Are Cell Phone Samples Needed for Studies of Walking Activity?

Ugo Lachapelle, Marc D. Weiner, and Robert B. Noland

The growth in cell phone–only households represents a challenge for the collection of survey data. Cell phone–only households have distinct sociodemographic characteristics, which may result in different travel behavior. To explore those differences, as well as to investigate the impact of including a cell phone component in active transportation research, a representative sample of New Jersey households was surveyed with a random digit dial survey that included 1,200 completed interviews (800 based on a statewide landline sample, 400 from a landline oversample of Jersey City) and 311 statewide cell phone interviews, of which 80 were cell phone–only respondents. The survey explored walking behavior and perceived characteristics of the pedestrian environment. Sociodemographic characteristics, the frequency of walking, and home location characteristics were compared with chi-square tests of significance between sample pairs as well as multivariate analysis (ordered probit). Cell phone–only respondents were typically younger and poorer, with a greater proportion of renters, carless households, and minorities. It was found that cell phone–only household members walked more frequently, but this finding was because of their distinct sociodemographic characteristics, not their cell phone use per se. The implication for any analysis of rates or trends in walking (and probably other travel behavior) is that cell phone–only households must be included through a cell phone sample supplementing a landline sample. However, in the absence of a cell phone supplement, multivariate analysis of the correlates of walking may not be overly biased if sociodemographics relevant to cell phone–only respondents are collected and included in the analysis.

The widespread introduction of and growing reliance on cell phones poses a new challenge to random digit dial (RDD) surveys, which are used for many transportation studies (*1–5*). A growing share of households no longer own landlines and instead rely entirely on cell phones, whereas others maintain a near-vestigial landline and primarily use their cell phones. Cell phones were first introduced to the consumer market in the early to mid-1990s. By 2005 cell phone–only households represented 8.4% of the U.S. population and have been steadily increasing (*3*). By 2008, that number was estimated to be 20.2%, the equivalent of 41 million adults and 14 million children under 18. By 2010, cell phone–only households represented 26.6% of American households (*6, 7*). At an estimated 115 million American households, this percentage amounts to 30,590,000 cell phone–only households or, with a mean household size of 2.6 persons, represents 79,534,000 Americans who cannot be contacted by traditional landline RDD survey contacting protocols (*8*).

The research question addressed here is whether this cell phone gap matters in terms of research on walking behavior and, if so, how? Pedestrian behavior has generated growing interest in both transportation and physical activity research. Health researchers are interested in the health benefits of an active lifestyle (*9*), and transportation practitioners seek to reduce vehicle travel and find ways to increase pedestrian accessibility to destinations, usually through land use, urban design measures, and other changes to the built environment (*9, 10*). Walking is particularly important in urban settings as an access mode or as a link to public transit (*11*).

Because the cost of conducting cell phone surveys is much greater than that for landline surveys, it is important to evaluate the extent to which landline samples capture a representative cross section of the population and, conversely, whether information relevant to the research question is lost by the exclusion of a cell phone sample. The objective of this study is to compare different subsamples of an integrated-dual-frame RDD survey. Based on a 2-year survey research effort, the analysis here includes a comparison of sociodemographic characteristics, walking behavior, and home location characteristics of three main subsamples: New Jersey statewide residents ($n = 800$), an oversample of urban Jersey City residents ($n = 400$), and a statewide cell phone sample ($n = 311$) consisting of cell-only ($n = 80$) and cell-and-landline respondents ($n = 231$). A cell phone sample enables proper representation of the 18- to 30-year-old group that is typically underrepresented in landline-only RDD phone surveys but that is more likely to walk than other age groups. Most important, this analysis helps researchers assess potential information loss when cell phone components are omitted for transportation research sampling protocols.

RESEARCH QUESTIONS AND HYPOTHESES

How do cell phone–only respondents differ from other telephone survey respondents? Do they walk more frequently? Do they live in areas with distinct built-environment characteristics? From previous research, it is hypothesized that cell phone–only households are distinct in sociodemographic composition and that they walk more frequently than the population reached by landlines. Finally, because a higher proportion of cell phone–only households is expected to be renters, they would be more likely to live near more central areas such as central business districts (CBDs) and have greater access to transit.

U. Lachapelle, Département d'Études Urbaines et Touristiques, École des Sciences de la Gestion, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-Ville, Montréal, Québec H3C 3P8, Canada. M. D. Weiner, Bloustein Center for Survey Research, and R. B. Noland, Alan M. Voorhees Transportation Center, Edward J. Bloustein School of Planning and Public Policy, Rutgers University, 33 Livingston Avenue, New Brunswick, NJ 08901. Corresponding author: U. Lachapelle, lachapelle.ugo@uqam.ca.

Transportation Research Record: Journal of the Transportation Research Board, No. 2285, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 56–65.
DOI: 10.3141/2285-07

Existing literature on trends in cell phone use is first reviewed, followed by a review of research on walking behavior. The sampling strategy and survey instrument are described and compared by using univariate analysis. Results are further confirmed in a multivariate analysis. Implications for crafting research protocols and sampling designs for research in transportation are discussed in the conclusion.

TRENDS IN CELL PHONE USE

Although cell phones have been patented since the 1970s, their widespread use began in the 1990s and they became ubiquitously distributed in the past 10 years (*12*). According to the Federal Communications Commission, there were approximately 24 million cell phone subscriptions in the United States in 1994 and 270 million in 2008 (*12*). In 2010, this number surpassed 300 million, with 302 million subscribers. In 2009 only 14.9% of households had landline service only, and 24.5% were cell phone–only households (*12,* Table 7.4).

