# Topology of the Landline Telephone Sampling Frame

Mansour Fahimi, Dale Kulp, and David Malarek
(Marketing Systems Group)

Published in the December 2009 Issue of Survey Practice

http://surveypractice.org/2009/12/09/topology-landline-sample

## Overview

Based on a series of studies conducted by the authors during the past two years, it has been reported that the landline telephone sampling frame may no longer hold some of the properties based on which list-assisted Random Digit Dial (RDD) sampling methodology was developed. Among other fundamental changes, the digital transition of the telephone network infrastructure has undermined the relevance of the 100-series telephone number banks for construction of RDD sampling frames. Today, local telephone exchanges are not servo-mechanical in nature anymore and hence such banks no longer serve as physical building blocks for telephone number assignments. Consequently, telephone companies are now less systematic in their assignment of numbers to new customers, Tucker and Lepkowski (2008).

With departure from the above AT&T-dominated structure and emergence of various innovations in the telecommunication industry, the landline telephone sampling frame has become subject to unfolding new realities. On the one hand, the residential density of the 100-series telephone banks has been changing, which is a manifestation of a series of interrelated factors. On the other hand, with the growing number of alternative providers of voice services there are emerging issues that have confounded the operational definitions that the RDD methodology has relied upon for frame construction and identification of residential landlines.

## Introduction

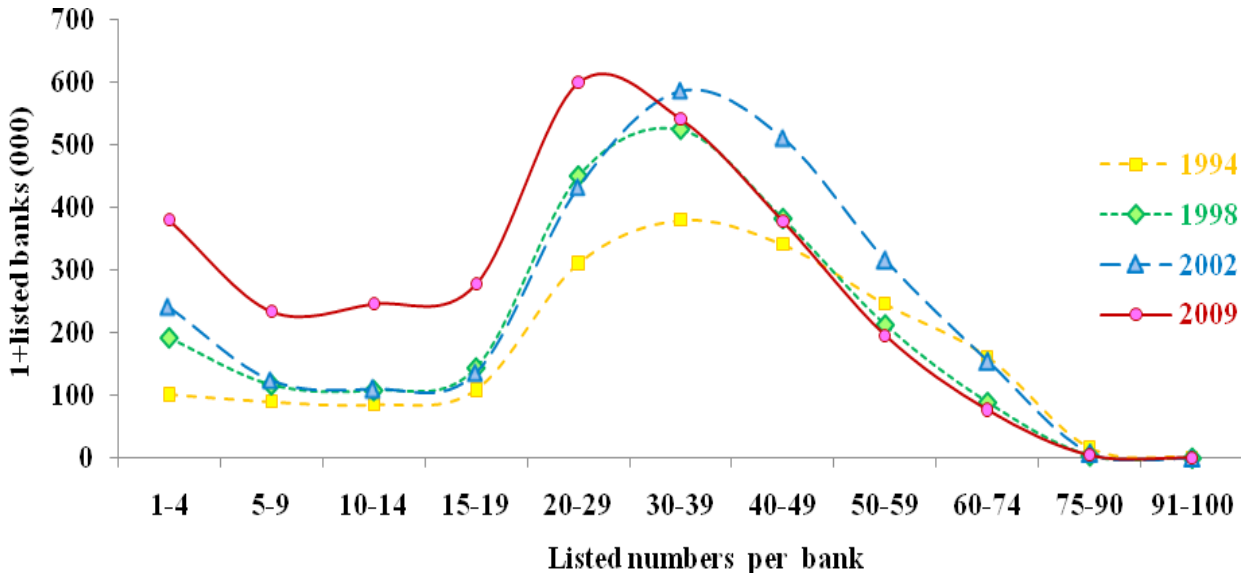
The Mitofsky-Waksberg method of RDD sampling (Waksberg 1978) was a major breakthrough in telephone survey research methodology, since it improved the efficiency of telephone sampling while allowing researchers to sample from both listed and unlisted telephone numbers. However, operational complexities led researchers to look for further refinements and examine alternative designs. In particular, Casady and Lepkowski (1993) codified a truncated design alternative (list-assisted RDD) that has been in use for a number of years, which only included telephone numbers from 100-series telephone banks with one or more listed numbers (1+listed banks). As such, a two-stage cluster sampling was replaced by a single-stage epsem sampling method that could produce survey estimates with smaller sampling variances. Of course, these impressive gains were realized at the expense of a presumed modest undercoverage that could be easily tolerated when time and cost saving considerations were kept in balance. Brick et al. (1995) estimated that only 3.7 percent of all landline telephone households were not covered in the list-assisted RDD frame at that time.

