Florida Department of Transportation (FDOT)
Customer Survey Research Analysis

August 23, 2013

Dr. Jason Barabas*

Dr. Mary Stutzman**

Carlos Paez***

Executive Summary

Many scholars and researchers believe it is more difficult than in the past to conduct high quality surveys. Of particular concern are nonresponse trends in which survey participation rates drop to low levels, especially among some demographic subgroups (e.g., minorities and younger respondents). This report explores strategies the survey research field has investigated to address underrepresentation issues. Specifically, the report (1) identifies strategies to make surveys more descriptively representative, (2) evaluates the relative tradeoffs of the main alternatives, and (3) conducts a pilot study to further assess the costs, feasibility, and outcomes associated with alternative survey strategies. Based upon these efforts, a series of recommendations are made for future FDOT Customer Satisfaction Surveys. In particular, greater use of the mail mode and available sample information as well as continued experimentation with the internet are recommended to address underrepresentation. Use of oversampling, repeated attempts for individuals in selected demographic categories, and more creative survey designs are also worth consideration.

* Associate Professor, Department of Political Science, Florida State University
** Director, Survey Research Laboratory, Florida State University
*** M.A., Department of Political Science, Florida State University

Summary of Problem

The Florida Department of Transportation (FDOT) has conducted biennial Customer Satisfaction Surveys of Florida residents since 2001. Past surveys have typically utilized telephone calls to landlines with a goal of 400 complete interviews within each of seven geographic districts. In recent years, FDOT experienced underrepresentation with respect to younger residents (ages 18 to 34) and minority groups (e.g., Hispanics, African-Americans). In other words, respondents who complete the customer surveys have been older and less diverse than the state population as a whole according to U.S. Census Bureau background data. Gender diversity may also be an issue. The underrepresentation bias likely stems from new communication technologies such as cell phones and caller identification as well as respondent fatigue from increased attempts to contact individuals by marketing and advertising companies (e.g., Toepoel, Das, and Van Soest 2008). As a result of these factors and others, traditional telephone surveys with landline samples might not be the best way to gather a representative sample of Florida residents to assess their views.

The purpose of this report is not to investigate past surveys per se, although the pilot study utilizes many of the same questions and it targets one of the seven transportation districts (i.e., District 5 near Orlando) so limited comparisons will be possible. Instead, we review empirical trends and findings in the scholarly literature that have implications for future FDOT surveys, especially as they relate to the representativeness of the surveys. In the sections below, we take up the issue of response rates and nonresponse patterns before continuing to survey data collection modes, sampling frames and strategies, as well as weighting and estimation procedures. The sections that follow review the pros and cons of the various options as well as some results from a pilot study before making a series of recommendations.

Trends in the Survey Methodology Literature

There are many reasons why surveys of any type (e.g., telephone, mail, internet) could produce questionable (i.e., biased) results. As such, when they focus on issues of data quality, methodologists encourage practitioners and end-users to consider the "total survey error" perspective (Groves 1989; Groves and Lyberg 2010; Weisberg 2005). This perspective reminds us that everything from sampling methods to subtle variations in question wording could threaten the validity of survey data.[1] However, this report focuses on the choices related to the mode of data collection and accompanying issues of representativeness without considering other important parts of the survey process such as questionnaire design or survey administration (e.g., Holbrook et al. 2009; Malhotra 2008; 2009; Malhotra, Krosnick, and Thomas 2009; Smyth et al. 2006). When it comes to mode and representativeness, a logical place to begin is with the issue of survey response rates and patterns of nonresponse.

[1] Regrettably, the total survey error approach is more of a checklist of possible threats rather than a single quantifiable score over the various dimensions. Therefore, total error quantification is not feasible. However, survey methodologists are trying to develop representation metrics, which are discussed later.

Response Rates and Nonresponse Patterns

Few topics have garnered more attention recently than the decision to participate in a survey. The summary statistic for measuring survey participation is a "response rate," which (broadly speaking) is the percentage of people who respond to a survey out of all those who were invited and could have responded. There are several different methods of calculating response rate statistics (American Association for Public Opinion Research 2011), but nearly all tell the same story—response rates have been dropping precipitously. This is true irrespective of survey mode and it has been occurring virtually everywhere across the globe.

More specifically, several recent reports have focused squarely on the issue of survey nonresponse. The National Research Council's Committee on National Statistics released a report in early 2013 titled, "Nonresponse in Social Science Surveys: A Research Agenda." The Pew Research Center issued a report, "Assessing the Representativeness of Public Opinion Surveys" (Kohut et al. 2012). Likewise, the Annals of the American Academy of Political and Social Science published an edited volume titled, "The Nonresponse Challenge to Surveys and Statistics" (Massey and Tourganeau 2013).

The consensus view to emerge from these reports is that response rates have declined (e.g., Singer 2006; Blom and Kreuter 2011; Groves and Couper 1998; deLeuw and deHeer 2002; Groves et al. 2002; Stoop 2005). More specifically, in the 1970s and 1980s, response rates above 60% were not unheard of in random digit dial surveys; more recently, though, the response rates for a typical telephone survey dropped to under 10% (Kohut et al. 2012). In fact, one study suggests that non-response is increasing by roughly 0.5 percentage points per year (Brick and Williams 2013). Another study put the figure higher, potentially double that rate, especially for telephone surveys (Curtin, Presser, and Singer 2005). In an alarming fashion, nonresponse rates have increased, "in all types of cross-sectional surveys, with nothing to suggest that the trend has plateaued," and with important implications for the FDOT surveys, the NRC report suggests "…the data also clearly show that the recent rates of increase in nonresponse have been substantially greater for RDD [Random Digit Dial] telephone surveys than for face-to-face surveys" (National Research Council 2013, 1-16).[2] Also, this decline is not specific to the U.S. Similar declines in response have been observed in all advanced countries (National Research Council 2013, s-1).

Part of the reason behind the increase in nonresponse stems from new communication technologies which have made it easier for respondents to evade interview requests (National Research Council 2013, 4-5; Brick and Williams 2009). However, there has also been a big increase in the number of surveys and survey questions (Presser and McCulloch 2011), which contributes to respondent fatigue. People are being asked for their opinions more than ever before. Moreover, potential respondents are also distrustful of survey requests because they fear identity theft; the

[2] Specifically, RDD surveys experience reductions in response at nearly "three times" the rate of face-to-face surveys.

worry is that if they reveal too much about themselves or their families, they might help thieves make fraudulent financial transactions in their names or expose themselves to other mischievous acts (Presser and Singer 2007).

Four variables have been found to be highly correlated with nonresponse rates: the percentage of families with children under age six, single person households, the violent crime rate of the surrounding area, and travel time to work (National Research Council 2013, 1-22; Brick and Williams 2013). However, the direction of the associations can be surprising. People who have families with children under 6 or who live in areas with high violent crime rates have *higher* rates of survey participation than those who do not. The findings might be the case because these individuals are home more often and hence more likely to be available for a survey. In contrast, respondents living in single-person households and those who have a travel time to work of 45 minutes or more tend to participate in surveys *less* (i.e., more survey nonresponse), consistent with the hypothesis that they are "too busy to respond" (Brick and Williams 2013, p. 49). Still, these factors may be markers for some other differences. In other words, Brick and Williams (2013) point out that "many of these variables are themselves highly correlated" and it may be that some other factor is causing factors like having a long commute and the propensity to respond.

Consequently, the increase in nonresponse begs the question of why it is occurring independent of technological changes; in other words, why would someone choose not to respond to a survey? There is no consensus in the literature, but several have proposed theories of nonresponse. One explanation is that deciding to respond to a survey is a cost-benefit analysis in the sense that "People respond to surveys when they conclude that the rewards outweigh the costs" (National Research Council 2013, 1-2). More specifically, two main theories have been proposed. Social capital theory, forwarded by Robert Putnam (1995; 2001), suggests a variety of factors at the societal and individual level that make interaction with others (as would occur in several survey modes) less likely. There are indicators like networking opportunities and reduced levels of trust. One study supporting the social capital perspective, by Abraham, Maitland, and Banichi (2006), found that people who are weakly integrated into their communities are less likely to respond because they are less likely to be contacted. Yet, Brick and Williams (2013) have argued for a rigorous investigation of social capital theory and nonresponse rates (p. 20).

A second perspective is leverage-saliency theory. Groves et al. (2004; also see Maynard et al. 2010) argue that people vary in the importance they assign to the survey request (Groves et al. 2012). More specifically, individuals decide to participate if the expected utility is higher than not responding. Consequently, there are a variety of aspects which can make a survey more appealing (e.g., cash incentives, shorter interviews) and people might react differently to these at various points in their life. There may also be ethnic and racial influences on a person's assessment of the survey (National Research Council 2013, 1-24). For example, immigrants might be wary of responding to a survey from a branch of the state government. However, one of the most relevant considerations is whether the topic is deemed interesting by the potential

respondent. Groves, Presser, and Dipko (2004) argue that the odds of cooperating with a survey request are roughly 40 percent higher for topics that are of interest to them (also see Groves et al. 2006).[3]

Irrespective of the causes of survey nonresponse, the broader concern related to nonresponse rates involves what it means for the resulting representativeness of a population being sampled. Indeed, underrepresentation in a sample can produce biased estimates that will lead to erroneous inference and conclusions about the subject of interest.[4] The Pew Research Center (Kohut et al. 2012) has undertaken efforts to measure the impact of growing nonresponse in RDD telephone surveys; somewhat surprisingly, the Pew team found reason for cautious optimism. They compared results from a standard five-day survey with results from a "rigorous" survey conducted over a longer period of time. In the end, even though the response rates for the standard survey were only 9% (compared with 22% for a "high effort survey"),[5] the Pew researchers found little evidence suggesting that unit nonresponse within the range of response rates obtained seriously threatens the quality of survey estimates. In particular, the results showed that across dozens of comparable items, the standard survey and high effort variants produced estimates that were statistically indistinguishable.[6]

However, while the standard survey approximated U.S. Census and other large government survey estimates for a variety of outcomes, they did unearth some evidence of education and participatory biases among those who responded to the standard survey (Kohut et al. 2012). That is, standard survey respondents were more highly educated and participated more than the U.S. population in government surveys. The biases were in the double digits (+11 percentage points for college graduation and +28 percentage points for volunteering for an organization or +21 percentage points for contacting public officials).

Nevertheless, one problem underlying conclusions about nonresponse bias in surveys is that defining the degree of bias is difficult (e.g., Fricker and Tourangeau

---

[3] A third perspective which blends elements of the others is social exchange theory. Some of the costs might be psychological as well as monetary, such as trust and conformity to group norms (Dillman 1978; 1991; 1999).

[4] It is important to note, however, that a number of recent studies have suggested that there is no consistent relationship between survey response rates and bias in survey estimates (see Curtin, Presser, and Singer 2000; Keeter et al. 2000; Merkle and Edelman 2002; Groves 2006; also see Olson 2006 or Peytchev, Baxter, and Carley-Baxter 2009).

[5] High efforts included a long field period, use of trained expert interviewers, increased cash incentives, additional call-back attempts, attempts to convert refusals with cash incentives, etc. Both standard and high-effort surveys included cell phone samples. For both types of surveys, the cell phone portions of the samples had lower response rates than the individuals contacted via landlines (standard survey: 10% landline vs. 7% cell; high-effort: 27% landline vs. 16% cell).

[6] Data from standard and high-effort surveys were weighted. The seven items where the standard survey samples appeared to be different showed that traditional survey respondents were A) less likely to have voted in 2002, B) less likely to be Republican, C) less likely to have 3+ adults in the household, D) less likely to be registered to vote, E) less likely to trust most people, F) less favorable in their opinion of Jews, and G) more likely to be ideologically moderate (See Table 3, Keeter et al., 2006, p. 763).

2010; Wagner 2012). In particular, how unrepresentative any sample is depends on what researchers use as a benchmark. As the National Research Council 2013 reports, "…response rates can be misleading as measures of survey representativeness. The fact that response rates have fallen…means only that the potential for nonresponse bias has increased, not necessarily that nonresponse bias has become more of a problem" (also see Peytchev 2013).[7] These issues will be discussed in more depth later in the report, but all of this underscores the need for detailed knowledge on the population of interest. Indeed, while it is hard to deduce problems with survey representativeness after the survey has already been completed, with a forward looking design ("prospective") and high quality data collection effort it might be possible to construct some of the more recently developed indicators of representativeness, such as representativity indicators (van der Grijn et al. 2006; Cobben and Schouten 2007), missing at random indicators (Wagner 2008; Andridge and Little 2009), or balance indicators (Sarndal 2011). However, it is important to note that many of these techniques are not fully developed or even available yet as they often appear in unpublished conference papers or public addresses.

It is also the case that the decline in response rates threatens the validity of inferential statistics based upon data collected through survey (National Research Council 2013, ix; Peytchev 2012).[8] That is, many of the inferential statistics used in survey research are premised upon random sampling, not samples in which the respondents self-select. All of this demands more attention to the tradeoffs from various choices related to the mode of the survey interview.

Survey Data Collection Modes

Survey mode refers to the communication style or technique used to query respondents. The primary options available to survey researchers when it comes to mode are telephone, mail, and the internet. As of the early part of the 21st Century, telephone survey methods predominate but these techniques are under assault for their increasingly low response rates (e.g., Blumberg and Luke 2007; Boyle et al. 2009). Increasingly, researchers are returning to techniques based upon surveys sent via mail since the internet and phone methods each have their own weaknesses (e.g., Heerwegh and Lossveldt 2008). Here we focus on the rise of cell phones and what that has meant for representation issues in the context of telephone surveys.