The National Health Interview Survey (NHIS) has become the survey research community's leading reference resource for landline and cell phone coverage estimates. NHIS interviews are conducted in person and thus reach the designated sample without reference to telephone ownership status and completely disentangle the telephone status from the research protocol. In addition to substantive health data, the interviewer records whether the respondent's household has landlines and cell phones. Beginning in 2007, that survey started including a question on which phone household members used the most, allowing the identification of "cell phone–predominant" households (*7*).

On the basis of this survey, as of 2010 approximately 60% of the U.S. population has access to both a cell phone and a landline. One important reason for this statistic is that individuals subscribing to landlines are often required to register a landline as part of an Internet or cable deal (*13*), whether they use it or do not use it. Many report using the cell phone primarily, even when they have a landline. Over 24% of households with both cell phones and landlines were considered wireless mostly (cell phone–predominant) households. They made up nearly 15% of all U.S. households (*3*). Cell phone–predominant households, despite having a landline, receive most of their calls on their cell phones.

For the purpose of this study, a cell-only respondent is defined as a respondent who does not have any means of telephone communication other than at least one cell phone. In a recent survey conducted by the Pew Internet and American Life Project, 23% of Americans were considered cell-only (*13*). The proportion of cell-only households has also been growing at a fast pace. According to the Cellular Telecommunications Internet Association (*6*), cell phone–only households went from 8.4% of American households in 2005 to 26.6% in 2010. In 2008, cell phone–only households consisted of 20% of households, corresponding to 18% of the total population (*3*). Together, cell phone–only and cell phone–mostly households now represent nearly a third of all U.S. households (*3, 14*). Omitting such a sizeable proportion of the population from a sampling plan, especially if the proportion is known to have characteristics different from those of the general population, may significantly bias survey estimates and may potentially provide inaccurate estimates of the determinants of travel behavior.

Because of the lack of representation of cell phone–only and cell phone–mostly households, as well as the exclusion process of certain numbers in list-assisted RDD, survey coverage may capture less than 70% of all households in the United States (*14*). This proportion has considerable implications for the statistical validity and reliability of the data. Coverage bias may exist if persons with and without landlines are different with respect to the variables of interest (*2*).

RELATIVE COSTS OF LANDLINE VERSUS CELL PHONE SAMPLING

One of the challenges in conducting surveys is to balance tradeoffs between study costs and sample precision. Including a cell phone component considerably increases survey costs: a cell phone interview costs roughly two-and-a-half to five times the cost of a landline interview primarily because interviews take longer and often respondents are paid a cash incentive, ostensibly to compensate for the cost of the air time. Interviews take longer because they typically require more dialing to reach respondents, require more screening time, and have more quota failures (respondents who do not fit the study's inclusion criteria, for example, being 18 or older), and the sample frame of cell phone numbers is generally not screened for known business or out-of-service numbers, so more numbers are dialed than in a sample frame precleaned of these nonproductive numbers (*14–16*). All of these factors increase the cost per completed interview. If one is interested in a particular geographic area, costs can be even greater, since cell phone users may have nongeographically defined numbers, especially if they have moved between regions and maintained the same cell phone number.

Similar issues apply to many voice over Internet protocol services, which may even extend numbers beyond international boundaries. A detailed comparison of cost per completed interview (*1*) shows that the cost per sampled telephone number and the cost per completed interview were, respectively, $10.85 and $64.25 for landline surveys, $5.79 and $74.18 for cell phone surveys, and $5.10 and $195.78 for cell-only households reached after screening. The difference in cost for completed surveys is striking and attests to the importance of balancing cost versus sampling and data collection precision in the design of transportation studies.

A compelling illustration in the transportation literature of concern for proper survey sampling can be found in research by Sen et al. (*14*), who compare two sampling strategies: the active contact method (cell phone) with the passive contact method (mail surveys) in terms of efficiency, data collection effort, response rate, and cost per interview of different sampling strategies. Sen and colleagues found that cell phone sampling involved more data collection effort but in turn yielded a higher response rate. Their RDD cell phone survey reported 42% of cell-only households and 58% of combined cell-and-landline households. The mail survey, however, reported 30% of cell-only households and 40% of cell-and-landline households. Hence, RDD cell phone samples were more likely to capture cell-only and cell-landline households than a mail survey. Mail surveys, however, captured a more comprehensive coverage including cell-only, landline-only, a mix of both, and no-phone households. Data collection efforts for the cell phone survey were more extensive than for address-based surveys, and response rates were higher for cell RDD (19%) than for mail surveys (8%) (*14*). The current authors found no research on address-based samples that employ telephone matching.

CELL PHONE SAMPLING IN EXISTING SURVEYS: DEMOGRAPHIC IMPLICATIONS

Since at least 2001, various government-funded surveys relevant to transportation and walking behavior have incorporated cell phone sampling to assist in dual-frame (i.e., combined landline and cell phone) weighting. Two key questions are relevant: how have these major surveys adapted their sampling plans to accommodate the advent and rise of cell phone–only and cell phone–mostly households, and is it useful to use the cell phone–only and cell phone–mostly categories as analytical categories in their own right? The evidence suggests that the inclusion of a cell phone sample is specific to the research question at hand, and telephone-use status is a proxy that captures the differential adaptation of various demographic segments to changing technology.

The National Household Travel Survey (NHTS), sponsored by FHWA, is an extensive nationwide computer-assisted telephone interview survey that uses list-assisted RDD to collect data about the travel behavior of American households. To its credit, the survey was early in exploring the impact of cell phones on survey research; the NHTS included questions on cell phone ownership for the first time in its 2001 field administration. In 2009, for the first time, the NHTS included a cell phone sample frame as a test of methods (*17*). The survey team justified this inclusion by suggesting the need to understand if travel patterns of cell phone–only households were significantly different from those of households reached via their landlines. This sample allowed survey sponsors "to determine the feasibility of conducting the NHTS interview by cell phone, and also provided some data for research on the differences in demographic characteristics and travel behavior between households that have landlines and those that have only cell phones" (*17*). These data and the results of any analysis are not yet publicly available since they are still being analyzed by FHWA.