Concerned about the longevity of the above pivotal assumption in light of the unfolding changes in US telephony, the authors conducted a series of studies and reported that the residential undercoverage rate for the list-assisted RDD frame has risen sharply (Fahimi et al. 2008). However, a study by Boyle et al. (2008) suggests this rate has remained virtually unchanged during the past two decades. Interestingly, a recent study by Weiss et al. (2009) reports depending on how the dispositions of unresolved telephone numbers are imputed, the undercoverage of this frame can be subject to very high rates in certain parts of the nation. Confronted by these seemingly inconsistent conclusions, this paper provides a brief summary of the results from ongoing investigations by Marketing Systems Group to gain a better understanding of the current structure of the landline telephone sampling frame.

## Changes in the Residential Density of Telephone Banks

Tucker and Lepkowski (2008) reported that the proportion of residential numbers in 1+listed banks has been on a decline. The following figure shows how the residential distribution of such banks has changed, both with respect to the location and scale parameters, across the years. While the residential density of listed banks was much flatter in 1994 with an average of about 35 listings per bank, this distribution has become more peaked in recent years with a 2009 average of about only 26 listings per bank and a sharp increase in the number of banks with lower residential density. As will be discussed later in this section, listed banks with low residential density are particularly vulnerable to losing their listed status because in many instances it takes only a handful of listed numbers becoming unlisted for their corresponding banks to get demoted to a 0-listed status and removed from the traditional 1+listed RDD frame.

Figure 1. Changes in the residential density of the 1+listed banks via number of listings per bank by year (Source: Marketing Systems Group)



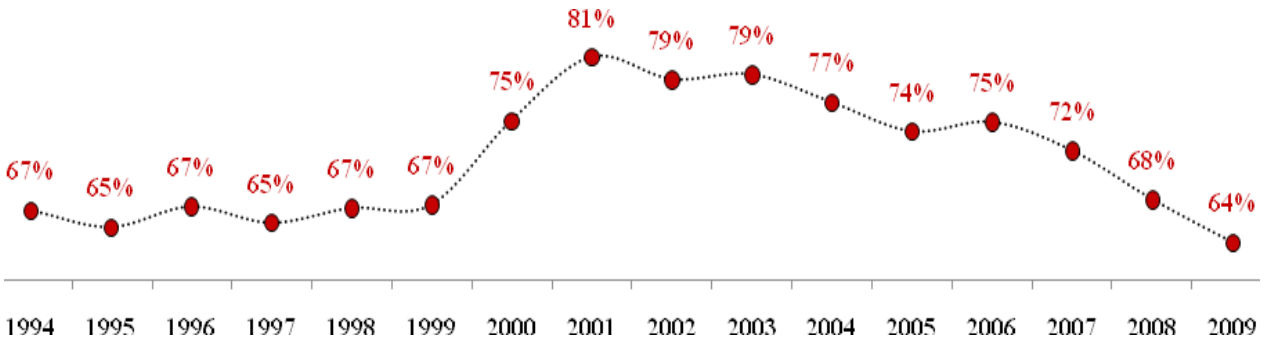

At a macro level, the above dilution in the residential density of 1+listed banks is due to the disproportional growth in the number of telephone banks as compared to the increase in number of households. Further complicating the situation, as seen from the following figure, has been the uneven growth of banks with listed and unlisted telephone numbers since 1990.