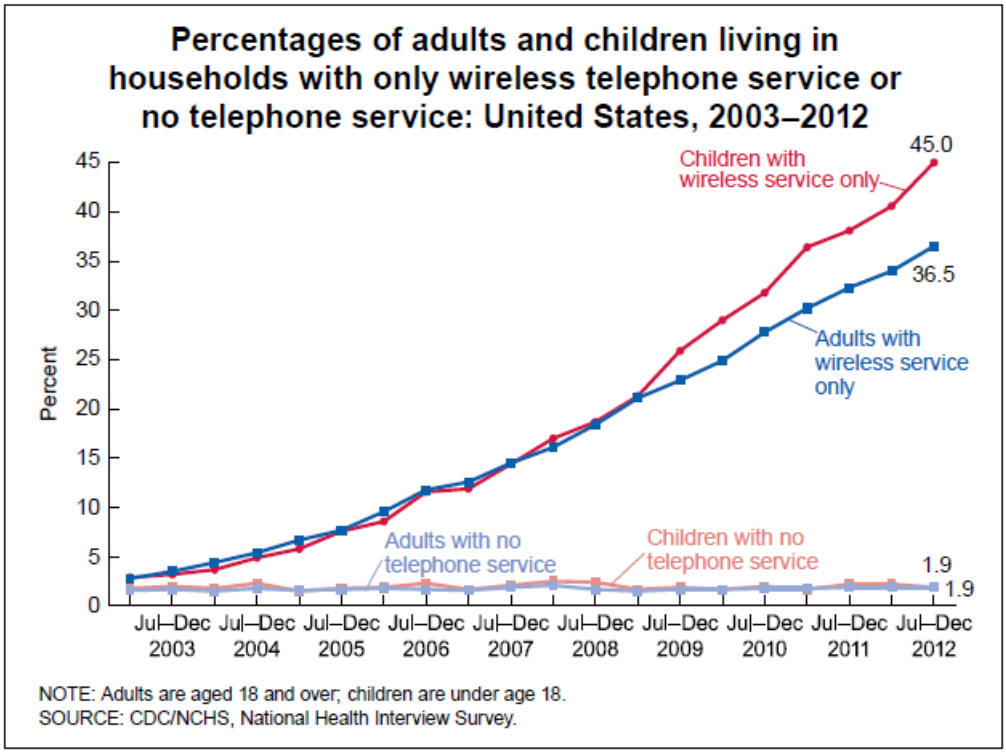
Some of the most reliable statistics on cell phone usage have come from studies sponsored by the Centers for Disease Control (CDC). In a recent CDC report, Blumberg and Luke (2013, p. 1) provide details on the federal government's most up-to-date

[7] Indeed, a recommendation from the National Research Council (Recommendation 2-1) is that "Research is needed on the relationship between nonresponse rates and nonresponse bias and on the variables that determine when such a relationship is likely."

[8] There are some ways to increase response rates, such as advanced letters ahead of the actual survey (National Research Council 2013, 1-11) or financial incentives (Singer and Ye 2013), but some of those methods are not practical or possible for FDOT surveys. However, Medway and Fulton (2012) found that having a government sponsor on the survey can increase response (see Table 2), something which will be a feature of the pilot study later (also see Groves et al. 2012).

estimates of cell phone only ("CPO") homes in the America. The authors note that during the second half of 2012, over one-third of American households had no landlines and used cell phones exclusively.[9] The trends for adults and children are depicted in the figure below. The child (i.e., under 18) figures are instructive as they suggest younger populations are even more dependent upon cell phones.

Figure 1. Blumberg & Luke 2013 Estimates of Cell-Phone Only (CPO) Population



Other highlights from Blumberg and Luke (2013, p. 2) report on cell phone usage included the following:

- Six in 10 adults aged 25 to 29 (62.1%) lived in households with only wireless telephones. This rate is greater than the rates for adults aged 18 to 24 (53.2%) or 30 to 34 (56.7%);

- Adults living in the South (39.7%) were more likely than adults living in the Northeast (23.6%) to be living in households with only wireless telephones. This is a marked increase from the first six months of 2009 when the percentage of wireless only adults in the South was estimated at 25.0% (Blumberg and Luke 2013, Table 2). State and county-specific estimates for Florida and neighboring states are shown in the appendix, but the overall state average for cell-phone only in 2011 was 34.4% with a standard error of 1.2 percent, almost exactly what

[9] In a different report using state-by-state estimates from the Marketing Systems Group, the cell phone only population in Florida was estimated at 38.2%, higher than the US average (34%) as well as neighboring states such as Georgia (34.6%) or Alabama (27.7%).

it was for the neighboring states of Alabama and Georgia (Luke et al. 2012, Table 1);

- Hispanic adults (50.5%) were more likely than non-Hispanic white adults (32.9%) or non-Hispanic black adults (39.0%) to be living in households with only wireless telephones, and;

- Among households with both landline and wireless telephones, 31.4% received all or almost all calls on the wireless telephones, based on data for the period July–December 2012. These wireless-mostly households make up 15.9% of all households.

This last point about the wireless-mostly population has been echoed in previous studies. Tucker, Brick, and Meekins (2007) found that approximately 16% of American homes received a great majority – if not all – of their calls on cell phones, including households with a landline. Their study explored the potential for biases in RDD surveys resulting from the increases in cell phones by presenting estimates of the percentage of households with different types of telephone service, including the percentage of cell-only households, and giving demographic profiles of households by type of telephone service.

In a more recent study, Ansolabehere and Schaffner (2010) investigated how the CPO population growth in the past several years is a cause of concern for researchers who depend mostly on landlines to conduct public opinion studies. The authors noted that CPO populations may differ from landline households in other ways besides age. Their evidence supports a significant conclusion: CPO households are not only a function of age, but of other factors like family structure and residential mobility.

Ansolabehere and Schaffner's findings parallelled those of other studies like Blumberg and Luke (2007). In their article exploring coverage of low-income and young adults in traditional telephone surveys, Blumberg and Luke brought to light a number of issues relating to bias. In particular, the authors found evidence suggesting that in 2006, CPO households encompassed: (1) 17% of adults with incomes below 200% of the federal poverty thresholds, (2) 25% of individuals between the ages of 18–29, and (3) 32% of low-income young adults. Moreover, Blumberg and Luke noted that traditional phone surveys significantly underestimated younger respondent health behaviors and health care coverage/access, even after controlling for relevant demographic differences. Still, as Mokrzycki, Keeter, and Kennedy (2009) point out in their research, although evidence demonstrates that CPO status seems to be the norm for individuals under age 30, respondents older than 30 actually displayed a faster rate of change from landlines to using only cell phones.[10]

[10] There were attitudinal and behavioral differences as well. The difference in presidential vote preference between the cell-only and landline-accessible voters in 30 + age group – the focus of the authors' work – was even greater than it was for younger voters.

Even as recently as the mid-2000s there was evidence to suggest that cell-phone populations did not respond differently on surveys. For example, Keeter et al. (2007, 773) argue that landline-only surveys produce almost identical results as compared to combined cell phone and landline surveys (for dual frame studies, see Brick et al. 2006); yet, even they acknowledged that while the non-coverage problem does not seem to alter estimates for the entire population, "it does produce biased estimates on certain variables for young adults, 25% of who are cell phone-only" (also see Link et al., 2007). All in all, survey researchers are trying to accommodate to the ever-changing culture of technology. In the same manner that landline surveys revolutionized sampling methods just three decades ago, cell phone surveys are making a significant change in the survey field.

Even though many people are switching to cell phones, survey researchers face formidable obstacles in surveying the cell-phone population. In particular, cell phone numbers are portable, meaning that someone who starts with a number in one area can bring it with them as they move to a different location. For instance, a study published in 2013 suggests that many people have cell phone numbers that are not reflective of where they actually live. In particular, Skalland and Khare (2013) report that, about 11.5 percent of cell-phone only adults reside in a state that differs from the state associated with their cell phone number. This number is consistent with previous estimates of 12 percent state-level geographic inaccuracy for the cell-phone only population found by Christian, Dimock, and Keeter (2009) and Benford et al. 2012 in their AAPOR conference papers. Importantly, the biases take two forms. While Skalland and Khare (2013) found that 11.5 percent of adults contacted by cell-phone in a national study reside in a different state (i.e., "overcoverage"), the same percentage reside in the state but do not appear on any given state's cell phone sampling frame (this is called "undercoverage"). For Florida, the estimates are 9.8% overcoverage of those who live in a different state compared with 8.0% undercoverage of those who live in the state but are not included. Skalland and Khare (2013) report that respondents with inaccurate sampling states tend to be younger (i.e., age 18 to 29), highly educated (college or above relative to less than h.s. graduate), and live in houses that they do not own.

At the county-level, geographic inaccuracy of cell-phones is much more problematic. While the Skalland and Khare (2013) study did not go beyond the state level, Christian, Dimock, and Keeter (2009) found that 43 percent geographic inaccuracy of the cell phone only (i.e., wireless only) adult population. Benford et al. (2012) found an almost identical geographic inaccuracy percentage (44%) at the county-level in their study (also see Montgomery et al. 2011). In contrast, Christian, Dimock, and Keeter (2009) found that only 1 percent of landline adults resided in a state that differs from the state associated with the landline phone number.

The implications of cell phone geographic inaccuracy are considerable for studies like the FDOT customer surveys that attempt sample individuals within districts that follow county boundaries. That is, studies that are sensitive to where individuals actually live must attempt to account for geographic inaccuracy bias when polling via cell phones. As Christian, Dimock, and Keeter (2009) write:

> The mobile nature of wireless phones creates a significant problem for geographic sampling, particularly as the size of the area being surveyed gets smaller. Because numbers are not associated with physical addresses, there can be a large amount of error in the geographic information associated with cell phone numbers. To address this issue, respondents who do not live in the area may be identified and screened out of the survey, but people who live in the area but have cell phone numbers from a different area will not be covered in the sampling frame. The size of the error is relatively small at the regional and state level but compromises the ability to accurately sample smaller geographic areas.

Skalland and Khare (2013) go further. They write, "…researchers must not assume that cell phone users actually reside in that state; instead, researchers should collect the state of residence from respondents and screen out those who reside in a different state. This will lead to increased survey costs, because more cell phone numbers will need to be sampled and dialed to complete the target number of interviews for residents of the state" (p. 53-5). They also expect the geographic inaccuracy trends to worsen as more individuals drop landline phone service. As a remedy, they urge researchers to consider sampling techniques based upon physical addresses. Skalland and Khare conclude, "In light of these challenges, many researchers are turning to address-based sampling (ABS) as an alternative to RDD telephone surveys (Link et al. 2008; Iannachione 2011). Because the sampling unit is an address, there will be virtually no geographic sampling inaccuracy in an ABS design, and the issues related to geographic sampling inaccuracy discussed in this article would not be present" (2013, p. 65). As such, the pilot study later will employ an address-based sampling system that should be less prone to geographic inaccuracy.

All things considered, then, there is little doubt that traditional landline-use trends in the U.S. are changing dramatically, a development that has implications for survey representation. Ehlen and Ehlen (2007) attempt to better understand the recent cell phone-only population trend by creating a model, which assumes that what individuals do "today" is conditional on the strength of current incentives to change their behavior, their propensity to respond to current incentives, and the strength of their tendency to persist in established behavior (p. 721). They use the model to forecast demographic characteristics of the CPO population, concluding that the adoption of the wireless lifestyle is not a simple "fad," but a rather significant trend in the American population (also see Lavrakas et al., 2007).

On the matter of other survey modes beyond the telephone, Messer and Dillman (2011) provide research on how to survey the general population over the internet, using address-based sampling (ABS) and mail contact. The authors reported results on five potential methods to obtain online responses: (1) sending web and mail modes of response sequentially, (2) providing a prepaid $5 incentive, 3) offering an instruction card for responding over the web, (4) sending the follow-up request by Priority Mail, and (5) providing an additional $5 incentive with this follow-up request.[11] Messer and Dillman found that using an initial web request followed by a mail request with the

[11] Results are evaluated from the standpoint of response rates, demographic representativeness of respondents, and survey costs and data collection times for web and mail modes.

prepaid incentive increased the likelihood of obtaining Internet responses. Furthermore, the mail follow-up to the initial web request substantially improved sample representativeness and increases response rates. Notwithstanding, the authors also noted that the mail-only design consistently obtained higher response rates and a demographically similar sample compared to "web-plus-mail." This is a significant finding underscoring the advantages of using the mail mode, which will factor into the pilot study discussed later.

It is also important to note that there is other empirical research focusing on the data quality of internet-based methods compared to other forms of surveying (e.g., Baker et al. 2010). Heerwegh and Loosveldt's (2008) research hypothesized that web survey respondents would be more likely to produce lower quality data than other techniques, such as face-to-face interviews. Indeed, the data supported the authors' hypothesis: web survey respondents were shown to (1) produce higher "don't know" response rate, (2) differentiate less on rating scales, and (3) produce more item nonresponse than face-to-face survey respondents.

Likewise, Chang and Krosnick (2009) also compare sample representativeness and response quality between internet surveys and RDD interviewing. The authors conducted a national field experiment in which identical questions were administered simultaneously by RDD telephone interviewing and by the internet. The mode difference underscores what is a common sampling difference. The RDD sample is considered a "probability" sample with all possible numbers having a known and quantifiable chance of being selected. Such samples are often considered "scientific" in the sense that they permit the use of inferential statistics. In contrast, most internet samples tend to use nonprobability selection techniques where individuals decide whether to take the survey from an unknown population of those who could have participated. Such surveys are often considered inferior by survey methodologists.[12] Results showed that the probability samples were more demographically representative of the national population than the nonprobability samples. At the same time, the telephone data produced more measurement error and more social desirability response bias than did the internet data. The probability internet sample, on the other hand, exhibited great sample composition and self-report accuracy. Overall, Chang and Krosnick's (2009) work suggests that internet data collection from a probability sample yields more accurate results than telephone interviewing and internet data collection from nonprobability samples (also see Yeager et al. 2011).

[12] The American Association for Public Opinion Research (AAPOR) commissioned a task force in 2010 to examine the rise in self-selection surveys (i.e., those in which people opt-in themselves), many of which are on the Internet. The AAPOR Task Force (2010) provides a detailed discussion of the inferential issues related to non-probability panels, and specifically recommended that, "Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values" (p. 5). Indeed, AAPOR's recent definition of codes and outcome rates states that, "For non-probability samples, response rate calculations make little sense, given the broader inferential concerns. Further, for many of these surveys, the denominator is unknown, making the calculation of response rates impossible (cf. Callegaro and DiSogra, 2008)."