In the health literature, two large-scale survey efforts implemented by the Centers for Disease Control and Prevention (CDC) continue to generate evidence on cell phone use from a physical activity perspective: the NHIS and the Behavioral Risk Factor Surveillance System (BRFSS).

Every 3 months, the CDC releases estimates for 15 key health indicators using the NHIS, a face-to-face interview survey that captures information on wide-ranging health and personal data; in 2003 the NHIS began to probe household telephone access and usage. In a comparison of health outcomes across the telephone use categories, wireless-only households were more likely to binge drink and smoke but also more likely to report an excellent or very good health status and to engage in regular leisure-time physical activity. They were also less likely to have ever been diagnosed with diabetes (*3, 7*). Given knowledge of the demographics of cell-only households, these variations are likely functions of the age distribution across the categories of telephone users.

BRFSS implemented a cell phone component in all states and territories in 2009. BRFSS is a nationwide health survey with a physical activity component and different modules that can be added at the request of states. In their comparison of the prevalence of obesity in the 2000 BRFSS and the 2000 National Health and Nutrition Examination Survey, Yun et al. (*18*) suggested that the increase in cell phone–only households raised the need to reconsider the validity of the BRFSS contacting protocol to track trends in obesity. This development prompted the CDC to expand their methodology to rely on dual-frame (i.e., combined landline and cell phone) samples. The 2008 version of BRFSS expanded the landline sample frame to a dual-frame sample in 18 states as a pilot study (*16*) and moved to a full-scale dual-frame sample of all 50 states in 2009. Again, differences in health-related behavior such as smoking, binge drinking, and engaging in physical activity were found between cell phone respondents and landline respondents (*16*) with cell phone respondents being significantly more physically active based on univariate analysis. The telephone usage category is a proxy for other more dominant demographic characteristics. Indeed, once sociodemographic characteristics were controlled for, the relationship between phone use category and likelihood of active transportation was no longer statistically significant.

In their analysis of landline and cell phone samples of public opinion surveys, Link et al. (*1*) found that compared with landline-only samples, cell phone–only samples were more likely to be men, African–American, Hispanic, under the age of 34, employed, of lower income, and not married. Zuwallack (*4*) found similar results in his dual-frame survey sample; cell phone–only households were younger and had a higher proportion of minorities. These are some of the same groups that are typically underrepresented in landline surveys because of differential nonresponse, the lower propensity of low-income population to answer surveys (*1*). Similar findings are presented by Blumberg and Luke (*7*) in their analysis of the NHIS with the addition of renters, residents of the Midwest, and adults living with unrelated adult roommates.

From these efforts the authors suspect that effects ostensibly attributable to telephone use status are actually the function of sociodemographic factors, particularly age, housing, urbanicity, and employment.

DATA AND METHODS

Sampling

A 2-year survey was conducted; in November 2009, 1,200 completed landline interviews were collected, 800 from an area-code-proportional statewide survey of New Jersey households and 400 from an oversample of Jersey City; in November 2010, 311 New Jersey statewide cell phone interviews were drawn from a cell phone frame of which 80 were cell phone–only respondents. The survey explored walking, sociodemographics, and perceived characteristics of the pedestrian environment. Weather conditions were similar during both field periods.

The rationale for oversampling Jersey City was that more potential walk-accessible destinations are expected to be present in reasonable proximity in large urban centers such as Jersey City. This sample also provides another point of comparison to assess the statewide cell phone sample frame against an urbanized population. The basic eligibility criterion was defined as being 18 years of age or older. Eligible participants for the Jersey City oversample had lived in Jersey City for more than 1 year. To be part of the cell phone sample, respondents had to have been reached on a cell phone. An integrated dual-frame sampling was used; for the 2010 sample, it was assumed that households for which a completed interview was obtained via cell phone were cell phone–predominant households. This sampling is consistent with estimates from the NHIS that suggest that approximately 25% of households with both cell phones and landlines predominantly use their cell phones (*3*).

Response rates, calculated by using the American Association of Public Opinion Researchers Approach 3, were 20.9% for the 2009 statewide landline sample, 19.9% for the 2009 landline Jersey City oversample, and 23.3% for the companion 2010 cell phone sample.

The rates were calculated for each sample by using the following equations (*19*):

$$\text{RR3} = \frac{I}{\left[(I+P)+(R+NC+O)+e(\text{UH}+\text{UO})\right]}$$

$$e = \frac{(I+P+R+NC+O)}{\left[(I+P+R+NC+O)+\text{NE}\right]}$$

where

I = complete interviews (and screenouts),
P = partial interviews,
R = refusals and breakoffs,
NC = noncontacts,
O = other,
e = estimated eligibility of unknowns,
UH = unknown households,
UO = unknown other, and
NE = not eligible.

Weighting schemas were calculated separately for each sample with an [(age × sex) × race] function, and analyses were run with and without weights. This sample closely matched U.S. estimates for 2010 (*7*). A Spanish language option was available and about 5% of all interviews were conducted in Spanish. The cell phone sample collected was limited in size because of budget constraints. It would have been preferable to obtain a larger sample to enable more subgroup analysis. However, this limitation does not have any implications for the analysis that follows, which has robust and useful results.