Figure 2. Number of 1+listed and unlisted banks since 1990 by year (counts in millions)


At a micro level, however, there are several interrelated factors that are germane to the changes in the residential density of 1+listed banks. These include a decline in the proportion of telephone numbers that remain listed or newly assigned residential numbers that emerge as listed from the onset of assignment, as well as an increase in the number of landline numbers

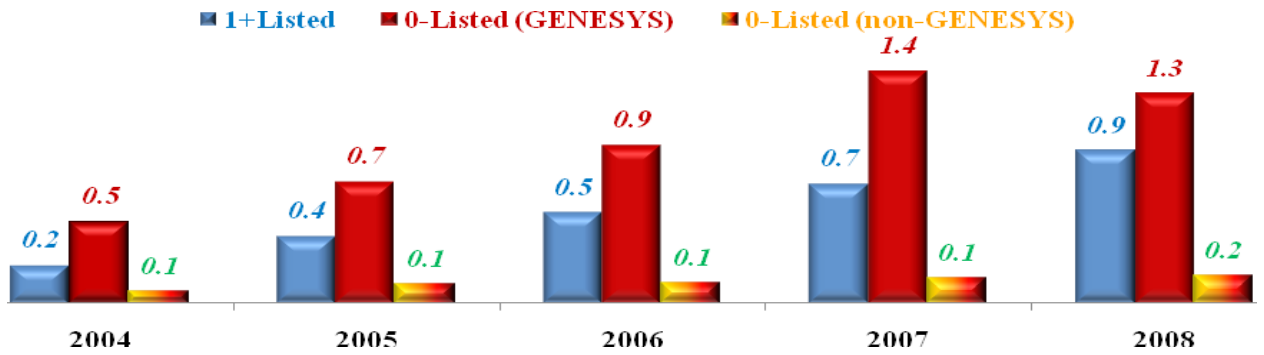
that are ported to cellular devices. These are among the reasons why some of the 1+listed banks with low residential density are becoming 0-listed and eventually removed from the 1+listed frame. The following figure shows estimated percentages of listed residential numbers since 1994, according to which in 2009 this percentage is lower than it has ever been in over a decade.

Figure 3. Estimated percentages of listed residential numbers by year since 1994 (denominator all households)




And, the following figure shows the growth in estimated number of ported landline numbers to cellular devices by bank type. Here, unlisted banks are separated into those with at least one listed number at the exchange level (0-listed GENESYS) and those with no listed numbers even at the exchange level (0-listed non-GENESYS).

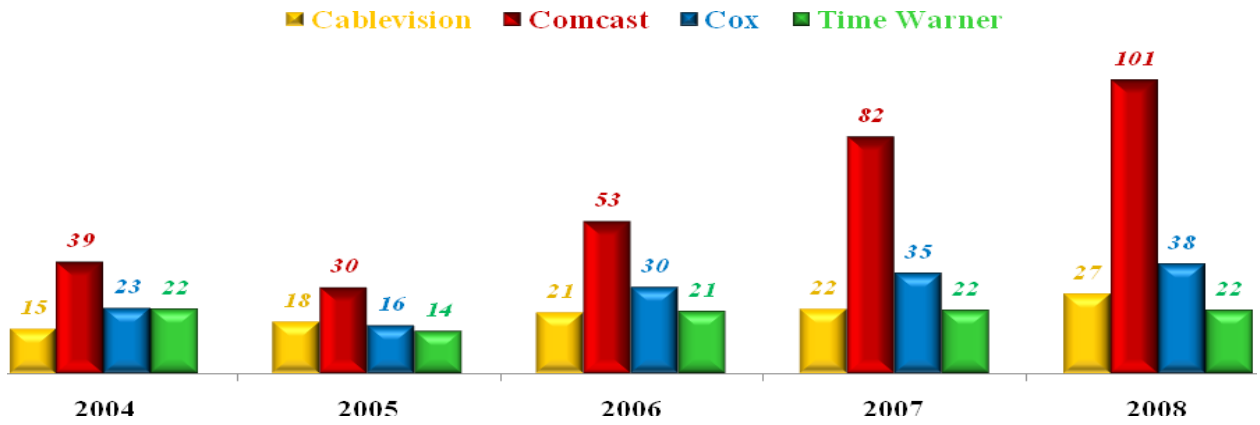
Figure 4. Estimated number of ported landline numbers to cellular devices by bank type (counts in millions)






