criterion required that the screener respondent be an adult who used or shared the cell phone dialed. Ineligibility at this stage could occur for two reasons. The telephone number was for a landline or we had contacted an adolescent who did not share their phone with an adult household member. The analysis shows that the eligibility rate for the ported numbers (55.69%) was lower than for the RDD cell (78.96%). Ineligibility among the ported numbers was most often the result of respondents reporting that the phone numbers was a landline (36.88% of screener respondents). In contrast, the largest portion of ineligible respondents in the RDD cell phone sample was the result of reaching a teen phone (20.33% of screener respondents).

4.2.3 Sample Stratification

We also reviewed the distribution of the sample by sampling and self-reported stratum. In CHIS, there were 44 sampling strata that cover the 58 counties in California. The first 41 strata were comprised of single counties while the smallest 17 counties were combined into three sampling strata. Due to the imprecise geographic information on phone numbers in cell phone frames, there is often a large rate of misclassification between sampling strata and respondents' self-reported data. Because of this misclassification, it is more difficult to accurately target specific geographies in a cell sample than a landline sample. In dual-frame telephone surveys with targets based on geography and by sample type, this misclassification makes it much more difficult to manage and reach completion goals for the cell phone component.

The geographic accuracy rate (the complement of the misclassification rate) can be defined as the proportion of numbers where sample strata correspond to the reported geographic strata. For the ported numbers, this rate can be used to evaluate whether the respondent had moved to a different geographic area in California. For the RDD cell cases it is less clear as to what this rate indicates. This rate could also be the result of inaccuracy in the geographic assignment for the telephone number. The analysis shows that about 10 percent of the ported numbers reported a geographic location different than the strata they were sampled from (90% accuracy). In contrast, the accuracy rate is 79 percent for the RDD cell sample.

4.3 Data from Ported Respondents

In the last part of our analysis, we compared estimated proportions of 38 variables (20 key survey variables and 18 demographics) from adult respondents in CHIS. Table A-1 and A-2 in the Appendix shows the estimates for the ported numbers when compared to the RDD cell (Table A-1) and landline respondents (Table A-2). In addition to these estimates, Tables A-1 and A-2 include the associated *p* values that were used to determine if the differences of the estimates between the two groups are statistically significant.

The data from Tables A-1 and A-2 are graphically represented in Figures 3 and 4. In both figures, the horizontal axis corresponds the estimated proportion for the ported respondents, while the vertical axes corresponds the estimated proportion computed using the landline respondents (Figure 3) or RDD cell phone respondents (Figure 4). Both figures include a 45° reference line that is an indicator of what the results would look like if the estimates had been the same.

Figure 3 shows that the estimates comparing ported and RDD cell respondents are highly correlated (R^2= 0.86). However, many estimates fall far from the 45° reference line.

Furthermore, the results in Table A-1 show that half of these variables are significantly different at the 95 percent confidence level. Among these differences, the health-related survey questions are slightly less likely to be significant when compared to the demographics. For the health-related variables, 9 of 20 were significant. In contrast, differences for 10 of the 18 demographic variables were statistically significant.

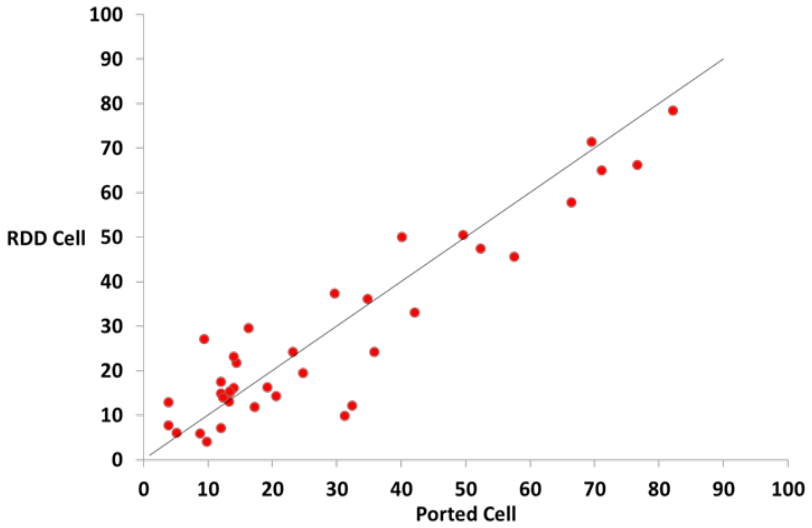

Figure 3: Comparison of estimates of percentages of ported respondents to RDD cell respondents

Figure 4 shows the estimates for the same 38 variables, when comparing the ported and landline respondents. The estimates of percentages from ported number respondents are highly correlated to estimates from the landline sample (R^2= 0.99), and the pattern is in close proximity to the 45° reference line. This result is confirmed by the data in Table A-2. The difference between the estimates of the ported and landline respondents is statistically significant for only 1 of the 38 variables (estimate for percentage of income of $20,000 or less) at the 95 percent confidence level.