In addition to the three main types of survey modes—telephone, mail, and internet—researchers have explored a variety of other methods of survey data collection. There are mixed modes in which researchers employ one or more of the methods (e.g., Rao, Kaminska, and McCutcheon 2010; Porter and Whitcomb 2007; Sakshaug, Yan, and Tourangeau 2010). For instance, Medway and Fulton found in their meta-analysis of 19 experiments using web surveys vs. mail that mail surveys with a web option have lower response rates, typically on the order of 13 percentage points (2012, 740).[13] Other times researchers employ designs that shift the mode in reaction to other data (so called "paradata") that is collected during the course of the interview (e.g., Couper 1998; Couper and Lyberg, 2005; Bates et al., 2008; Lynn and Nicolaas, 2010; Olson, 2010).

Also, new technologies have prompted the use of instant polls of consumers at the point of their transaction or perhaps via their smartphones. Other variants, such as offering customers a telephone number to phone in their views are likely to generate data of dubious quality. For instance, some of the earliest polling efforts allowed respondents to opt-in (e.g., the million-plus members of the Literary Digest Poll of the 1930s) or employed quota sampling methods whereby interviews could choose who they wanted to interview subject to meeting a few overall goals or criteria. Both of these methods have been discredited and each leads to well publicized failures to predict the actual state of public opinion (e.g., Herbst 1995; Berinsky 2005). The bottom line is that alternative sampling modes are often accompanied by non-standardized techniques for survey participation. The resulting samples are non-probability samples; whether they are representative is hard to determine. All of this means that even a sample that is large and diverse may still yield flawed estimates of public opinion if respondents are allowed to opt-in or out without detailed knowledge of this process.

Finally, surveys utilizing any mode can be conducted in languages other than English. The fact that many minorities in the U.S. are not English speakers presents yet a significant challenge to survey researchers. Brick et al. (2012) attempt to address this issue in their article evaluating methods to improve response rates from Spanish speakers in two-phase mail surveys. Indeed, the authors point out that a considerable amount of surveys are subject to non-response bias and underrepresent close to 20% of families that do not speak English in their households. In the U.S., Spanish speakers comprise 60% of those speaking some language other than English at home. Brick et al. (2012) devised a set of language treatment experiments to improve response rates from Spanish speakers, which were then implemented in the 2011 National Household Education Surveys Program (NHES) Field Test. These language experiments included treatments that varied the languages used in the materials (i.e. both English and Spanish, or English only), whether the materials in both languages were mailed in all

[13] As for why surveys with a web-option experience lower response rates, some possible explanations include the following: (1) increased complexity for respondents, (2) a disruption in the process of responding to the survey, and (3) online formats may have more implementation failures. Also, dual or mixed modes—such as the use of landlines and cell phones—can make it more challenging to combine the various types from a statistical point of view (Peytchev and Neely 2013).

mailings or just in the follow-up, the way the instruments presented the language alternatives, and whether the language efforts were targeted at certain addresses or sent to all households. Evidence from the NHES data showed that sending survey materials to all sampled households in both languages significantly increased participation from Spanish-speaking individuals, without jeopardizing responses from English-only households (see also Bouffard and Tancreto, 2006). Once again, this is a significant finding considering the size of the Hispanic and non-English speaking populations in Florida.

Sampling Frames & Strategies

Often, sampling strategies are synonymous with mode choice. For example, the "dialing" part of random digit dialing (RDD) defines it as a telephone survey. Likewise, once respondents are selected on the basis of their residential address—such as takes place in Address-based Sampling (ABS)—then the mail mode is the most practical unless efforts are undertaken to augment the data. However, some lists of voters will contain both address and telephone information (and potentially email addresses too), which means that no single mode must be used.[14]

Telephone surveys are the most dependent upon procedures to select the appropriate respondent (Kish 1965; also see Battaglia et al. 2008). Screening question formats affect response rates. In particular, it is best to refrain from revealing the eligibility criteria to potential respondents. For instance, Tourangeau, Kreuter and Eckman's (2012) article on motivated underreporting in screening interviews explores how underrepresentation can be attributed to interviewers, survey respondents and question content (also see Berinsky 2005 for strategic misreporting to other survey questions). The authors explain how screening questions can easily overlook eligible household members, a phenomenon that can be attributed to interviewer motivation, respondent motivation, or some combination of the two. Tourangeau et al. test several hypotheses about this underreporting by conducting an experiment as part of a telephone survey targeting 35 to 55 year-olds. They find that the format of the screening questions affected the proportion of households screening into the survey, as well as the response rates to both the screening and main interviews. In particular, revealing the eligibility criteria to respondents makes them less likely to report an eligible household member. Further, evidence from the experiment suggested that interviewers with higher screener completion rates tended to find fewer households with eligible members, indicating that they may have screened out eligible households to achieve high response rates to the screener. Thus, RDD may lead to more underrepresentation relative to other modes which do not employ the same screening methods.[15]

[14] However, some modes will be preferred. In their study of Leon County in 2011 with local voter registration lists, Jerit, Barabas, and Clifford (2013) found that only 7% of the registrants supplied an email address. The authors opted to use the mail mode to survey respondents given the much more complete sample coverage. The pilot study detailed later in this report discusses the results of largely unsuccessful attempts to recover email addresses from the voter list via commercial data suppliers.

[15] It is possible, however, some screening protocols could lead to false positives (i.e., people indicating they are eligible when they are not). This concern is related to geographic eligibility, which is discussed later.

A somewhat different, yet still concerning matter is researched by Merkle et al. (2009); in their article, the authors explore how RDD samples are commonly purged of telephone numbers listed as businesses in an attempt to increase interviewer productivity. However, removing these business numbers has an unintended consequence, namely, an increase in household non-coverage. Their data showed that these purged numbers were in fact mostly eligible households that were being left out of the sample. Overall, then, a substantial amount of evidence from the literature suggests that researchers must exercise caution when utilizing RDD surveying methods. Naturally, there is no consensus on the potential effects of nonresponse in survey research, particularly within RDD landline only surveys. Still, recent academic and government publications seem to be recognizing the potential bias of landline-only telephone surveys.

Scherpenzeel and Toepoel (2012) examine the area of sampling strategies by looking at the potential effects of incentives, contact mode, and information in survey research. They conduct an experiment to determine the best recruitment strategy for an online household panel. More specifically, the authors tested whether contact mode, incentive amount, timing of the incentive, content of the advance letter, and timing of the panel participation request had a significant effect on the quality of the population sample. One interesting detail about this study is the fact that it controlled for whether or not the households involved had a fixed telephone landline. Accordingly, Scherpenzeel and Toepoel (2012) contacted respondents on the phone, via computer assisted telephone interviewing (CATI), and face-to-face, via computer assisted personal interviewing (CAPI). Ultimately, the incentives tested by the authors were shown to have stronger effects on response rates when they were distributed in advance with screening letter than when they were paid later. Interestingly enough, however, the 20-euro and 50-euro incentives did not significantly increase response rates beyond those at the 10-euro level, which produced the highest response rates. Lastly, the results demonstrated equivalent contact and recruitment rates for CATI and CAPI for households with a known telephone number, with significantly lower rates for households without a known telephone number.[16]

In moving beyond RDD sampling, one of the most exciting developments concerns address-based sampling (ABS) and registration-based samples (RBS). In address-based sampling, survey respondents (typically households) are selected from lists of addresses maintained by the U.S. Postal Service (Brick et al. 2011). In the RBS sampling framework, researchers sample respondents from lists of registered voters compiled by the states (e.g., Fullerton, Dixon, and Borch 2007; Green and Gerber 2006). Since the implementation of motor voter registration laws and reforms following the disputed 2000 presidential election, the quality of these voter files has increased considerably (McDonald 2007).

[16] The Scherpenzeel and Toepoel study was conducted in the Netherlands, a country which still has about 70 percent of the population with a landline connection. However, the authors report that the proportion of households with landlines is “decreasing rapidly”. For additional insights into incentives, see Petrolia and Bhattacharjee 2009.

Schwartz and Richards (2003) explored underrepresentation in an applied political setting – the 2002 elections in New York and Pennsylvania. More specifically, they compared the surveying methods of Registration-based Sampling (RBS) and Random Digit Dialing (RDD), which they ultimately used to predict the outcome of the gubernatorial election in both states. The authors found that the final RBS polls in both states very closely approximated the actual election outcome (also see Green and Gerber 2006; Barber et al. 2011). A closer look at the demographics and results, however, uncovered similarities in some areas and differences in others. In fact, Schwartz and Richards (2003) noted that there was some underrepresentation of urban areas in RBS data as compared with RDD. Nevertheless, the significantly greater efficiency of RBS sampling, its savings in polling costs, and the substantial accuracy found in the study, all make a strong case for further consideration of RBS methods. As such, the 2013 FDOT pilot study utilizes RBS extensively.

Weighting and Estimation Procedures

There are two primary ways of contending with nonresponse problems in surveys.[17] One popular remedy is to apply corrections such as post-stratification weighting; with this procedure, attempts are made to "fix" the data after it has been collected (e.g., Brick et al. 2007; Massey and Tourangeau 2013b). More specifically, weighting takes place when responses from subgroups deemed to be underrepresented are adjusted so that they "count" more toward the aggregate results (see Lohr 2010). Weighting has been used in previous FDOT surveys (see the Appendix for an illustration of how to weight a survey). However, often some subgroups are so heavily underrepresented that the weights are unrealistically large if not difficult to calculate since the data are often recorded at highly aggregated-level (e.g., age categories).[18] Thus, the problem with weighting is that (A) it requires background/auxiliary data which may or may not be available, and (B) one can never ensure that all relevant sources of nonresponse have been corrected.

A second approach to deal with nonresponse entails the use of detailed information beyond what is collected in the survey. In particular, there are a number of techniques available to adjust responses statistically for patterns of nonresponse. One variant is the Heckman (1979) sample selection model (also see Greene 2012 or

[17] It is important to note that some researchers argue that pursuing reluctant respondents may not be worthwhile (e.g., Hox, de Leeuw, and Chang 2012) in the sense that there are no systematic differences in reliability or validity between those who respond early in a survey vs. those who respond later. The pilot study is designed to go a step further by studying participants versus non-participants rather than participants only.

[18] The customer survey from 2011 was weighted using four age categories (18-34, 35-54, 54-64, and 65+). After the sample was grouped into these age categories, the 2010 Census data revealed the county's population in each of these groups for the counties within each of the seven FDOT districts. After ratios were calculated of the sample to the population, a computer program was written in SPSS to compute and apply the weights. However, the 2011 survey had almost no representation of individuals in the 18-24 age range which required collapsing ages 18-34 into one category for weighting purposes. There was also some missing gender data, making it hard to weight on that dimension.

Fullerton, Dixon, and Borch 2007). With sample selection models, analysts first construct a model of who is likely to respond to the survey based upon background information that exists for everyone in the sampling frame. In the second stage, the responses of those who have participated in the survey are analyzed after they have been adjusted for the nonresponse patterns in the first stage. Another statistical technique is to use matching methods to select respondents within a sample who are descriptively representative (Guo and Fraser 2010). A related technique simply imputes the missing data from unit nonresponse (Peytchev 2012). Yet another strategy entails using statistical models to "control" for differences in sample composition, particularly with a technique known as multilevel or hierarchical modeling (Gelman and Hill 2007). With these models, researchers can "borrow strength" by using detailed information they have in some areas to improve the precision of estimates in different areas with less information. However, all of the aforementioned models require statistical expertise in the estimation and evaluation stage.

Pros and Cons of Survey Strategies

In comparing the major survey modes (telephone, mail, and internet), we do so from the perspective of FDOT and with an eye toward reducing the participation problems acknowledged earlier with past surveys. Also, it should be noted that there is tremendous variation in how any given survey organization conducts surveys (e.g., Nickerson 2007), so it could be that particular advantages or disadvantages are magnified. Likewise, costs can vary greatly depending upon the capabilities of the provider and the particular methods used. Having said all of this, we nonetheless endeavored to compare the main surveying techniques across nine evaluative criteria.[19]

[TABLE 1 ON NEXT PAGE]

[19] By mail, we are assuming the use of a pre-constructed list such as sampling from the voter registration list ("RBS" which is used later) or some other list of individuals/households. Alternatives include address-based sampling (ABS) or other lists available from vendors. Using the voter registration list is attractive because it permits a detailed account of nonresponse from the perspective of the individual, making it possible to determine who responds and who does not. In the future it could also be possible to target specific households through the geographic information systems to target specific areas or types of households (see Beimer and Peytchev 2012). With respect to the internet, we have in mind an email invitation which links to a survey delivered electronically via a service such as Qualtrics, Survey Monkey, or a related company. Sometimes, companies such as YouGov, Harris Interactive, or SSI offer both the pool of potential respondents and the technology to host the survey.

Table 1. Comparison of Survey Modes across Evaluative Criteria

	Landline Telephone	Cell/Wireless Telephone	Mail	Internet/Email
Response Rates	●●	●●	●●●	●
Representativeness	●●	●●	●●●	●
Coverage	●●●	●	●●	●
Financial Costs	●●	●	●●	●●●
Speed/Timeliness	●●	●	●	●●●
Data Entry Ease	●●	●●	●	●●●
Weighting/Correcting	●●	●●	●●●	●
Respondent Accuracy/Trust	●●	●●	●●●	●●
Respondent Safety	●●●	●	●●●	●●

● = lowest/worst
●● = medium/middle
●●● = highest/best

Response rates. The overwhelming conclusion from the literature is that telephone response rates are low and declining rapidly, even despite the introduction of cell phone numbers in the samples (e.g., Kohut et al. 2012; NRC 2013). However, internet surveys often fare worse. Accordingly, internet surveys are last due to their low response rates, something confirmed by others (e.g., Adey and Cornelius 2006, p. 110). It should be acknowledged that many times response rate calculations are impossible with internet surveys with unknown/opt-in sampling frames (e.g., Callegargo and Disorgra 2008). The experience with the pilot study reported later confirms the ranking of this mode low with respect to response rates. The telephone modes are given the middle scores, particularly because some of the groups most likely to not respond to a survey are those that FDOT seeks to reach. As Adey and Cornelius (2006) write, "There is evidence that young people, the elderly, the poor and poorly educated, the uninsured, and individuals who have certain disabilities (such as hearing loss) or tire easily may be less likely to participate in telephone surveys." (p.110-1). In contrast, mail surveys are rated more highly since they tend to produce comparatively higher response rates.[20]

[20] For example, Breenan and Charbonneau (2009) used a mail survey in New Zealand and obtained response rates of 66.5 percent after three follow-ups. However, the incentives used in this study (chocolate) might have something to do with the elevated response rates. Yet, even the control condition (without chocolate) had response rates of 62.3 percent with three follow-ups. In contrast, the telephone surveys in the report generated by Kohut et al. (2012) revealed 10 percent response rate for landlines and a 7 percent response rate for cell phones in 2012 (the combined rate was 9 percent). In the U.S. Montaquila et al. 2013 report higher responses rates with mail surveys, particularly when they have a screening step to determine eligibility and when the topic is made salient. Other studies have the two modes closer but still show an advantage for mail surveys, particularly when more than one mailing is used; in particular and with respect to response rates, "…the mail survey performing significantly better in five of the six states: Washington (+6.7 percent), Texas (+5.4 percent), California (+4.5 percent), Illinois (+4.1 percent), and New Jersey (+3.7 percent). In North Carolina, the second mailing markedly improved the response rates for the mail survey, yet the rate was still significantly lower than that obtained by the telephone survey" (Link et al. 2008, pages 17-18).