Undoubtedly, one of the profound changes in US telephony during recent years has been the emergence of alternative service providers for voice and data, such as cable companies. For example, Competitive Local Exchange Carriers (CLECs) are now important providers of telephone service, accounting for 34% of all the residential numbers in the 0-listed GENESYS and about 72% in the 0-listed non-GENESYS banks (Fahimi et al. 2008). The following figure shows the growth in the number of 100-series banks dedicated to major cable companies over the years. For reasons yet to be fully investigated, the "time-to-listing" for these providers can vary, resulting in some residential numbers remaining unlisted until they are reported to directories.

Figure 5. Growth in the number of 100-series banks dedicated to major cable companies (counts in thousands)



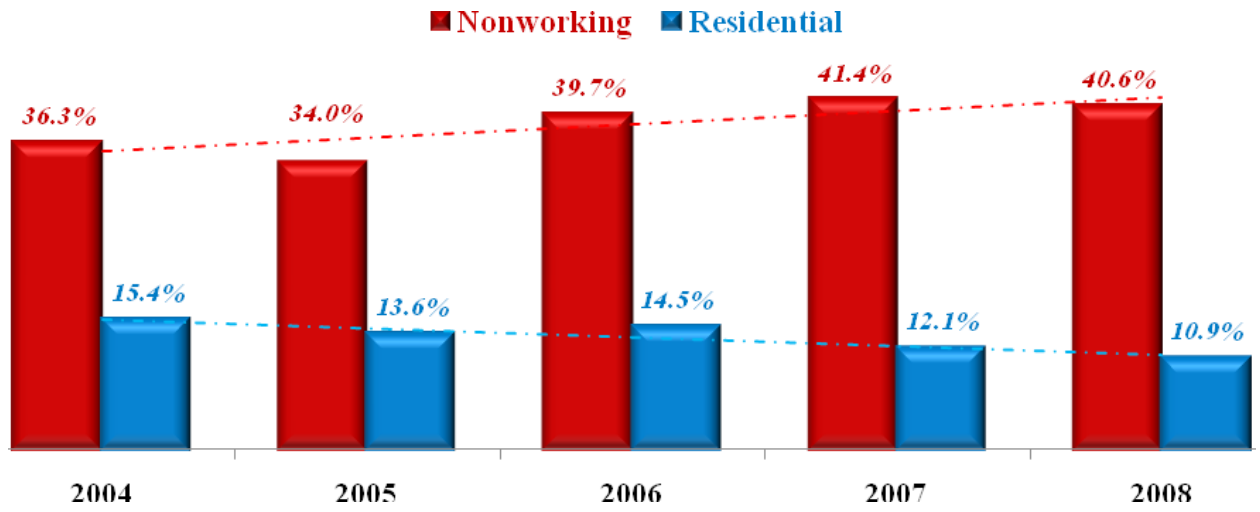

Moreover, there is a growing number of Voice over Internet Protocol (VoIP) service providers that convert voice signals to digital signals to travel over the Internet and then get converted back to voice signals at the receiving end. The website www.voip-info.org shows a list of over 180 residential VoIP providers in the US, which consists of new providers and long-time service providers such as Regional Bell Operating Companies (RBOCs) and Incumbent Local Exchange Carriers (ILECs). With many of these providers, such as Magic Jack[◊] that claims to be signing over 250,000 new subscribers every month, incoming residential telephone numbers will not be listed in any public telephone directories. That is, such numbers are unlisted and many of them end up in 0-listed 100-series banks that are outside of the traditional RDD frame. What's more, when these numbers are assigned within the listed banks they can gradually erode the residential density of such banks, eventually demoting some of them to become completely unlisted and out of the 1+listed frame. Also, with VoIP-based services users are no longer tied to local area codes.