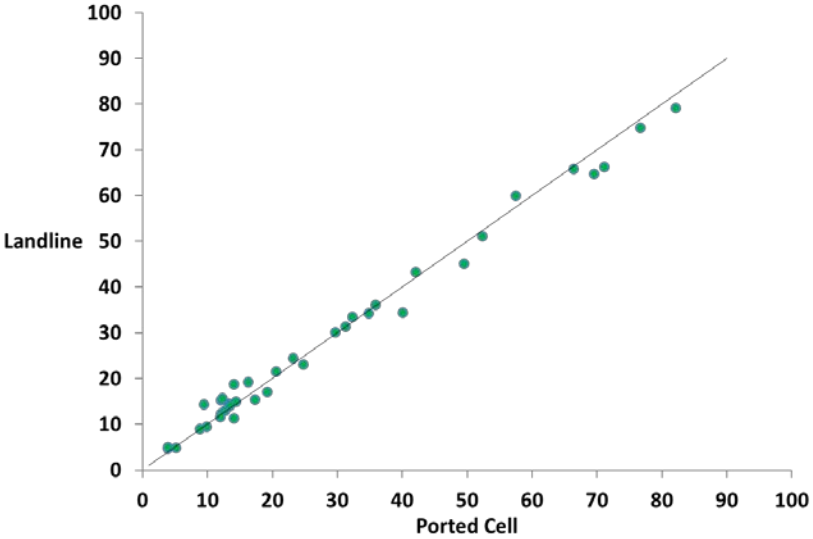

Figure 4: Comparison of estimates of percentages of the ported respondents to landline respondents

5. Discussion

Ported telephone numbers require special attention in telephone surveys. To comply with Federal regulations, these numbers must either be removed or dialed using cell phone protocols. This paper looks at these numbers in more depth and provides guidance on which approaches are most appropriate to use in future telephone surveys similar to CHIS.

Data from the last three cycles of CHIS showed a continued growth of ported numbers in the last six years. We found that that from CHIS 2009 to CHIS 2013, the estimated total of ported numbers increased by more than five-fold, the percentage of numbers in the landline sample classified as ported cell phones increased by a factor greater than 10, and the proportion of cell phones that were ported from landline among all cell phone numbers doubled.

Results from the CHIS 2013 screener interview showed that ported telephone numbers were more likely to be a working residential cell phone. Ported number respondents were also more likely to live in California and reside in the stratum where they were sampled, and were less likely to be a teen-only phone number. However, over one-third of respondents initially identified as being a ported number reported that they were reached on a landline. This result was an issue in CHIS because there were separate screener interviews for the landline and cell samples, and they were administrated using different CATI systems. In CHIS, there was no mechanism to switch telephones from one system to the other during the screener interview. This not only results in a loss of efficiency but could result in a hidden bias.

When comparing the CHIS 2013 adult interview data, we observed that the characteristics of the ported cell phone respondents are quite different than the RDD cell respondents. The analysis showed that among 38 estimates, significant differences were found in half of the estimates. In contrast, the comparison of estimates of ported respondents to estimates of landline respondents identified just one significant difference at the 95 percent confidence level.

Overall, the results of the analysis suggest the exclusion of ported numbers from telephone samples could be viewed as a minimally biased and cost-effective way to handle these numbers. Not only do the ported cases require additional efforts when dialed and weighted, but a high proportion of ported numbers were misidentified. Although they are increasing rapidly, we found that the characteristics of ported number respondents have minor differences when compared to the landline sample with which they share a common sampling frame. Based on these results we believe that employing weighting adjustments at the screener level to account for the presence of these ported numbers in landline frame rather than dialing and weighting them as part of the cell phone sample is a reasonable procedure currently.

References

Brick, J.M., Edwards, W.S., and Lee, S. (2007). Sampling telephone numbers and adults, interview length, and weighting in the California Health Interview Survey cell phone pilot study. *Public Opinion Quarterly*, 71, 793-813.

Brick, J.M., Flores Cervantes, I., Kali, J., Norman, G., and Grant, D. (2012). A review of the cell phone sample component of the California Health Interview Survey. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, [CD-ROM]. Alexandria, VA: American Statistical Association.

Flores Cervantes, I., Jones, M., Alvarez Rojas, L., Brick, J.M., Kurata, J., and Grant, D. (2006). A Review of the Sample Design for the California Health Interview Survey. *Proceedings of the American Statistical Association Section on Survey Research Methods, 2006 Joint Statistical Meetings*, [CD-ROM]. Alexandria, VA: American Statistical Association.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646-675.

Link, M., Town, M., and Mokdad, A. (2007). Telephone number portability and the prevalence of cell phone numbers in random digit-dialed telephone survey samples. *International Journal of Public Opinion Research*, *19*(4)*:* 504-511.