Representativeness. One advantage of mail surveys is that all respondents, irrespective of their technological background, can complete the survey. Perhaps for this reason, people from all walks of life are likely to respond, especially if the burdens are not too great (e.g., the survey is not too long, on a topic of interest). It is also possible, by tracking responses, to determine who returns a survey and who does not. Random digit dial surveys and cell phone surveys, especially those at higher geographic units (i.e., states rather than counties) can be representative or adjusted to become more so.

Coverage. Many people do not have email addresses or refuse to provide them to researchers. Likewise, some mail surveys start with populations like the voter registration lists that are subsets of the entire population (i.e., the registered population might differ from the unregistered population).[21] Landline telephone surveys bypass many of these concerns by randomly selecting respondents from telephone exchanges, providing a high degree of (potential) coverage since those who have email or not and who are registered or not would be included as long as they have a working landline phone number.[22] However, landlines are becoming portable as well, so the geographic inaccuracy could be a problem here too. As noted earlier, though, wireless phone lines are often highly geographically inaccurate, especially at the county-level, making it hard to determine the geographic origin of the respondent (Skalland and Khare 2013). Similar problems emerge with internet surveys since most are drawn from an unknown sampling frame. But it is possible, with added financial resources, to recruit probability samples for online surveys (e.g., Scherpenzeel and Toepoel 2012).

Financial Costs. Even though the price of making a phone call is dropping, telephone methods are expensive because of the staffing and software needs inherent in this method. In other words, to reach any given respondent often entails multiple calls and the staffing resources to accomplish this are not insignificant. Also, as response rates drop on the telephone mode, it requires more calling attempts, which costs more relative to earlier years. It is also best to conduct phone polls via a computer-assisted telephone interview system (CATI), which are expensive to purchase and maintain. Cell phone samples might help improve the representativeness of the survey, but they also increase the costs of a telephone survey (Link et al. 2007). In comparison, mail surveys are costly because of the postage and due to the staffing needed to distribute and track the mailings, but most of the costs come upfront and postal costs are only incurred for surveys that are returned. Finally, the internet offers a cheap way to deliver surveys, but there are costs, among others, for the survey software license and for the costs of designing the instrument online. However, once the surveys have been deployed, responses come back in a highly usable format.

Of course, the financial costs depend upon the survey organization, particularly their capabilities and expertise. For instance, the authors of this report solicited a bid for the

[21] The pilot study attempts to move beyond this concern by including an attempt to study registered and non-registered populations. For more on the extent to which mail delivery sequence files provide survey coverage, see Iannacchione, Staab, and Redden 2003.

[22] Screening for telephone service can help improve coverage in a phone survey (Kennedy 2007; also see Kennedy, Keeter, and Dimock 2008).

survey work on the pilot study. A private survey firm in Gainesville, Florida by the name of Independent Data Collection Center prepared cost estimates for mail, telephone, and email surveys of nine Florida counties (Brevard, Flagler, Lake, Marion, Orange, Osceola, Seminole, Sumter, and Volusia) on a 40-50 question survey (closed-ended) with 2-3 additional open-ended responses of roughly 15-20 minutes in length. The cost estimates below are for 500 completions for each mode.

Mail Survey:	$40,988.52	(Price per completion: $81.98)
Phone Survey:	$22,585.00	(Price per completion: $45.17)
Internet/Email Survey:	$ 9,995.00	(Price per completion: $19.99)

These estimates, obtained during the spring of 2013, were preferred pricing (given to someone with a survey research background) and assume no questionnaire development (i.e., that the questions have already been formulated). Also, they do not include extra services like coding the open-ended comments or preparing the survey report.[23] Lastly, additional survey waves for the mail would cost more,[24] and the cost estimates also do not include sales tax.

The Survey Research Laboratory at Florida State University was able to field a comparatively larger set of surveys for less than what is reported above. For the mail survey, the costs are likely closer to $24 per completion (i.e., based upon a cost of $13,046 for 588 completions in the primary mail sample). However, these cost figures tend toward being conservative. For instance, the FSU researchers there were able to leverage existing expertise and resources to offset costs (e.g., past experience with the FDOT survey, compiling voter registration lists, computing techniques, research assistance, the university site licenses for statistical software and internet surveying capabilities). As such, these outside cost estimates should be viewed as approximate and they would likely change in the future. Given the experience of the FSU team on the pilot project, the internet survey costs *appear* to be low given what it takes to procure email addresses for a high quality sample of actual voters and the costs of programming the survey and fielding it electronically. However, *in reality* the effort does not produce much in terms of completions (as will be discussed later), at least with the procedures employed.[25]

Speed/Timeliness. For quick turnout, the internet mode appears to enjoy an advantage. As stated above, surveys can be deployed quickly (once they are designed) and the data come back in a usable format very quickly. This means several time intensive

[23] On the telephone survey, coding 2 open-ended questions is $1.75 per survey per questionnaire, for a total cost of $1,750 for two questions on a 500 person survey. Preparing a survey report would be $4,500 extra and would include an executive summary, the key findings, numerical results for each question with written descriptions, representative verbatims, conclusions, and recommendations.

[24] For example, a second wave mailing from IDCC would cost an additional $40,178.55, for a total price including data processing of $93,056.07.

[25] It might be possible to contract with sampling firms such as SSI or YouGov/Polimetrix to obtain internet samples from their panels. The representativeness of these samples can be questioned since they are nonprobability samples typically, but the number of completions will likely be higher.

steps can be cut out related to data entry. However, if multiple attempts are made to contact individuals, it could add time to the process.

Data Entry Ease. Once respondents respond to an internet survey, the data have already been "entered" into a database. Data entry would also be automated with a CATI system in telephone surveys. With the mail modes, the extra step of entering data could result in errors as well as additional costs.[26] However, it might be possible automate the process of entering mail surveys (e.g., via scanning) so that this disadvantage is offset.

Weighting/Correcting. Technically, data from all three modes can be weighted. However, with a household address system like address-based sampling (ABS), the household level unit precludes weighting on characteristics like age. However, the mail mode using registration based sampling (RBS) starts with a definition of the sampling frame that is highly useful for both weighting and statistical corrections. In other words, surveyors tend to know who they are inviting and who responds from the invitation list. Deviations can be corrected with more certainty than would be the case with telephone or internet surveys with unknown samples.

Respondent Accuracy/Trust. In the mail format, respondents can see all the introductory materials as well as what they are being asked to provide prior to completing the survey. The ability to see the entire survey is likely of benefit for increasing trust. Of course, survey length could moderate this effect. For instance, if respondents "see" a lengthy survey going on for several pages, they might be disinclined to respond. In contrast, telephone surveys often disguise their length unless they are short and then this information can be made known at the outset to improve responses (e.g., Chang and Krosnick 2010; Presser and Singer 2007).[27]

Safety. The safety of the respondents from physical or economic harm is another valid consideration. On this dimension, mail surveys appear to be the safest. Internet surveys have the potential to carry computer viruses while telephone surveys conducted via wireless devices are potentially physically dangerous. With respect to bodily safety, according to the National Safety Council (NSC), "23 percent" of all crashes each year involve cell phone use, resulting in 1.3 million crashes nationally. Distractions, along with alcohol and speeding, are now leading factors in fatal and serious injury crashes. The NSC also notes that researchers observing more than 1,700 drivers found that 3 out of 4 drivers using a cell phone committed a traffic violation. Perhaps not surprisingly, then, talking on a cell phone while driving makes someone four times as likely to crash. While a growing number of drivers are turning to hands-free devices, studies show hands-free devices provide no safety benefit. The NSC believes that "It's the conversation, not the device, that creates the danger."[28] Furthermore, the NSC has

[26] With a CATI system, the data entry costs can be reduced significantly.
[27] There are also incentives that can be used to enhance participation in a mail survey (Brennan and Charbonneau 2009). For the effects of visual stimuli in surveys, see Christian, Dillman, and Smyth 2007;Couper, Conrad, and Tourangeau 2007; or Toepoel and Couper 2011.
[28] http://www.nsc.org/safety_road/Employer%20Traffic%20Safety/Pages/NationalDistractedDriving.aspx

produced a white paper detailing the potential liability employers face if employees have accidents while on the phone and driving.[29] The liability issues for state governments and agencies were not explicitly addressed, but the legal and ethical considerations should be reviewed in depth before proceeding.

As of 2013, cell phones are not banned while driving in Florida. If such a ban were to be instituted, contacting respondents might be comparatively safer. In other words, if cell phones were taken away from drivers in an effective manner, then placing survey calls over cell phones would be relatively safer (i.e., the calls would be going to non-drivers so the driving-while-being-interviewed dangers would be averted). Note, however, that the Florida Department of Motor Vehicles explicitly discourages talking while driving even though it is legal.[30] If surveying via cell or wireless device was to be attempted in Florida, it would be important to quickly determine whether the respondent was driving or about to drive during the interview. Then it would be possible to reschedule the call or terminate the call. Importantly, though, this safety conscious decision might decrease survey response if respondents are not successfully recontacted.

The criteria above are not exhaustive and sometimes improvements on one dimension results in reductions for another. Also, we consciously do not consider efforts within each mode to boost participation (e.g., multiple survey attempts, incentives, bilingual materials).[31] For example, even inserting a deadline for returning the survey might boost response rates (Martin 2009). Also, mixed modes are possible, as are adaptive designs that evolve over the course of the field period (National Research Council 2013).

2013 Pilot Surveys in FDOT District 5

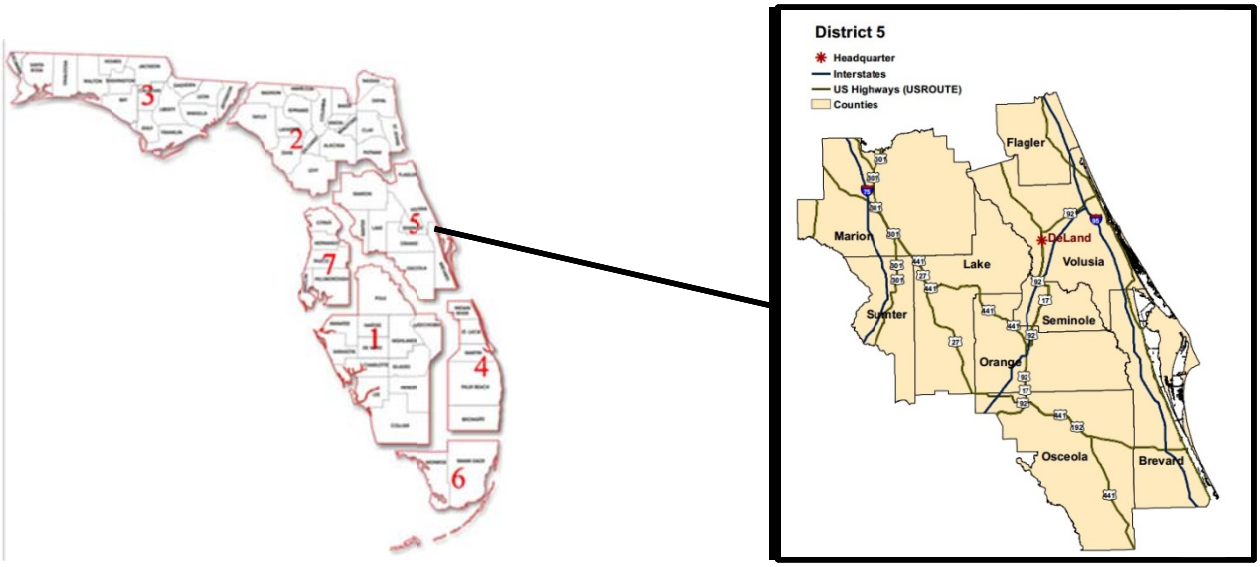
The primary goal of the pilot study was to investigate alternatives to the existing methods used by FDOT in their Customer Satisfaction Surveys (i.e., traditional landline telephone interviewing). While there were a lot of potential alternatives—many of which were discussed in the literature review—the pilot study employed a cost-conscious hybrid design using different sampling techniques and modes. In particular, the main data collection effort centered upon a mail survey, but there was also a more limited attempt to interview respondents via the internet/web using an email invitation.[32]

The pilot study was also limited geographically. Given the budget constraints and the desire to understand the representation issues in depth, the research team at Florida State University decided to focus the pilot study on a mail survey of FDOT District 5, which encompassed nine counties of Brevard, Flagler, Lake, Marion, Orange, Osceola, Seminole, Sumter, and Volusia. These counties have a mix of urban and rural

[29] http://www.nsc.org/safety_road/Distracted_Driving/Documents/CorpLiability_wp.pdf
[30] http://www.dmvflorida.org/cell-phone-distractions.shtml
[31] There is also evidence suggesting that shorter surveys generate higher response rates (Galesic and Bosnjak 2009).
[32] The mail and internet surveys were timed to be delivered during the week of April 8, 2013 and the field period for all modes ended on June 28, 2013, but most responses were received weeks earlier.

populations as well as sizeable minority populations. District 5 is illustrated on the map below in Figure 1.