Modeling and Analysis

The key dependent variable is the frequency of walking during the past month; from the six original possible answers, the responses were coded into four categories: "more than once a day," "once a day," "several times a week," and "no more than once a week." The independent variables fall into two categories: sociodemographic predictors and location and built-environment measures. The household-reporting respondent (the informant) was asked to report her or his ethnicity, age, education, and gender as well as household information including number of children if any, housing type, rent versus own, and car ownership. The household's self-reported total annual income was coded into five categories (see Table 1). Working

TABLE 1 Sample Description

	Landline Sample Frame		Cell Phone Sample Frame		
Variable	Statewide (%)	Jersey City (%)	Cell and Landline (dual service) (%)	Cell Phone Only (%)	Total (%)
Renter	20.00	58.75	25.11	52.50	32.76
Minorities	32.25	71.25	43.29	57.50	45.60
Women	53.63	53.25	54.55	50.00	53.47
Has children	35.63	39.75	47.62	45.00	39.05
Carless household	7.73	32.89	5.88	18.99	14.69
Household income					
<$25,000	12.72	27.61	6.15	34.85	16.86
$25,000 ≤ income < $50,000	18.92	22.39	17.95	40.91	20.89
$50,000 ≤ income < $100,000	34.82	26.69	36.41	15.15	31.83
$100,000 ≤ income < $150,000	17.33	9.51	22.05	4.55	15.30
>$150,000	16.22	13.80	17.44	4.55	15.13
Age (years)					
18 to 30	7.12	15.18	28.18	44.74	14.62
31 to 40	14.79	21.95	15.00	18.42	16.92
41 to 55	31.64	28.46	33.18	23.68	30.61
56 to 70	27.95	23.04	19.09	11.84	24.37
71 and older	18.49	11.38	4.55	1.32	13.48
Education					
High school or less	24.77	32.47	25.66	40.51	27.79
Less than a college degree	26.58	21.13	28.32	35.44	25.89
College degree or more	48.65	46.39	46.02	24.05	46.32
Lives in single-family home	73.50	18.50	69.70	38.75	56.52
Employed full time	46.63	50.75	54.98	42.50	48.78
Goes to school	1.25	4.00	7.79	10.00	3.44
Has CBD within 10-min walk	41.88	48.50	41.99	42.50	43.68
Has transit stop within 10-min walk	46.13	82.75	47.62	55.00	56.52
Frequency of walking					
Less than weekly	13.87	5.06	12.44	5.26	10.86
Several times a week	31.65	21.07	25.35	25.00	27.51
Once a day	23.11	20.79	17.05	17.11	21.20
More than once a day	31.37	53.09	45.16	52.63	40.43

full-time and going to school were also considered dichotomous variables. Of particular interest, in light of the body of research on the enabling effect of built environments on walking (*9*), were respondents' self-reported dichotomous measures of ≤10-min walk access to their municipality's CBD and to a public transit stop or station.

With the screening questions, a sample indicator variable identifying the different subsets of the New Jersey samples was created: "statewide landline, 2009;" "Jersey City landline oversample, 2009;" "cell phone and landline, 2010;" and "cell phone–only, 2010." Each subset's sociodemographic characteristics and walking behavior were compared. Preliminary univariate tests of significance of difference using chi-square tests for pairs of samples were conducted as follows: statewide landline versus cell phone only; Jersey City landline versus cell phone only; and cell-and-landline (dual-service households in cell phone sample frame) versus cell phone–only sample.

Frequency of walking was then modeled in a multivariate framework with ordered probit models. Indicator variables for sample type were assessed while sociodemographic characteristics were controlled for. The reference category was the statewide landline. A positive association between sample indicator and dependent variable would suggest that once sociodemographic characteristics were accounted for, being part of the cell phone sample drives up the mean walking frequency. Analyses were conducted with STATA 11 with and without survey weights; inclusion or omission of weights did not substantively affect the results. Weighted estimates are provided.

Respondents were asked to report the nearest intersection to their home and the municipality and zip code where they resided. With this information, completed interviews were mapped by subsample to visualize their distribution within the state of New Jersey. In Figure 1, a three-panel map shows that the statewide landline and cell phone sample respondents are generally well distributed throughout the state, matching up with the 2000 census municipal-level population density. Tertiles of population density were used as a backdrop to show where populations concentrate. Thus, visually there is no systematic variation in where these samples reside compared with the general population.

A chi-square test comparing the statewide sample to 1) the cell-and-landline sample and 2) the cell phone–only sample, across tertiles of population density, is provided in Table 2. Cell phone–only samples were significantly more likely to be found in higher-density areas than the statewide landline sample. The cell-and-landline sample was not significantly different from the statewide sample.