Tucker, C., Brick, J.M., and Meekins, B. (2007). Household telephone service and usage patterns in the U.S. in 2004: Implications for telephone samples. *Public Opinion Quarterly*, 71, 3-22.

Appendix A

Table A-1: Comparison of estimates of percentages of RDD cell respondents and ported respondents for 38 variables in CHIS 2013

	RDD cell (n=6,982)	*Ported (n = 770)*	*p value*
Health-related variables			
Has asthma	14.46	12.69	0.459
Has diabetes	7.05	12.03	0.008
Has high blood pressure	24.14	35.85	< 0.001
Smoked 100 cigs in lifetime	36.12	34.80	0.651
Had alcohol past 12 months	71.47	69.53	0.460
Covered by MediCal	15.25	13.46	0.484
No usual source of health care	17.48	11.99	0.019
Condition limits physical activity	14.34	20.56	0.027
Diagnosed heart disease	4.11	9.84	0.001
Blind or deaf	5.93	8.80	0.058
Has difficulty learning, remembering, etc.	13.13	13.27	0.954
Covered by Medicare	12.14	32.35	< 0.001
Heath rating - Fair/Poor	19.43	24.73	0.108
Feel nervous - All/Most of time	6.08	5.13	0.403
Delayed care in past 12 months	16.12	14.02	0.380
Feel save in neighborhood - Always	47.41	52.34	0.133
No fast food past week	33.07	42.05	0.001
No sex partners past 12 months	16.33	19.20	0.215
More than one sex partner past 12 months	12.84	3.85	< 0.001
Did not see MD past 12 months	21.82	14.41	0.002
Demographics			
Country of birth - USA	64.92	71.10	0.064
Country of birth - Mexico	15.35	13.39	0.420
Marital status - Married	45.53	57.50	< 0.001
Marital status - Never Married	29.57	16.25	< 0.001
Education - Less than HS	14.85	12.03	0.229
Education - HS grad or GED	24.21	23.20	0.734
Education - Graduate/Professional	11.86	17.28	0.006
Male	50.51	49.56	0.761
White	57.76	66.42	0.006
African-American	7.64	3.89	< 0.001
Hispanic	37.30	29.66	0.014
All/most calls on cell phone (Dual users)	50.09	40.09	0.027
No children (under 12) in HH	66.15	76.63	< 0.001
No adolescents (12-17) in HH	78.44	82.14	0.121
Income 20K or less	23.11	14.05	< 0.001
Income more than 135K	13.97	12.32	0.342
Age 18-29	27.07	9.42	< 0.001
Age 65+	9.92	31.22	< 0.001

Table A-2: Comparison of estimates of proportions of landline respondents and ported respondents for 38 variables in CHIS 2013

	Landline (n=32,007)	*Ported (n = 770)*	*p value*
Health-related variables			
Has asthma	12.89	12.69	0.928
Has diabetes	12.27	12.03	0.895
Has high blood pressure	36.11	35.85	0.922
Smoked 100 cigs in lifetime	34.26	34.80	0.846
Had alcohol past 12 months	64.74	69.53	0.062
Covered by MediCal	14.07	13.46	0.802
No usual source of health care	11.52	11.99	0.829
Condition limits physical activity	21.52	20.56	0.719
Diagnosed heart disease	9.36	9.84	0.774
Blind or deaf	8.96	8.80	0.917
Has difficulty learning, remembering, etc.	14.49	13.27	0.617
Covered by Medicare	33.42	32.35	0.652
Heath rating - Fair/Poor	22.99	24.73	0.588
Feel nervous - All/Most of time	4.80	5.13	0.772
Delayed care in past 12 months	11.28	14.02	0.258
Feel save in neighborhood - Always	51.06	52.34	0.691
No fast food past week	43.34	42.05	0.663
No sex partners past 12 months	17.09	19.20	0.369
More than one sex partner past 12 months	4.66	3.85	0.415
Did not see MD past 12 months	14.90	14.41	0.824
Demographics			
Country of birth - USA	66.23	71.10	0.140
Country of birth - Mexico	14.23	13.39	0.737
Marital status - Married	59.90	57.50	0.369
Marital status - Never Married	19.17	16.25	0.294
Education - Less than HS	15.18	12.03	0.177
Education - HS grad or GED	24.37	23.20	0.680
Education - Graduate/Professional	15.34	17.28	0.313
Male	45.18	49.56	0.181
White	65.78	66.42	0.823
African-American	4.92	3.89	0.207
Hispanic	30.03	29.66	0.907
All/most calls on cell phone (Dual users)	34.45	40.09	0.220
No children (under 12) in HH	74.89	76.63	0.416
No adolescents (12-17) in HH	79.06	82.14	0.183
Income 20K or less	18.76	14.05	0.016
Income more than 135K	15.63	12.32	0.055
Age 18-29	14.39	9.42	0.074
Age 65+	31.39	31.22	0.948