Figure 1. Florida Department of Transportation, District 5



The main part of the pilot study focused on surveying residents from voter registration lists obtained by the Florida Division of Elections (FDOE). These lists are maintained by the county supervisors of elections and updated periodically. After applying some initial screening criteria to remove ineligible cases,[33] the sample population was over 2.4 million people distributed into the nine counties as shown on the left side of the figure below. In a procedure known as Registration Based Sampling (RBS), the random sample drawn from the population is shown on the right of the tables below. Each county has a number of potential respondents proportional to its size. In that sense, the sampling procedures employed here produce highly representative sampling frames in a way that might be different had telephone exchanges been used (i.e., if phone prefixes straddle county lines).

[33] The list was purged from 2,550,132 to 2,440,388 by removing respondents who were under age 18 as of March 1, 2013 (i.e., 17 year olds who are allowed to register early were removed), individuals in the law enforcement or legal community who have "suppressed" data files omitting address information, and individuals who report living out of state or the country even though they are still registered to vote in Florida.

Table 2. County Representation for FDOT District 5

	2013 Florida Div. of Elections		Random Sample	
County	**N**	**%**	**N**	**%**
Brevard	370,156	15.2%	1,532	15.3%
Flagler	72,258	3.0%	307	3.1%
Lake	213,615	8.8%	870	8.7%
Marion	229,562	9.4%	907	9.1%
Orange	689,786	28.3%	2,836	28.4%
Osceola	173,577	7.1%	723	7.2%
Seminole	281,710	11.5%	1,140	11.4%
Sumter	74,108	3.0%	291	2.9%
Volusia	335,616	13.8%	1,394	13.9%
Total	2,440,388	100.0%	10,000	100.0%

The sample was highly diverse in that more than 50% of respondents were female, 26% identified with a minority group (including a more than 50% of the respondents in Osceola County). Roughly 18% of the registered voter sample from the Division of Elections was under age 30, but this figure rose to almost a quarter in some counties.

The results reported in the next subsection go beyond the core FDOE sample in a two main ways. First, faculty members at FSU marshaled non-FDOT research funds in an attempt to supplement the original study to see what the unregistered population was in this area and whether their responses would differ.[34] To obtain a sample of the unregistered population, FSU staff worked with a private firm, Catalist, to obtain a random sample of non-registered voters in the nine FDOT district 5 counties as well as a companion sample of registered voters from the same company.[35] This duplication of the registered voter sample (i.e., two samples ostensibly of the same registered population for the nine counties) was undertaken to determine if the Florida Division of Elections lists were similar to what could be obtained from the private firm.[36] In general, the samples from the private data collection firm were similarly diverse. Due to resource constraints, however, there were far more mail surveys attempted of the FL Division of Elections sample than for the Catalist sample (i.e., 5,000 vs. 2,000).[37] The registered and unregistered populations are highly similar (10.2% vs. 11.5% black for registered and unregistered on average; 12.3% vs. 12.2% Hispanic on average for registered and unregistered), but there are important exceptions. For example, in

[34] Prior to contracting, Catalist estimated that on average, 29.5% of the individuals living in the 9 counties comprising FDOT District 5 were unregistered. Numerically this is more than 1 million people, almost all of whom are 18 and older. Thus, the rationale was to try to see if the results and response rates would differ for the registered and unregistered populations.

[35] Catalist maintains a database of "clean and current" data for more than 265 million individuals in the U.S. The appendix contains additional information on their firm and capabilities.

[36] For instance, the private firm promised more detailed and regular updating of people who had moved, died, etc. However, the quality of "big data" is not well documented.

[37] Since the Catalist samples are "exploratory," the discussion of responses to the survey will focus on the highest quality sample obtained from the Florida Division of Elections. Also, the pricing given to academic researchers is less than it would be for non-academic researchers.

Osceola County, roughly 42% of the Hispanic population is registered compared with 36.5% of the unregistered (as compared with an estimate of 47.8% Hispanic from the Census Bureau in 2012).[38] This suggests that in some cases the voter registration population is more reflective of the underlying diversity.

The second way in which the pilot study was augmented involved an attempt to recover internet addresses for the FDOE registered voter list as well as the names supplied by Catalist.[39] FSU personnel investigated several commercial email look-up services (e.g., Spokeo, Emailfinder.com), but these were rather limited in that they required looking up names individually instead of submitting names in bulk, and even premium subscriptions were capped at levels far below the sample sizes needed for the pilot study. Consequently, arrangements were made with a contractor, AcquireWeb, to access their data files in an attempt to match names in the sample with known emails.[40] This process involved several steps to confirm with federal anti-spamming regulations (i.e., individual were given a chance to opt-out prior to receiving surveys).[41] Ultimately, and as shown in Table 3 below, this process yielded internet addresses for about 4,500 individuals out of more than 62,999 originally (7.2%), 421 of the 5,000 individuals in the original the FDOE sample (8.4%), 2,478 of the 29,000 private firm registered voters (8.5%), and 1,618 of the 28,999 unregistered individuals (5.6%).[42]

Table 3. Email Availability by Sample Type

	Internet Sample							
	All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
	N	**% of Above**	**N**	**% of Above**	**N**	**% of Above**	**N**	**% of Above**
Original	62,999	--	5,000	--	29,000	--	28,999	--
Has Email	4,517	7.2%	421	8.4%	2,478	8.5%	1,618	5.6%
Responded	103	2.3%	15	3.6%	64	2.6%	24	1.5%

As the table shows, even though internet addresses were obtained for thousands of individuals (i.e., 4,517 “has email”), only 2.3% (n=103) of those people responded to

[38] http://quickfacts.census.gov/qfd/states/12/12097.html

[39] Again, this was technically beyond the scope of the pilot study and financed by faculty member research funds but the results are illustrative and included since they speak to the issues of representation across the various modes and samples.

[40] For more on the consumer identity integration industry, see http://www.acquireweb.com/. This firm was one of several recommended by the firm supplying the names of voters and non-voters (i.e., Catalist). In personal correspondence with Rich Paterra of Acquire, typically they are able to match email addresses 8 to 15 percent of individuals and 12 to 25 percent of households.

[41] There were only 20 “opt-outs” out of thousands of emails sent during the initial screening.

[42] Even though the recovery operation produced emails for fewer than 10% of the sample, bouncebacks (i.e., undeliverable) email messages, were low (approximately 163). Some email addresses were particularly prone to bouncebacks (e.g., Hotmail) due to the capabilities of the Internet firm and their spam filtering technologies. Most of the email completions were obtained quickly. A second round of email surveys did not significantly increase responses.

a survey across the various types of samples.[43] In some cases, as with the FDOE sample (n=15), the actual number of respondents was so low that responses can be highly unstable (i.e., most statistics and survey textbooks recommend samples of at least two to ten times larger). Consequently, the next section will provide some descriptive statistics on Internet participants but the section on the substantive responses for the survey questions will focus on the mail sample respondents.

Survey Participation and Non-Participation

For the mail survey taken together (i.e., from all three types of samples), the response rate was 10.4% using AAPOR Response Rate 1 definition (i.e., 728 completions across all types of mail survey respondents out of 7,000 surveys sent). This number likely would have been better by about 5-10% percentage points if a second mailing had been employed based upon past studies conducted by the FSU Survey Research Lab. The costs of secondary mailings are typically the same or lower than the original given that fewer mailings need to be sent, however, the cost to mail survey materials is a fixed cost (see, for example, footnote 24).[44] Yet, even with one mailing, the response rate was better than the 2011-12 Resident and Commercial Driver customer survey which achieved a response rate of 6.8% by mail. Yet, the response rate in 2013 was less than that achieved in residential surveys of 2011 and 2009. For example, in 2011, there was a 19.7% response rate statewide for the landline telephone survey.[45] The 2011 number itself is a decline from the previous survey in 2009 which achieved a 24% response rate.[46] In that sense, it is worth noting that the trend in the two telephone surveys was downward in the response rate with more non-contacts and ineligible numbers. These patterns suggest telephone surveying is becoming more difficult.

The overall surveys had slightly more women participate than men (52% vs. 48%, n=728), but the percentages were higher in female composition for the private firm sample (57% female vs. 43% male, n=243) than they were for the FDOE sample (51% vs. 49%, n=588). However, the 2011 BEBR study was much more skewed with respect to gender, with 62% of the respondents being female. In that sense, all of the 2013 pilot study samples were more representative with respect to gender (i.e. they were closer to the 53% female in the FDOE voter database for the nine counties of FDOT district 5 (n=2,440,388).

[43] Another 30+ individuals opened the survey but did not complete any questions; we discard these observations, effectively treating them as if they did not respond to the survey.
[44] The survey sponsor needs to decide the cost/benefit of doing a second mailing. It should enhance the response rates, but does add cost to the survey effort. There are also options short of re-sending the questionnaires, such as "prompting" responses with a relatively inexpensive postcard reminder.
[45] This calculation makes use of AAPOR response rate 1 with 2,715 completions out of 27,775 numbers attempted, of which 13,981 were not working numbers and 7,657 were no contacts (e.g., answering machine, no answer, busy), and ,422 individuals who refused to answer the survey.
[46] There were 2,807 completions out of 22,616 numbers attempted with 3,172 refusals and 5,950 non-contacts, and 10,687 not eligible.

Table 4 below shows the responses by race and ethnic categories. The top half of the table shows the racial and ethnic composition of the nine counties in FDOT District 5 for all 2.4 million individuals in the voter file as well as the various samples utilized in the pilot study. As expected, the samples appear to be highly representative, although the Catalist registered voters had more non-Hispanic Whites while the non-voter sample had higher percentages of blacks and Hispanics. Again, though, these figures are for the overall samples, not those who actually responded to the surveys.

The bottom half of Table 4 reports the racial and ethnic composition of survey respondents in the 2011 landline telephone survey from BEBR as well as for each of the samples in the 2013 mail and internet surveys. In terms of race and ethnicity, most respondents in the mail mode were self-identified whites (78.8% overall), followed by Hispanics (6.9%) and non-Hispanic blacks (5.2%). The registered samples, both FDOE and Catalist, are not as diverse as the non-voter sample with respect to Hispanic ethnicity as compared to the initial composition or what was achieved in 2011 via telephone interviewing. However, the non-registered voter sample, even though rather small, did have more Hispanic representation in the mail condition than what was observed in 2011. The last set of entries for Table 4 shows the racial and ethnic composition of the individuals, 103 in all, who responded to the internet survey. Non-Hispanic whites make up the vast majority of those who responded via the internet. There was some minimal presence of Hispanics and black respondents, but both racial/ethnic categories together constituted just over 10% of the sample.

[TABLE 4 APPEARS ON THE FOLLOWING PAGE]

Table 4. Racial and Ethnic Categories by Sample Type for FDOT District 5, Spring 2013

			Overall Sample, 2013							
	Entire District FDOE Voter Data		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
White, Non-Hispanic	1,671,818	68.5%	48,197	68.9%	6,840	68.4%	21,493	71.6%	19,864	66.2%
Hispanic	327,894	13.4%	9,831	14.0%	1,371	13.7%	4,023	13.4%	4,437	14.8%
Black, Non-Hispanic	277,582	11.4%	9,137	13.1%	1,146	11.5%	3,600	12.0%	4,391	14.6%
Asian, Indian, Pacific	45,945	1.9%	2,055	2.9%	169	1.7%	757	2.5%	1,129	3.8%
American Indian	8,370	0.3%	201	0.3%	32	0.3%	114	0.4%	55	0.2%
Multiracial	7,424	0.3%	23	0.0%	23	0.2%	--	--		
Other/Unknown	101,355	4.2%	555	0.8%	419	4.2%	13	0.0%	123	0.4%
Total	2,440,388	100.0%	69,999	100.0%	10,000	100.0%	30,000	100.0%	29,999	100.0%

			Mail Survey Respondents, 2013							
	BEBR 2011 Landline, Dist. 5		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
White, Non-Hispanic	312	75.0%	574	78.8%	453	79.1%	95	83.3%	26	63.4%
Hispanic	36	8.7%	50	6.9%	42	7.3%	4	3.5%	4	9.8%
Black, Non-Hispanic	30	7.2%	38	5.2%	31	5.4%	6	5.3%	1	2.4%
Asian, Indian, Pacific	7	1.7%	15	2.1%	12	2.1%	1	0.9%	2	4.9%
American Indian	1	0.2%	7	1.0%	4	0.7%	2	1.8%	1	2.4%
Multiracial	7	1.7%	4	0.5%	3	0.5%	1	0.9%	0	0.0%
Other/Unknown	23	5.5%	40	5.5%	28	4.9%	5	4.4%	7	17.1%
Total	416	100.0%	728	100.0%	573	100.0%	114	100.0%	41	100.0%

			Internet Survey Respondents, 2013							
	BEBR 2011 Landline, Dist. 5		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
White, Non-Hispanic	312	75.0%	91	88.3%	14	93.3%	57	89.1%	20	83.3%
Hispanic	36	8.7%	7	6.8%	0	0.0%	4	6.3%	3	12.5%
Black, Non-Hispanic	30	7.2%	5	4.9%	1	6.7%	3	4.7%	1	4.2%
Asian, Indian, Pacific	7	1.7%	0	0.0%	0	0.0%	0	0.0%	0	0.0%
American Indian	1	0.2%	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Multiracial	7	1.7%	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Other/Unknown	23	5.5%	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Total	416	100.0%	103	100.0%	15	100.0%	64	100.0%	24	100.0%

The next table, Table 5, shows the samples and responses by age category. For reference purposes, data from the Census and the most recent 2011 FDOT study by BEBR are provided as well. Overall, the samples are well-balanced with respect to age category; i.e., they look like the Census numbers. Among those who responded, though, some familiar descriptive biases emerge. There are fewer young people and more respondents age 65 or older. However, the 2013 pilot study produce more response among the youngest cohort of 18 to 34 year olds as compared to the 2011 District 5 survey (i.e., 11.6% via mail in 2013 vs. 7.0% via landline telephone in 2011). Among the non-voter sample, both mail and internet, the percentage of respondents in the youngest cohort was over 20%.