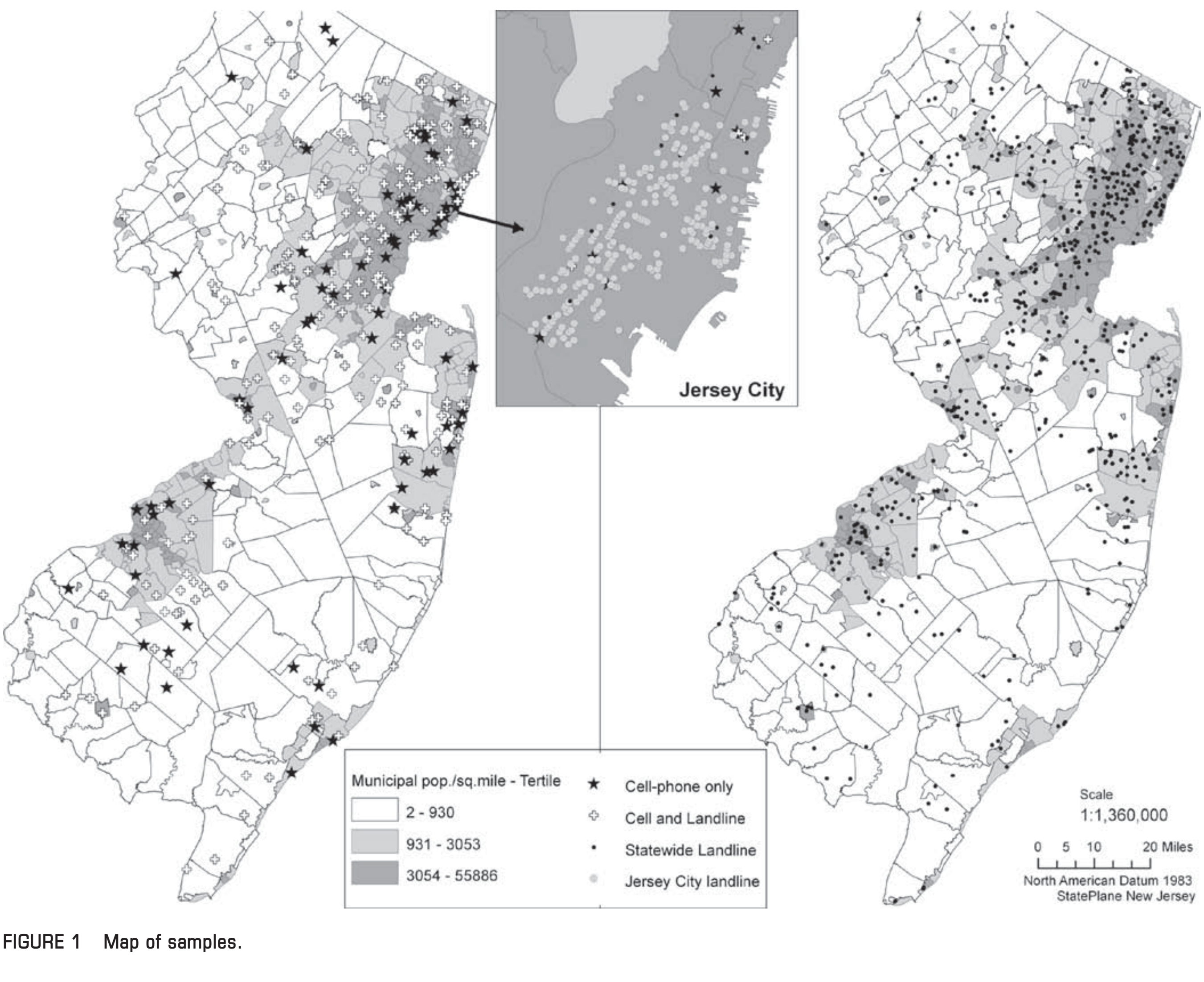

FIGURE 1 Map of samples.

TABLE 2 Assessing Differences in Distribution of Each Sample Across Tertiles of Density with Statewide Landline Sample as Reference Category

Population Density Tertile	Statewide Landline (%)	Cell and Landline (dual service) (%)	Cell Phone Only (%)	Total (%)
Low	46.50	50.56	40.98	46.36
Medium	46.50	42.13	40.98	45.51
High	6.99	7.30	18.03	8.13
Total	100.00	100.00	100.00	100.00
Pearson chi-square (2 df) between statewide landline and . . .	—	1.0932	9.3248	—
Significance	—	.579	.009	—

NOTE: df = degrees of freedom; — = not applicable.

UNIVARIATE ANALYSES

The sociodemographic characteristics of each sample are presented in Table 1. Pearson chi-square tests of significance for pairs of samples across sociodemographics, housing and environment, and walking frequency are presented in Table 3. The cell phone and combined cell phone and landline sample frame had respondents who were younger, more often renters, students, minorities, and carless respondents compared with the statewide landline sample.

Cell phone–only respondents had lower household incomes, had fewer households with children, were less likely to be women, and were less educated than the other samples. The proportion of cell phone–only respondents without a car was also considerably higher than the state average but much lower than that for the Jersey City sample. With respect to residential location, roughly the same proportion of cell phone–only households lived within a 10-min walk of a CBD as well as closer to transit stops or stations. They were also much less likely to live in single-family homes as opposed to apartment buildings and other multifamily residences.

TABLE 3 Chi-Square Test of Significance Between Pairs of Samples

Variable	Cell Phone–Only Sample Versus: Statewide Landline (p-values)	Jersey City Landline (p-values)	Cell and Landline (dual service) (p-values)
Renter	.000	.302	.000
Minorities	.000	.015	.028
Women	.536	.595	.482
Has children	.097	.383	.686
Carless household	.001	.015	.001
Household income	.000	.002	.000
Age	.000	.000	.038
Education	.000	.001	.002
Lives in single-family home	.000	.000	.000
Employed full-time	.480	.178	.054
Goes to school	.000	.025	.539
Has CBD within 10-min walk	.914	.327	.937
Has transit stop within 10-min walk	.129	.000	.255
Frequency of walking	.002	.832	.332

Are cell phone users actually more active than others, or is this relationship merely captured by differences in group composition? Table 3 provides chi-square tests of significance for cell phone–only respondents paired with other subsamples. Cell phone–only respondents walked considerably more frequently than the statewide landline sample and about as much as Jersey City respondents or as the cell-and-landline households. Column 1 provides chi-square significance levels for a comparison between the cell phone–only sample and the New Jersey statewide landline sample. Both samples were not significantly different in terms of gender, employment status, having children, being employed full-time, and distance to transit and the CBD. The samples were significantly different on all other characteristics, including the frequency of walking.

Column 2 provides significance levels for a comparison between the Jersey City landline sample and the cell phone–only sample. Again, there were no significant differences between samples for gender, employment status, having children, and percent renters. There were also no significant differences in the frequency of walking.

In Column 3, the cell-and-landline sample is compared with the cell phone–only households. Gender, going to school, and having children were not significantly different across the two categories within the cell phone sample frame. There was also no difference in the frequency of walking.