## Concluding Remarks (for now)

The telephone network infrastructure in the US has undergone a number of fundamental changes during the past two decades, some manifestations of which are depicted in the above figures. These changes, many of which continue to unfold, can have significant ramifications for RDD sampling methodology. Most notably, the traditional 1+listed frame may no longer exclude an ignorable percentage of landline residential numbers as new residential numbers get assigned outside of this frame or existing numbers migrate from it to 0-listed banks. Predominantly, this coverage gap appears to be due to a decrease in the residential density of the 100-series banks as well as an increase in alternative dial-tone providers that have much lower listed rates for households. Recapturing this coverage will require developing sampling frames that are more inclusive, even though such expansions will entail lower residential hit rates and additional costs for screening efforts.

Estimating the extent of bias due to the above undercoverage is subject to a number of challenges, including definitional issues as well as those related to the dynamic nature of what is being measured. With the increase in the number of prefix-level assignments of landline telephone numbers for mixed use applications and the steady increase in the number of households that have multiple lines, it is now more difficult to detect residential lines or separate noncontact from nonworking dispositions for RDD samples. The following figure provides a depiction of these issues based on examination of national RDD samples generated and screened at Marketing Systems Group using the GENESYS-CSS sample screening procedure. While there is a rise in the percentage of nonworking numbers, the rate of residential numbers detected in RDD samples is on a decline. Clearly, these rates are somewhat specific to the employed screening procedure; nonetheless, the consistency of the resulting trends speaks to the stated concern.

Figure 6. Growth of nonworking and decline of residential numbers detected in national RDD samples via GENESYS-CSS screening system (denominators are the initial sample sizes)




The above complications are bound to multiply as more households come to rely on cellular, VoIP, as well as landline phone services for voice communications. Consequently, it is important to carry out tracking studies that can assess the ongoing changes in the structure of the landline telephone sampling frame and develop more effective screening procedures that can reverse the cost drain associated with the decreased residential hit rates that go hand-in-hand with more inclusive frames. Among others, it will be informative to conduct investigations that can offer a better understanding of the emerging changes introduced by the alternative service providers for voice communications.

## References


Boyle, J., M. Bucuvalas, L. Piekarski, and A. Weiss. 2009. "Zero Banks: Coverage Error in List Assisted RDD Samples." Survey Practice, January Issue, http://surveypractice.org/2009/01/29/zero-banks-coverage-error-in-list-assisted-rdd-samples/

Brick, J. M., J. Waksberg, D. Kulp, and A. Starer. 1995. "Bias in List-Assisted Telephone Samples." Public Opinion Quarterly, 59: 218-235.

Casady, R. J. and J. M. Lepkowski. 1993. "Stratified Telephone Survey Designs." Survey Methodology, 19:103-113

Fahimi, M., D. Kulp, and M. Brick. 2008. "Bias in List-Assisted 100-Series RDD Sampling." Survey Practice, August Issue, http://surveypractice.org/2008/09/25/bias-in-list-assisted-100-series-rd

Mitofsky, W. 1970. "Sampling of Telephone Households." Unpublished CBS News memorandum.

Tucker, C. Lepkowski, J. M. (2008). Advances in Telephone Survey Methodology. New York: John Wiley & Sons, Inc.

Waksberg, Joseph. 1978. "Sampling Methods for Random Digit Dialing." Journal of the American Statistical Association 73:40–46.

Weiss, A., M. Battaglia, J. Boyle, A. Hyon, and D. Kulp. 2009. "Unlisted Banks in New York City: Coverage Error and Bias in Urban Areas from RDD Samples Based on Hundreds Banks with Listed Numbers." Presented at the American Association for Public Opinion Research, Hollywood, FL.