[TABLE 5 APPEARS ON THE FOLLOWING PAGE]

Table 5. Age Categories by Sample Type for FDOT District 5, Spring 2013

			Overall Sample, 2013							
Age	2010 Census		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
18-34	805,432	27.8%	15,208	26.3%	2,572	25.7%	8,607	28.7%	4,029	22.7%
35-54	1,009,394	34.8%	19,276	33.4%	3,226	32.3%	9,348	31.2%	6,702	37.7%
55-64	458,464	15.8%	9,066	15.7%	1,694	16.9%	4,641	15.5%	2,731	15.4%
65+	627,331	21.6%	14,218	24.6%	2,508	25.1%	7,404	24.7%	4,306	24.2%
Total	2,900,621	100.0%	57,768	100.0%	10,000	100.0%	30,000	100.0%	17,768	100.0%

			Mail Survey Respondents, 2013							
Age	BEBR 2011 Landline, Dist. 5		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
18-34	29	7.0%	84	11.6%	64	11.2%	12	10.5%	8	21.6%
35-54	125	30.2%	163	22.5%	131	22.9%	22	19.3%	10	27.0%
55-64	92	22.2%	169	23.3%	135	23.6%	29	25.4%	5	13.5%
65+	168	40.6%	308	42.5%	243	42.4%	51	44.7%	14	37.8%
Total	414	100.0%	724	100.0%	573	100.0%	114	100.0%	37	100.0%

			Internet Survey Respondents, 2013							
Age	BEBR 2011 Landline, Dist. 5		All		Florida Div. of Elections		Catalist Reg. Voters		Catalist Non-Voters	
Category	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
18-34	29	7.0%	12	11.7%	2	13.3%	4	6.3%	6	25.0%
35-54	125	30.2%	37	35.9%	5	33.3%	24	37.5%	8	33.3%
55-64	92	22.2%	16	15.5%	3	20.0%	9	14.1%	4	16.7%
65+	168	40.6%	38	36.9%	5	33.3%	27	42.2%	6	25.0%
Total	414	100.0%	103	100.0%	15	100.0%	64	100.0%	24	100.0%

Note: Some survey respondents did not provide their birth year. In these instances, background age information from the original sample files (i.e., from FL DoE or Catalist) was used to fill-in the missing age responses.

Table 6 reports survey response rates for specific demographic subgroups of interest. The table contains columns for the various types of mail survey samples as appeared in Tables 4 and 5. The rows are also similar, although some racial categories have been collapsed for ease of presentation. Overall, Hispanics, blacks, and respondents age 18 to 34 are the least likely to respond. For example, the 0.5% entry for Hispanics in the “All” samples in the upper left corner of Table 6 uses the 50 Hispanic survey participants across all three types of samples divided by 9,831 who were invited to participate (i.e., 50/9,831=.005). The other values in Table 6 are constructed in a similar fashion, dividing the number of respondents in each racial/ethnic or age category by the total number sent mail surveys. Viewed in this manner, it is apparent that the sample of registered voters from the Florida Division of Elections had comparatively higher rates of participation than did the samples from Catalist, the private vendor. Nevertheless, it was still the case that Hispanics, blacks, and younger respondents aged 18 to 34 participated half as often (or less frequently) than respondents in other racial/ethnic and age categories.

Table 6. Response Rates for Mail Survey by Race/Ethnicity and Age Category

	Response Rates			
	All	Florida Div. of Elections	Catalist Reg. Voters	Catalist Non-Voters
Race/Ethnicity				
Hispanic	0.5%	3.1%	0.1%	0.1%
Black, Non-Hispanic	0.4%	2.7%	0.2%	0.0%
White, Non-Hispanic	1.2%	6.6%	0.4%	0.1%
All Others	2.3%	7.3%	1.0%	0.8%
Age Category				
18-34 years old	0.6%	2.5%	0.1%	0.2%
35-54 years old	0.8%	4.1%	0.2%	0.1%
55-64 years old	1.9%	8.0%	0.6%	0.2%
65+ years old	2.2%	9.7%	0.7%	0.3%

Collectively, the 2013 pilot study was able to produce modestly higher degrees of descriptive representation than in the past, especially with respect to gender and age. The racial and ethnic composition still lagged the population values but some types of samples, especially the non-registered voters, were closer to the benchmarks. Of course, many of the individuals in the low participation subgroups are both minorities and younger. Thus, it is important to statistically control for these factors simultaneously. In a multivariate statistical model predicting participation in the mail survey, blacks and Hispanics are each roughly 5 percentage points less likely to participate than all other racial groups ($p < .01$, two-tailed test). Likewise, respondents age 18 to 34 are statistically less likely to respond to the survey as compared with those age 35 to 54 (+3 percentage points, $p < .05$), those age 55 to 65 (+12 percentage points, $p < .01$), and individuals 65 and older (+14 percentage points, $p < .01$). Thus, underrepresentation still exists, but with additional survey waves or oversampling of key subgroups then there is reason to suspect that the numbers could be made to be more representative.

Responses to FDOT Transportation Questions

The responses to the questions appear in attachment 2 (the FDOT version of the survey is in attachment 1). The 2013 survey results have been divided into three parts (S1=Registered voters from FDOE sample, S2=Registered Voters from private firm Catalist, S3=non-registered voters from private firm Catalist). The last set of entries titled, “2013ALL,” reports the results for respondents across all forms of the mail survey.

In general, the responses in the 2013 pilot study were highly similar to what has been achieved in the past for FDOT District 5. In other words, the collective 2013 results are typically within a few percentage points of the past results for 2011 and the prior years. There are deviations from this within the 2013 pilot study. For example, the first item on the spacing of exits and crossroad signs had a much higher percentage strongly agree (31% for S2 in 2013) as compared to the 2013 overall percentages (20%) and what had been achieved in prior years (23% in 2011, 17% in 2009, etc.).

Some of the strongest exceptions to the overall pattern of continuity are for travel through an FDOT construction zone in the last 12 months in Part II. Nearly all respondents irrespective of the sample type reported travel through a construction zone in 2013 (between 95 to 98%) as compared with estimates in the 80-90 percentage range in previous years. Of course, it could be that more construction is occurring in FDOT District 5 relative to the previous years, so it would be important to determine if the context changed before concluding that the pilot study is reaching a different subset of people. For other items, like travel time, littering, signal timing, and others, the percentages are remarkably similar to what was achieved in the past.

One area of departure with previous surveys was in the questions devoted to the need for bike lanes along roads or separated from roads. In general, there were fewer respondents who preferred designated bike lanes on roadways in 2013 (65% agree or strongly agree in 2013 across all samples) versus what the survey showed in 2011 (89% in the same two categories). Yet, the 2013 survey produced roughly the same support for bike paths separated from roads (90% in the strongly agree or agree category) as compared with the prior survey in 2011; however, there were many more responses in the “strongly agree” category in 2013 (54% overall) as compared with 2011 (44%), suggesting that the intensity of support was different and higher in 2013 for separated bike paths. Similar patterns we observed for the 2013 pilot study in terms of sidewalks along side of roads or separated from them.

Several factors were likely to have contributed to these differences. First, the 2013 pilot study did not include a filter question as previous years did asking a yes/no question on whether or not more bike lanes are needed in general. This change was due to an effort to streamline the survey in order to accommodate other questions that will be discussed next. Second, some versions of the 2013 survey provided background information and statistics on bike safety which lead to statistically discernible increases in support for bike lanes and paths as well as more support for sidewalks adjacent to and separate from roadways. Once these differences are taken into account, the distribution of responses looks similar to the historical patterns.

There were other points of departure from previous years. In particular, the 2013 respondents reported awareness of public transportation services near place of work a higher rates (Part IV, item B, 70% in 2013 vs. 64% in 2011), but most of the increase was among respondents from the private firm sample rather than the FDOE. Likewise, there was widespread variation in use of public transportation across the various types of samples in 2013, with the non-registered respondents being particularly likely to report use of public transportation. However, the sample sizes are very small (only 3 respondents who pasted the filter question preceding it). Another question that generated different responses was for the alternative travel modes, which much higher percentages reporting walking and carpooling than in the past (67% in 2013 for walking vs. 21% in 2011; 23% for carpooling in 2013 vs. 11% in 2011). Some of this could have been due to the layout of the mail survey as compared to the telephone mode. For other questions on overall satisfaction with aspects of the state highway system (Part V of the survey), the responses appeared to be similar to the past results.

Survey Sponsorship Experiment

Recent research studies suggest that surveys with a governmental sponsor tend to have higher rates of response. In particular, Medway and Fulton (2012) found that having a government sponsor listed on a survey can increase participation. In their review of more than a dozen published studies offering respondents a web-option, the surveys that mentioned a governmental sponsor experienced higher response rates than those that did not, typically on the order of a few percentage points ($p < .10$). In separate article, Groves et al. (2012) conducted an experiment with survey sponsorship attributed to either the March of Dimes charity or the University of Michigan.[47] Here the variations were across a charitable organization versus a state university (i.e. a quasi-governmental entity). Overall, 19.8% of the 6,000 mail surveys were returned, but the aggregate figure masks important underlying variability. When the March of Dimes was listed as the sponsoring organization, only 12.4% of 2,025 surveys were returned. In contrast, when the University of Michigan was listed as a sponsor, approximately 23% to 24% out 3,975 mail surveys were returned completed irrespective of whether the survey was cast as a "volunteering survey" or a "labor force survey." Groves et al. (2012) go on to note that these differences in survey response can be correlated with underlying support for the organization (i.e., those who had given money to or volunteered for the March of Dimes in the past), suggesting that who responds may be related to their underlying opinions.[48]

Given these findings in the survey research literature, the pilot study attempted to vary survey sponsorship in an effort to uncover differences in participation rates. However, due to Human Subjects Committee guidelines at Florida State University not all sponsorship variations were possible. In particular, all surveys being conducted by Florida State University researchers must disclose their origin (i.e., the surveys are being *fielded* by FSU Survey Research Laboratory), but it was possible to randomly alter the survey sponsor in this instance given the various funding sources. Consequently, there were two versions of the survey, one with the Florida Department of Transportation logo atop and another with Florida State University's seal at the top.

---

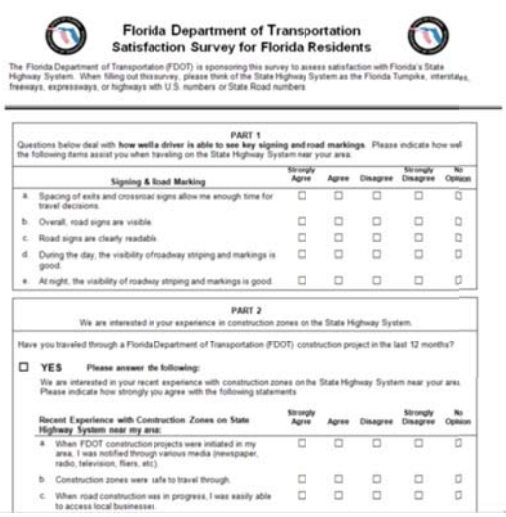
[47] The March of Dimes is a nonprofit organization devoted to preventing birth defects, premature birth, and infant mortality. Groves et al. (2012, pgs. 514-5) explain their protocols as follows: "One questionnaire used the March of Dimes logo and was titled 'The March of Dimes Survey of Giving and Volunteering'; a second questionnaire included the same questions but used the University of Michigan logo and was titled 'The University of Michigan Survey of Giving and Volunteering'; a final questionnaire included many of the same questions as the other two versions but used the University of Michigan logo and was titled 'The University of Michigan Survey of Jobs and the Labor Force.' The mailing protocol for each survey consisted of an advance letter, a paper questionnaire package (containing a cover letter, a twelve-page questionnaire, and a postpaid return envelope), a reminder postcard, and one follow-up questionnaire package for nonrespondents."

[48] Respondents who had historically supported the March of Dimes in the past based upon behavioral measures were more likely to respond to the survey irrespective of survey sponsorship. However, Groves et al. (2012, pgs. 517-8) write, "…the magnitude of this nonresponse bias is larger for the March of Dimes condition than for the University of Michigan conditions. In other words, although those committed to the March of Dimes are overrepresented in all [sponsorship forms of] the surveys, the overrepresentation is particularly acute with the March of Dimes sponsorship. Thus, the university sponsor brings in not only more of the sample but also a more representative set of the initial sample on the support variables."

The surveys were virtually identical otherwise.[49] Figure 2 shows each version of the survey with the differential sponsorship and logo information.

Figure 2. Questionnaires with Different Survey Sponsorship

FDOT Sponsorship



Florida Department of Transportation
Satisfaction Survey for Florida Residents

The Florida Department of Transportation (FDOT) is sponsoring this survey to assess satisfaction with Florida's State Highway System. When filling out this survey, please think of the State Highway System as the Florida Turnpike, interstates, freeways, expressways, or highways with U.S. numbers or State Road numbers.

PART 1

Questions below deal with **how well a driver is able to see key signing and road markings**. Please indicate how well the following items assist you when traveling on the State Highway System near your area.

Signing & Road Marking	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. Spacing of exits and crossroad signs allow me enough time for travel decisions.	☐	☐	☐	☐	☐
b. Overall, road signs are visible.	☐	☐	☐	☐	☐
c. Road signs are clearly readable.	☐	☐	☐	☐	☐
d. During the day, the visibility of roadway striping and markings is good.	☐	☐	☐	☐	☐
e. At night, the visibility of roadway striping and markings is good.	☐	☐	☐	☐	☐

PART 2

We are interested in your experience in construction zones on the State Highway System.

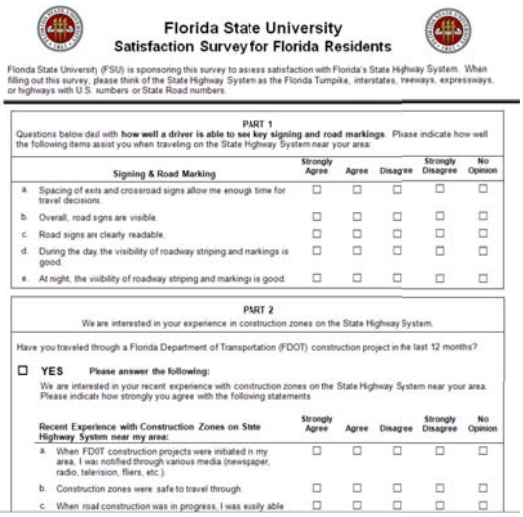
Have you traveled through a Florida Department of Transportation (FDOT) construction project in the last 12 months?