MULTIVARIATE ANALYSIS

The reported frequency of walking during the past month was modeled in a multivariate framework; results are shown in Tables 4 and 5. Results of multivariate ordered probit regressions are presented for the entire sample for which all variables were available. The frequency of walking was modeled as a function of the variables that were significantly different between groups of interest. Model 1 tests associations between the frequency of walking and the sample type categories. Each sociodemographic characteristic was tested individually along with the sample type indicators (Models 2 through 13) and final Models 14 and 15 combine all variables. Significant control variables that changed the coefficient of the cell phone sample variable are interpreted as capturing portions of the effect of cell phone use relevant to walking frequency. Survey weights were used.

DISCUSSION OF RESULTS

With the statewide landline survey as a reference category, all three other samples, including the cell phone–only sample, were significantly positively associated with the frequency of walking in Model 1. The largest coefficients, as expected, were for Jersey City (more short-walk trips are feasible in denser urban areas). The significant positive relationship of cell phone–only households was maintained even when sociodemographic variables were introduced one by one in subsequent models (Models 2 through 13). Model 14 provides estimates when all significant sociodemographic characteristics are included. Being a renter, all age groups below 71, and having a CBD and a transit stop within 10 min from home were all individually positively associated with more frequent walk trips when the sample types were controlled

TABLE 4 Weighted Model Estimates for Frequency of Walking: Models 1–7

Model Includes	Model 1	Model 2 + Renter	Model 3 + Minorities	Model 4 + Women	Model 5 + Children	Model 6 + Carless	Model 7 + Income
Sample type							
Statewide landline 2009 [reference]							
Jersey City landline	0.499*	0.423*	0.418	0.499*	0.499*	0.493*	0.471*
Cell and landline	0.236**	0.230**	0.219**	0.237**	0.238**	0.236**	0.238**
Cell phone–only sample	0.336***	0.274**	0.291**	0.332***	0.338***	0.333***	0.271**
Renter	—	0.190*	—	—	—	—	—
Minorities	—	—	0.171*	—	—	—	—
Women	—	—	—	–0.06	—	—	—
Has children	—	—	—	—	–0.035	—	—
Carless household	—	—	—	—	—	0.026	—
Household income							
<$25,000 [reference]	—	—	—	—	—	—	—
$25,000 ≤ income < $50,000	—	—	—	—	—	—	–0.01
$50,000 ≤ income < $100,000	—	—	—	—	—	—	–0.225*
$100,000 ≤ income < $150,000	—	—	—	—	—	—	–0.196
≥$150,000	—	—	—	—	—	—	0.038
Age (years)							
18 to 30	—	—	—	—	—	—	—
31 to 40	—	—	—	—	—	—	—
41 to 55	—	—	—	—	—	—	—
56 to 70	—	—	—	—	—	—	—
71 and older [reference]	—	—	—	—	—	—	—
Education							
High school or less [reference]	—	—	—	—	—	—	—
Less than college degree	—	—	—	—	—	—	—
College degree or more	—	—	—	—	—	—	—
Lives in single-family home	—	—	—	—	—	—	—
Employed full-time	—	—	—	—	—	—	—
Goes to school	—	—	—	—	—	—	—
Has CBD within 10-min walk	—	—	—	—	—	—	—
Has transit stop within 10-min walk	—	—	—	—	—	—	—
Cut-point 1	–1.076***	–1.037***	–1.034***	–1.107***	–1.092***	–1.074***	–1.194***
Cut-point 2	–0.130**	–0.089	–0.086	–0.160**	–0.145**	–0.128**	–0.242*
Cut-point 3	0.400***	0.443***	0.446***	0.370***	0.385***	0.402***	0.291**
Observations	1,062	1,062	1,062	1,062	1,062	1,062	1,062
II (base)[a]	–1,536.4	–1,536.4	–1,536.4	–1,536.4	–1,536.4	–1,536.4	–1,536.4
II (model)[b]	–1,525.4	–1,522.2	–1,522.5	–1,525.0	–1,525.3	–1,525.5	–1,519.6
Chi-square	21.9	28.5	27.9	22.8	22.2	22	33.7
Pseudo-R^2	.007	.009	.009	.007	.007	.007	.011

NOTE: Results for Models 8–15 appear in Table 5. All models present significant improvement from the constant-only model at the $p < .001$ level. — = not applicable.
[a]II (base) = log likelihood of constant-only model.
[b]II (model) = log likelihood of fitted model.
*$p < .05$; **$p < .01$; ***$p < .001$.

for. Being a woman and living in a single-family home were both individually negatively associated with the frequency of walking, but these findings were not significant.

Model 15 only used significant or theoretically important variables. In Model 15, only the youngest age category and the walking distance to a transit stop or station were still positively associated with walking frequency once other variables were controlled for. The cell phone–only coefficient was no longer significantly associated with the frequency of walking but remained positive, albeit considerably lower. This finding suggests that any distinct effect on walking behavior associated with cell phone–only individuals is captured by their different sociodemographic characteristics and the location of the respondents' households. Once combined in Models 14 and 15, the sociodemographic characteristics associated with cell phone–only users absorbed the effect of the sample indicators. Thus, although the omission of a cell phone sample will be problematic for measuring trends in these travel behavior variables, integrating sociodemographic and location characteristics of the cell phone sample in multivariate analysis seems to adequately control for any sample-driven differences.

The objective of this work was to determine whether the sociodemographic characteristics and walking patterns of different sampling frames varied. This analysis provides important insights into the potential measurement errors in phone surveys conducted without cell phone samples. The analysis suggests that cell phone samples have distinct sociodemographic characteristics and walking patterns. These respondents came from lower-income households, were less educated and younger, were more likely renters not living

TABLE 5 Weighted Model Estimates for Frequency of Walking: Models 8–15

Model Includes	Model 8 + Age	Model 9 + Education	Model 10 + Single-Family Home	Model 11 + Employed Full Time	Model 12 + CBD in 10-min Walk	Model 13 + Transit Stop in 10-min Walk	Model 14 All Control Variables	Model 15 All Controls Except Children, Education, and Going to School
Sample Type								
Statewide landline 2009 [reference]								
Jersey City landline	0.446*	0.503*	0.411	0.491*	0.494*	0.402	0.275	0.284
Cell and landline	0.136	0.232**	0.235**	0.215**	0.233**	0.241**	0.128	0.143
Cell phone-only sample	0.201*	0.313***	0.287**	0.309***	0.345***	0.316***	0.082	0.11
Renter	—	—	—	—	—	—	0.049	0.049
Minorities	—	—	—	—	—	—	0.081	0.063
Women	—	—	—	—	—	—	−0.057	−0.068
Has children	—	—	—	—	—	—	−0.122	—
Carless household	—	—	—	—	—	—	−0.159	−0.159
Household income								
<$25,000 [reference]	—	—	—	—	—	—	—	—
$25,000 ≤ income < $50,000	—	—	—	—	—	—	−0.07	−0.044
$50,000 ≤ income < $100,000	—	—	—	—	—	—	−0.237	−0.214
$100,000 ≤ income < $150,000	—	—	—	—	—	—	−0.189	−0.17
≥$150,000	—	—	—	—	—	—	0.047	0.059
Age (years)								
18 to 30	0.546***	—	—	—	—	—	0.483**	0.448**
31 to 40	0.238	—	—	—	—	—	0.232	0.181
41 to 55	0.266*	—	—	—	—	—	0.244	0.225
56 to 70	0.235	—	—	—	—	—	0.218	0.219
71 and older [reference]	—	—	—	—	—	—	—	—
Education								
High school or less [reference]	—	—	—	—	—	—	—	—
Less than college degree	—	0.005	—	—	—	—	0.036	—
College degree or more	—	−0.096	—	—	—	—	−0.06	—
Lives in single-family home	—	—	−0.156*	—	—	—	−0.105	−0.112
Employed full-time	—	—	—	0.333*	—	—	0.088	0.141
Goes to school	—	—	—	—	—	—	0.194	—
Has CBD within 10-min walk	—	—	—	—	0.203**	—	0.108	0.109
Has transit stop within 10-min walk	—	—	—	—	—	0.254***	0.169*	0.185*
Cut-point 1	−0.833***	−1.122***	−1.194***	−1.071***	−0.994***	−0.966***	−0.977***	−0.958***
Cut-point 2	0.115	−0.176*	−0.245***	−0.125**	−0.043	−0.016	−0.013	0.003
Cut-point 3	0.650***	0.355***	0.286***	0.407***	0.488***	0.518***	0.533**	0.547**
Observations	1,062	1,062	1,062	1,062	1,062	1,062	1,062	1,062
II (base)[a]	−1,536.43	−1,536.4	−1,536.4	−1,536.4	−1,536.4	−1,536.4	−1,536.4	−1,536.4
II (model)[b]	−1,516.51	−1,524.3	−1,523.1	−1,523.3	−1,520.6	−1,517.8	−1,497.0	−1,499.6
Chi-square	39.8	24.2	26.7	26.3	31.7	37.4	78.9	73.6
Pseudo-R^2	0.013	0.008	0.009	0.009	0.01	0.012	0.026	0.024

NOTE: Results for Models 1–7 appear in Tables 4. All models present significant improvement from the constant-only model at the $p < .001$ level. — = not applicable.
$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

in single-family homes, and also tended to live closer to CBDs and to transit stops or stations. Differences in gender were not significant. In univariate analysis, these respondents walked more frequently than statewide landline users did, but not as frequently as the oversampled residents of Jersey City, an urban area with considerably higher density than the rest of the state.

However, once sociodemographic characteristics were controlled for, the walking patterns of respondents selected from a cell phone sample frame, whether they had a landline or did not, were not significantly different from other respondents. Hence, the effect was largely driven by the different sociodemographic characteristics of the samples.

For the purpose of calculating inferential statistics on the correlates of walking activity, these results suggest that not having a cell phone sample supplement may be acceptable and should not overly affect estimates provided there is adequate variation in the sample to capture the demographic distributions that would be collected were a cell phone sample supplement included and provided these relevant sociodemographic characteristics are collected. However, because of the difficulty of reaching younger and poorer respondents by traditional landlines, it will become increasingly difficult to generate adequate demographic distributions with landline-only sampling. As such, if the purpose of a survey is to determine travel trends and rates across a population, supplementing a landline sample frame with a cell phone sample frame is necessary to represent the population accurately. This interpretation concurs with work by Hu et al. (*16*), who found that health surveys carried out by telephone require a dual frame of landline and cell phone numbers to provide reliable and representative estimates of rates, trends, and prevalence. This requirement is particularly important in research on pedestrians for two reasons: first, sampling of pedestrians is typically made harder by the fact that the incidence rates are relatively low, especially when surveys assess specific travel purposes (*10*); and, second, because those more likely to engage in walking often have the same characteristics as those found in cell phone–only households.

Researchers should be cautious and particularly wary about using landline surveys to draw inferences about subpopulations that are more likely to be wireless only (*7*). Because of the demographics of cell phone–only households, research focusing on social disparities and research on the health or transportation consequences of poverty should consider including cell phone–only and cell phone sample frames to more accurately capture these underrepresented groups.