☐ YES **Please answer the following:**

We are interested in your recent experience with construction zones on the State Highway System near your area. Please indicate how strongly you agree with the following statements

Recent Experience with Construction Zones on State Highway System near my area:	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. When FDOT construction projects were initiated in my area, I was notified through various media (newspaper, radio, television, fliers, etc.)	☐	☐	☐	☐	☐
b. Construction zones were safe to travel through.	☐	☐	☐	☐	☐
c. When road construction was in progress, I was easily able to access local businesses	☐	☐	☐	☐	☐

FSU Sponsorship



Florida State University
Satisfaction Survey for Florida Residents

Florida State University (FSU) is sponsoring this survey to assess satisfaction with Florida's State Highway System. When filling out this survey, please think of the State Highway System as the Florida Turnpike, interstates, freeways, expressways, or highways with U.S. numbers or State Road numbers.

PART 1

Questions below deal with **how well a driver is able to see key signing and road markings**. Please indicate how well the following items assist you when traveling on the State Highway System near your area.

Signing & Road Marking	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. Spacing of exits and crossroad signs allow me enough time for travel decisions.	☐	☐	☐	☐	☐
b. Overall, road signs are visible.	☐	☐	☐	☐	☐
c. Road signs are clearly readable.	☐	☐	☐	☐	☐
d. During the day, the visibility of roadway striping and markings is good.	☐	☐	☐	☐	☐
e. At night, the visibility of roadway striping and markings is good.	☐	☐	☐	☐	☐

PART 2

We are interested in your experience in construction zones on the State Highway System.

Have you traveled through a Florida Department of Transportation (FDOT) construction project in the last 12 months?

☐ YES **Please answer the following:**

We are interested in your recent experience with construction zones on the State Highway System near your area. Please indicate how strongly you agree with the following statements

Recent Experience with Construction Zones on State Highway System near my area:	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. When FDOT construction projects were initiated in my area, I was notified through various media (newspaper, radio, television, fliers, etc.)	☐	☐	☐	☐	☐
b. Construction zones were safe to travel through.	☐	☐	☐	☐	☐
c. When road construction was in progress, I was easily able	☐	☐	☐	☐	☐

It was clear to most respondent who received the FSU banner at the top that FSU was sponsoring their survey (75%). However, most individuals who had the FDOT logo did not report FDOT sponsorship (27.9%). Nonetheless, having an FDOT sponsored survey still increased response rates by about 2 percentage points, which is statistically significant ($p < .05$, two-tailed statistical test of the mean response rate differences). Thus, there is a bit of irony that many people failed to perceive the FDOT survey sponsorship but they were somewhat more likely to return the survey.

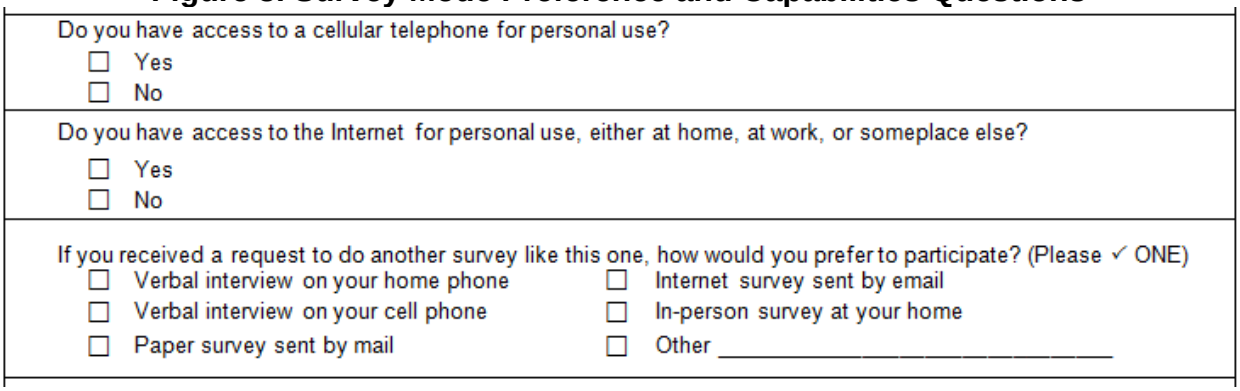
New Mode Potential and Preference Questions

Based upon the survey research literature (e.g., Olsen, Smyth, and Wood 2012), the pilot study also attempted to determine whether respondents preferred to be contacted in a particular manner. There were also questions about capabilities, such as whether individuals had access to a cell phone for personal use or the internet. A screenshot of the web survey appears below in Figure 3.

[FIGURE 3 APPEARS ON THE NEXT PAGE]

[49] The FSU surveys contained additional questions related to bike safety and some versions of the FSU surveys provided background information on bicycle safety as well. These results are beyond the scope of this report and will be detailed in other papers and projects.

Figure 3. Survey Mode Preference and Capabilities Questions



Nearly all respondents have access to a cell phone for personal use (95.6%, or 674 of the 705 mail respondents who answered the question). Similarly, 92% of the respondents had access to the Internet for personal use (651 of the 707 who responded). For the mode preference question, the vast majority (70%) of respondents prefer to be contacted via mail. The second most popular form of contact was via an internet survey sent by email (22.9%). In contrast, only 2.5% of respondents (n=17) wanted to be interviewed by telephone at home and only one respondent preferred to be interviewed by cell phone. The results are proportionally similar across the three types of mail survey respondent samples.

These results suggest that access is not an issue. When given a mail survey, most respondents prefer it. Of course, this is for the sample that chose to respond via mail. Still, it is not as if the respondents preferred some other mode. These findings are consistent with the published work by Olson, Smyth, and Wood (2012) who found that preference for web or phone predicts participation in these modes. Importantly, there was no clear effect for mail surveys. In particular, "…people who preferred phone or Web modes were just as likely to participate in a mail survey as those who preferred mail. As such, mail surveys appear to be robust to the effects of mode preference that we saw in other modes" (p. 631; emphasis added).

Discussion

In an effort to be responsive to the citizens of the state (i.e., the "customers"), the Florida Department of Transportation (FDOT) has invested considerable resources in determining satisfaction with its practices and efforts. Recently, however, changes in technology have made it more difficult to obtain responses from a representative sample. At the same time, people are busier than in the past; factors such as commute time are powerful predictors of survey nonresponse (i.e., individuals with commute times of over 45 minutes are less likely to respond). Ironically, people who commute often are likely to be the subgroup with experience and/or opinions about FDOT matters. Failing to include these people could have profound effects on transportation policy decisions. Irrespective of cause, diminishing survey diversity--particularly with respect to age, race, ethnicity, and gender--are concerns.

The FDOT surveys have historically been conducted via Random Digit Dial (RDD) calling to landline telephones. In the FDOT surveys and in others, telephone response rates are declining and unlikely to improve. If most people, and especially younger people, are abandoning landline phones, then a possible response would be to try to survey via wireless communication devices (e.g., cell phones). Including cell phones in addition to landlines is increasingly common but it does not solve the low response rate problem,[50] and doing so may even create additional problems.

First, it costs more to have survey researchers call cell phones and the end results may not produce higher response rates (e.g., Kohut et al 2012).[51] Second, respondents may incur charges of their own in responding to the call, although this is less of an issue with the increased popularity of unlimited calling plans. Third, and as reported earlier, many of cell phone numbers are geographically inaccurate. That is, since cell phone numbers are portable and many people do not change them when they change addresses, an attempt to call someone in a particular county or FDOT district could result in contacting someone who now lives in another county or state. Likewise, many people residing in an FDOT district might have cell phones registered in another state or locality, which would effectively exclude them from the survey since only some phone exchanges would be used. While the first problem is something that can be addressed via enhanced screening at the outset of the survey, the second problem is more difficult to correct since many people will simply go missing.[52] At the county-level, previous studies have found that more than 40% of the numbers are geographically inaccurate. Such inaccuracy would pose significant problems for the district-based FDOT surveys. Finally, and as discussed earlier, there are also considerable safety risks when it comes to surveying individuals via cell phone who may be driving. Given the mission of FDOT, these concerns regarding respondent safety are not trivial. Florida laws do not yet prohibit driving while talking on the phone,[53] but completing a typical FDOT survey on the phone will likely take more than ten minutes. That means some respondents will be driving at the time they are completing the survey, which could lead to traffic safety problems. Distracted drivers are a concern for state and local transportation authorities. The legal office at FDOT may want to consider whether having motorists respond to an FDOT survey while they are driving will create safety hazards that outweigh the benefits of increased representation for respondents who commute often.

[50] The Pew report (Kohut et al. 2012) found lower response rates among those reached via cell phones than in their landline samples, even though both groups were participating less relative to a decade ago.
[51] A conference presentation at the 2013 American Association for Public Opinion conference by Guterbock, Peytchev, and Rexrode reported that the number of hours per completion was roughly 50% higher for cell phones (1.04 vs. 1.52 on landline vs. cell) in 2013 and the number of completions is lower for cell phones (about 1 in 20 for cell phones vs. 1 in 15 for landlines). The team also reported that the costs for cell phone surveys of the same type were more in 2013 than in 2010.
[52] The extra screening to survey via cell phones will also increase costs.
[53] http://www.dmvflorida.org/cell-phone-distractions.shtml

Recommendations and Conclusions

Based upon the literature review and the pilot study, several recommendations are offered:

Mode. The mail surveys used here were relatively inexpensive and produced results comparable to other modes. Thus, FDOT should consider the use of mail surveys in the future. If resources permit, mixed modes (i.e., mail and phone or internet) could be used (e.g., Link and Lai 2011; Millar and Dillman 2011). On this point, it might be possible to use a design which gives people their choice of mode, either initially or after a screening period (e.g., Olson, Smyth, and Wood 2012). A related feature would be to conduct multiple waves (follow-up, possibly in a different mode). However, mixing modes can create difficulties when it comes to analyzing the data and in the creation of survey weights.

It is important to note that past efforts to explore mail and internet samples did not yield much in the way of diversity on race/ethnicity and age, especially when compared with the telephone surveys. An excerpt from the 2006 mode experiment report revealed the following:

- Mail survey and web-based panel respondents reflected less diversity with a higher percentage of whites [82% mail, 89% web-based panel vs. 73% phone] and fewer African-Americans [5% mail, 4% web-based panel vs. 9% phone] and fewer Hispanics [6% mail, 2% web-based panel vs. 11% phone].
- Respondents to the mail survey tended to be older residents with a higher proportion of those 65 and older [43% mail vs. 29% web-based panel and 25% phone] and a lower proportion of younger Floridians under the age of 35 [3% mail vs. 10% web-based and 13% phone.]

However, different sampling procedures might alter the results relative to what took place in 2006. For instance, the mail and internet surveys conducted as a part of the 2013 pilot study revealed much more diversity and representation among younger residents. Compared with the 2011 study, the sample had roughly the same proportion of respondents age 65 and older (42.5%). However, the internet respondents were younger (36.9% age 65 and older). The big improvements were among the youngest cohorts. Roughly 12% of the respondents in the mail survey were under age 35 in the 2013 pilot. This figure was the same in the internet sample, suggesting that young people participated in both modes and at higher rates than in the 2006 survey. Gender representation also improved in 2013.

High Quality Known Samples. Irrespective of the mode, it is helpful to know who is not responding. Traditional telephone surveys using RDD are often compared to Census figures and weights, if sufficient cases exist, can be created. However, in 2011 responses among the younger cohorts made it difficult to weight the estimates in this fashion. In contrast, starting from a known population has several advantages. Thus, one recommendation would be to work from established lists of Florida residents, either via the voter registration process or Address-Based Sampling from the U.S. postal service. The advantage of going with voter registration is that the list is (A) much

cheaper than ABS files, (B) there are physical addresses for individuals, and (C) a sizeable portion of the registered voters provide their telephone numbers and some counties (e.g., Leon) also attempt to collect email addresses; thus, a study that was based mainly on the mail mode could be extended to telephone or the internet.[54] Also, the voter registration lists are (a) regularly reviewed and "cleaned" by supervisors of elections, and (b) the files can be aggregated up to the household level by focusing on people who reside at the same address. Therefore, one recommendation is to sample from known populations (RBS or ABS).

There have been other attempts to use Registration-based Sampling (RBS). Tim Kellison of Florida State University sampled 5,000 registered voters from Columbus, Ohio. Kellison achieved a response rate close to 10% in a mail survey, but he used a lottery-style reward incentive, which may have increased participation. Also, the fact that the survey originated from Florida could have reduced response rates in a different state like Ohio. In contrast, a study by Jerit, Barabas, and Clifford (2013) in Leon County, Florida used voter registration lists as the sampling frame; for the mail survey conducted in the field, they witnessed a response rate of almost 20%. There were fairly dramatic differences between who was invited and who returned the survey, with most of the biases coming in previous voter turnout (i.e., those who returned the surveys were much more likely to have voted in the past, typically by 20 percentage points or more). Yet, demographically, there was a high degree of diversity. Roughly a quarter of the sample was black (compared with 35% among those who were sampled) and there was a high degree of participation among young respondents as well. In short, the combination of RBS and mail surveys has been tried before and these efforts have produced satisfactory results.

Oversampling. If there are remaining concerns about participation among subgroups, it might be fruitful to move to oversample these individuals. The RBS sampling frame permits explicit targeting of underrepresented groups, such as younger and minority communities.

Language Diversity. Employ multiple languages (English and Spanish). There is research suggesting that offering surveys in two languages boosts participation by Hispanic respondents without reducing responses among non-Hispanics (e.g., Brick and Williams 2012). However, additional languages will mean added costs.

Participation Enhancements. There are other measures, not undertaken here, which can boost participation. Aside from monetary incentives, survey researchers can send letters to respondents to warn them about the coming interview request (De Leeuw et al. 2007). Also, making the topic prominent and salient may help increase response rates as well as highlighting the FDOT origins of the survey as discovered in the sponsorship experiment.

[54] As an example of the possibilities for the telephone, 29% of individuals in the 9 Florida counties have a seven digit phone number in their record for the FDOE data (that is, 715,269 of the 2,440,388).