In their essay on improving research on walking and bicycling, Krizek et al. (*10*) underscore the importance of clear conceptualization, sound research design, measurement innovation, and strategic sampling. Whether a cell phone sample will be taken may be a strategic decision that researchers should consider carefully, and approaches should be tailored to the different age groups and income strata expected to participate in an activity. Although caller ID features and voice mail may lower response rates of cell phones, some believe that in the long run, cell phones may make survey respondents more accessible to researchers (*5*). Understanding the implications of this growing trend is necessary to conduct meaningful and representative survey research in this day and age.

CONCLUSIONS

Cell phone–only respondents were typically found to be younger, renters, from carless households, nonwhite minorities, and to earn lower income. They also tended to walk more frequently than did landline-using households. However, once controls for the sociodemographic characteristics of the cell phone–only sample were included in a multivariate analysis, the differences became nonsignificant. The distinct sociodemographic characteristics of cell phone–only households are associated with more walking, but cell phone users do not otherwise differ fundamentally in their residential location patterns or walking behavior.

While for descriptive and analytical purposes a cell phone frame makes a properly drawn and executed probability sample more representative, the costs are not insignificant. Hence, researchers should carefully examine their research questions and sample inclusion criteria in light of available resources to make a firm determination of the necessity of including a cell phone supplement in an RDD landline sample. As cell phone households become dominant, including cell phone sample supplements will become more and more necessary, particularly for studies targeting minorities and low-income populations.

Cell phone use is a pervasive and growing trend that influences the way telephone survey sampling is conducted. Whether in transportation planning or in health research, researchers need to consider seriously the impact of omitting a cell phone sample supplement from the typical RDD sample frame survey. These findings suggest that in order to identify trends in the population or calculate prevalence of walking and other physical activity, except under limited circumstances, researchers should deploy dual-frame samples to collect data from cell phone–only, cell phone–mostly, and landline telephone users. As communication technologies continue to evolve, researchers will need to be aware of the changing telephone landscape and how this may affect their desired sampling plans.

ACKNOWLEDGMENTS

Funding for survey design, deployment, and analysis was provided by the New Jersey Department of Transportation. The authors thank M. P. Simon and M. K. Simon-Weiner for encouragement and support.

REFERENCES

1. Link, M. W., M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad. Reaching the US Cell-phone Generation. *Public Opinion Quarterly,* Vol. 71, No. 5, 2007, p. 814.
2. Blumberg, S. J., and J. V. Luke. Reevaluating the Need for Concern Regarding Noncoverage Bias in Landline Surveys. *American Journal of Public Health,* Vol. 99, No. 10, 2009, p. 1806.
3. Blumberg, S. J., and J. V. Luke. *Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2008.* National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Md., 2009.
4. Zuwallack, R. Piloting Data Collection Via Cell-phones: Results, Experiences, and Lessons Learned. *Field Methods,* Vol. 21, No. 4, 2009, p. 388.
5. Kempf, A. M., and P. L. Remington. New Challenges for Telephone Survey Research in the Twenty-First Century. *Annual Review of Public Health,* Vol. 28, 2007, pp. 113–126.
6. *U.S. Wireless Quick Facts.* CTIA—The Wireless Association, Washington, D.C. http://www.ctia.org/media/industry_info/index.cfm/AID/10323. Accessed May 31, 2011.
7. Blumberg, S. J., and J. V. Luke. *Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January–June 2010.* National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Md., 2010. http://www.Cdc.gov/nchs/nhis.Htm.
8. *Quickfacts.* U.S. Census Bureau. quickfactscensus.gov/gfd/states/00000.html. Accessed June 26, 2011.

9. *TRB Special Report 282: Does the Built Environment Influence Physical Activity? Examining the Evidence.* Transportation Research Board of the National Academies, Washington, D.C., 2005.
10. Krizek, K. J., S. L. Handy, and A. Forsyth. Explaining Changes in Walking and Bicycling Behavior: Challenges for Transportation Research. *Environment and Planning B: Planning & Design,* Vol. 36, No. 4, 2009, pp. 725–740.
11. Hoback, A., S. Anderson, and U. Dutta. True Walking Distance to Transit. *Transportation Planning and Technology,* Vol. 31, No. 6, 2008, pp. 681–692.
12. *Trends in Telephone Service.* Federal Communications Commission, Washington, D.C., 2010.
13. Purcell, K., R. Entner, and N. Henderson. *The Rise of Apps Culture.* Pew Internet & American Life Project, 2010.
14. Sen, S., J. C. Zmud, and C. Arce. Efficiency and Effectiveness of Cellphone Samples in Transportation Surveys. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2105,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 44–50.
15. Miles, D. R., H. Herrick, and C. A. Ford. The North Carolina Child Health Assessment and Monitoring Program: Survey Methodology and Data Collection. *Statistical Primer,* No. 18, 2010, pp. 1–14.
16. Hu, S. S., L. Balluz, M. P. Battaglia, and M. R. Frankel. Improving Public Health Surveillance using a Dual-Frame Survey of Landline and Cellphone Numbers. *American Journal of Epidemiology,* Vol. 173, No. 6, 2011, pp. 703–711.
17. *2009 National Household Travel Survey User's Guide.* FHWA, U.S. Department of Transportation, 2011.
18. Yun, S., B. P. Zhu, W. Black, and R. C. Brownson. A Comparison of National Estimates of Obesity Prevalence from the Behavioral Risk Factor Surveillance System and the National Health and Nutrition Examination Survey. *International Journal of Obesity,* Vol. 30, No. 1, 2006, pp. 164–170.
19. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys,* 6th ed. American Association for Public Opinion Research, Lenaxa, Kans., 2009.

The statements or conclusions in this paper do not represent the policy or opinion of the New Jersey Department of Transportation. Any errors or omissions are the responsibility of the authors.

The Travel Survey Methods Committee peer-reviewed this paper.