In the end, resources may constrain research design decisions. Telephone surveys conducted via landline phones have been economical in the past. There are reasons to think that future phone surveys will be more expensive compared to what they were several years ago. The inability to reach respondents, via landlines or cell phones, will increase costs. As mentioned earlier, people are switching to cell phones but surveying in this manner is often more costly, potentially 40-50% more. At the same time, the cost of conducting mail and internet surveys has declined. The 2013 pilot study costs were estimated at nearly $24 per completion in the primary mail sample survey. Yet, the bulk of the completions were via mail. Part of the cost savings may be due to the low cost of acquiring samples from the FDOE relative to samples procured from private firms. Viewed in this manner, mail surveys may be more cost effective than telephone surveys in the future. However, there are numerous design decisions and enhancements that could make surveys using any mode more or less expensive. Also, costs are but one consideration along with others, such as representativeness and bias.

This report highlighted many difficulties in the field of survey research. One major factor is that many individuals are abandoning their landline telephones. This trend has repercussions for how surveys are conducted and who can be reached. A central concern is sample representativeness. This report presents a review of recent and relevant literature, along with some recommendations drawn from its studies, in order to assist FDOT with the arduous task of designing surveys and collecting representative, unbiased data. Ultimately, the most important lesson that should be taken from this review is that when surveying, one must exercise caution in every step of the way. Biases can arise for many reasons, from the data collection methods, to the format of the questions, the interviewers' techniques and incentives, and even the language of the survey. In the end, the recommendations outlined here should help improve future FDOT surveys across a wide array of criteria, especially as they relate to response rates and improving representation of minority groups as well as younger individuals.

Appendix

Cell-Phone Only Population by State, Race/Ethnicity, and Age

The text reports estimates on cell-phone only population from Blumberg and Luke 2013. Elsewhere, Blumberg et al. (2012) report model estimates from household telephone status for adults aged 18 and over. The table below is an excerpt of the Blumberg et al. report for three states: Florida, Alabama, and Georgia.

Appendix Table 1. Household Telephone Status for Adults Aged 18 and Over by Geograhphic Area, 2011

	Percent (standard error)						
	Wireless Only	Wireless Mostly	Dual Use	Landline Mostly	Landline Only	No Telephone Service	Total
Florida	34.4 (1.2)	17.3 (0.9)	25.7 (1.1)	11.8 (0.8)	8.8 (0.7)	2.1	100
Miami-Dade County	35.9 (3.2)	13.8 (2.3)	26.5 (3.4)	13.0 (2.3)	8.7 (2.2)	2.0	100
Duval County	41.3 (2.1)	17.7 (1.7)	22.2 (2.0)	7.0 (1.1)	8.7 (1.3)	3.2	100
Orange County	39.7 (3.2)	20.4 (2.6)	22.3 (3.1)	10.6 (2.0)	5.1 (1.7)	1.9	100
Rest of Florida	32.7 (1.4)	17.2 (1.2)	26.5 (1.4)	12.7 (1.0)	9.1(0.9)	1.9	100
Alabama	34.4 (1.8)	14.4(1.4)	25.3 (1.8)	15.6 (1.4)	8.3 (1.2)	1.9	100
Jefferson County	41.9 (2.9)	14.0 (2.0)	23.4 (2.7)	12.3 (1.9)	6.9 (1.7)	1.5	100
Rest of Alabama	33.2 (2.0)	14.5 (1.5)	25.6 (2.1)	16.2 (1.6)	8.6 (1.3)	1.9	100
Georgia	34.3 (1.6)	21.6 (1.4)	23.7 (1.6)	11.4 (1.1)	6.7 (0.9)	2.3	100
Fulton/DeKalb Counties	41.9 (2.9)	20.4 (2.4)	23.3 (2.8)	7.8 (1.6)	4.7 (1.4)	1.9	100
Rest of Georgia	32.7 (1.8)	21.8 (1.6)	23.8 (1.8)	12.2 (1.3)	7.2 (1.1)	2.3	100

Source: Blumberg et al. 2012.

As Appendix Table 1 shows, the overall estimates for Florida are similar to the statewide distributions in Alabama and Georgia. However, there is considerable variation within and across populous counties. In general, urban areas have higher percentages of adults who are wireless only. Appendix Table 2 shows the wireless-only population by racial or ethnic group.

Appendix Table 2. Percentage of Adults Living in Wireless-Only Households by Race/Ethnicity, 2009-2012

	Jan-Jun 2009	Jul-Dec 2009	Jan-Jun 2010	Jul-Dec 2010	Jan-Jun 2011	Jul-Dec 2011	Jan-Jun 2012	Jul-Dec 2012
Hispanic or Latino, Any Race(s)	28.2	30.4	34.7	38.4	40.8	43.3	46.5	50.5
Non-Hispanic Black, Single Race	21.3	25.0	28.5	31.1	32.5	36.8	37.7	39.0
Non-Hispanic White, Single Race	19.7	21.0	22.7	25.0	27.6	29.0	30.4	32.9
Non-Hispanic Asian, Single Race	18.0	20.6	18.8	27.0	27.7	31.6	33.4	34.4
Non-Hispanic Other, Single Race	20.6	26.5	16.1	31.9	33.8	44.1	43.4	43.9
Non-Hispanic, Multple Races	28.7	26.9	36.0	36.1	39.3	36.7	40.2	45.3

Source: Blumberg et al. 2012.

Appendix Table 3 shows the trend for age categories during the same time period. The results reveal that the wireless-only population has increased over the three year time span. The increases are especially large for minorities and younger cohorts.

Appendix Table 3. Percentage of Adults Living in Wireless-Only Households by Age Category, 2009-2012

	Jan-Jun 2009	Jul-Dec 2009	Jan-Jun 2010	Jul-Dec 2010	Jan-Jun 2011	Jul-Dec 2011	Jan-Jun 2012	Jul-Dec 2012
Age 18-24	37.6	37.8	39.9	45.5	46.8	48.6	49.5	53.2
Age 25-29	45.8	48.6	51.3	53.5	58.1	59.6	60.1	62.1
Age 30-34	33.5	37.2	40.4	43.8	46.2	50.9	55.1	56.7
Age 35-44	21.5	23.9	27	30.9	34.3	36.8	39.1	43.5
Age 45-64	12.8	14.9	16.9	18.8	21.6	23.8	25.8	28.4
Age 65 and older	5.4	5.2	5.4	7.7	7.9	8.5	10.5	11.6

Source: Blumberg et al. 2012.

Post-stratification Weights

When conducting a survey, having a representative sample of the population is important. But sometimes, despite best efforts, researchers may *oversample* some kinds of people and *undersample* others. In other words, the way a certain characteristic (such as age, education, race, sex, etc.) of a particular sample is distributed may differ from the way it is distributed in the population. For example, the sample may consist of 60 percent women, when women make up only 52 percent of the population. This can introduce bias into any estimates obtained from the sample data because statistical procedures will give greater weight to subgroups oversampled. In these situations, researchers can correct for these biases mathematically with a post-stratification survey weight.

In order to calculate a post-stratification weight, one must have an auxiliary data set (or “benchmark” data set) to compare and adjust the sample data. For example, for a survey of residents of Florida, one could use Census data or Current Population Survey data that show the demographic characteristics of the population of the state of Florida. If such a data set exists, the sample can be compared to the auxiliary data file, to make sure that the distribution of demographic characteristics (such as age, education, race, sex, etc.) is similar to that of the auxiliary data. If the distributions are close enough, there is probably no need to calculate post-stratification weights (or the benefit it likely to be small). However, if the two data sets differ by more than a few percentage points, it will often be helpful to calculate the weights. This kind of weight is called a *post*-stratification weight because it is created *after* the data have been collected. It is stratified because various known *strata* (such as age group or sex distribution) of the population are used to adjust the sample data to conform more to the population's parameters. The table below illustrates the steps needed to calculate a post-stratification weight by age group and sex.

Appendix Table 4. Illustration of Survey Weights

	Col. A	Col. B	Col. C	Col. D	
Age Group & Sex	Population N	Proportion of Population **Col. A / 8000**	Sample n	Proportion of Sample **Col. C / 200**	Weight **Col. B / Col. D**
Age 18-34: Male	2,000	.25	30	.15	1.67
Age 18-34: Female	1,760	.22	40	.20	1.10
Age 35-64: Male	1,440	.18	50	.25	.72
Age 35-64: Female	1,600	.20	60	.30	.67
Age 65+: Male	400	.05	10	.05	1.00
Age 65+: Female	800	.10	10	.05	2.00
Total	**8,000**		**200**		

The Catalist Database

Based in Washington, D.C., Catalist is one of the nation's leading providers of voter information. In particular, Catalist sells annual subscriptions for access to their continually updated database for the entire U.S. (all 50 States, plus the District of Columbia) with over 180 million registered voters. Importantly, Catalist also maintains data on 85 million unregistered adults, which is an extremely hard-to-reach and potentially unknowable population. Other voter-file vendors typically assemble their database only immediately prior to election cycles or when customer demand for the data is high. Other attractive features of Catalist include "synthetic data" through modeling offering insights about persons such as relative likelihood to turnout to vote, likelihood to be married or have a college degree, to name just a few. Catalist gives the capacity to generate individual-level models of partisanship, voter turnout likelihood, and propensity to take action or support a variety of issues.

Current clients of Catalist include the nationally known polling firm of Greenberg, Quinlan, and Rosner Research, Women's Voices/Women Vote, The Center on Budget and Policy Priorities, Emily's List, Families USA, the League of Conservation Voters, the Sierra Club, the ACLU, NARAL-Prochoice America Foundation, and more than two dozen U.S. representatives. In addition, data from Catalist was used by the Justice Department in its successful challenge to the Texas Voter ID law.

In addition to voting information, the Catalist subscription includes extensive geographical information system data. Their Q-tool application enables customers to directly and rapidly query (across states and nationwide), view, visualize, and export a wide range of data points that are synthesized into a standard format from multiple sources. Also, many Catalist subscriptions include the license to receive an analytics export of over 700 fields of commercial, census, and synthetic data, so that clients (and their consultants) may build their own custom models. Catalist stores customer data separate and apart from the Catalist database. Finally, the I/O Tool is a self-service web based interface that allows subscribers to match their data to Catalist's national database. For more information, see http://catalist.us/product.

References

Abraham, Katherine G., Aaron Maitland, and Suzanne M. Bianchi. 2006. "Nonresponse in the American Time Use Survey: Who is Missing From The Data and How Much Does it Matter?" *Public Opinion Quarterly* 70: 676-703

Aday, Lu Ann,and Llewllyn J. Cornelius. 2006. Designing and Conducting Health Surveys: A Comprehensive Guide, 3rd ed. San Francisco, CA: Jossey-Bass.

American Association for Public Opinion Research. 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. *7th Edition*. AAPOR.

American Association for Public Opinion Research. 2010. *AAPOR Report on Online Panels*. Deerfield, IL: AAPOR.

Andridge, R.H., and Roderick J. Little. 2009. "The Use of Sample Weights in Hot Deck imputation." *Journal of Official Statistics* 25 (1): 21-36.

Ansolabehere, Stephen, and Brian F. Schaffner. 2010. "Residential Mobility, Family Structure, and the Cell-Only Population." *Public Opinion Quart*erly 74 (Summer): 244-59.

Baker, Reg, Stephen J. Blumberg, J. Michael Brick, Mick P. Couper, Melanie Courtright, J. Michael Dennis, Don Dillman, Martin R. Frankel, Philip Garland, Robert M. Groves, Courtney Kennedy, Jon Krosnick, and Paul J. Lavrakas. 2010. "Research Synthesis: AAPOR Report on Online Panels." *Public Opinion Quarterly* 74 (Winter): 711-81.

Bates, Nancy, James Dahlhamer, and Elinor Singer. 2008. "Privacy Concerns, Too Busy, or Just Not Interested: Using Doorstep Concerns to Predict Survey Nonresponse." *Journal of Official Statistics* 24 (4): 591-612.

Battaglia, Michael P., Michael W. Link, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* 72 (Fall): 459-69.

Barber, Michael, Chris Mann, J. Quinn Monson, and Kelly D. Patterson. 2011. "Conducting Online Pre-Election Polls Using Registration Based Sampling." Paper Presented at the American Association for Public Opinion Research, Phoenix, AZ.

Benford, R., L. Piekarski, J. Lien, and Trevor Thompson. 2012. "Geographic Differences between RDD Cell and Landline Frames and Self-Report." Paper presented at the American Association for Public Opinion Research Conference, Orlando, FL, May 2012.

Berinsky, Adam J. 2005. *Silent Voices: Public Opinion and Political Participation in America*. Princeton: Princeton University Press.

Biemer, Paul B., and Andy Peytchev. 2012. “Census Geocoding for Nonresponse Bias Evaluation in Telephone Surveys: An Assessment of the Error Properties.” *Public Opinion Quarterly* 76 (Fall): 432-52.

Blom, Alan, and Kreuter, F. 2011. Preface*. Journal of Official Statistics* 27(2):151–152.

Blumberg, Stephen J., and Julian V. Luke. 2013. “Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, July-December 2012.” Available at http://www.cdc.gov/nchs/nhis.htm.

Blumberg, Stephen J., and Julian V. Luke. 2012. “Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, January–June 2012.” Available at http://www.cdc.gov/nchs/nhis.htm.

Blumberg, Stephen J., and Julian V. Luke. 2011. “Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, January–June 2011.” Available at http://www.cdc.gov/nchs/nhis.htm.

Blumberg, Stephen J., and Julian V. Luke. 2009. “Wireless Substitution: Early Release of Estimates Based on Data from the National Health Interview Survey, January–June 2009.” Available at http://www.cdc.gov/nchs/nhis.htm.

Blumberg, Stephen J. and Julian V. Luke. 2007. "Coverage Bias in Traditional Telephone Surveys of Low-Income and Young Adults." *Public Opinion Quarterly* 71: 734-49.

Blumberg, Stephen J., Julian V. Luke, Nadarajasundaram Ganesh, Michael E. Davern, and Michael H. Boudreaux. 2012. “Wireless Substitution: State-Level Estimates from the National Health Interview Survey, 2010-2011.” In *National Health Statistics Reports*. Hyattsville, MD: National Center for Health Statistics.

Blumberg, Stephen J., Julian V. Luke, Gestur Davidson, Michael E. Davern, Tzy-Chyi Yu, and Karen Soderberg. 2009. “Wireless Substitution: State-Level Estimates from the National Health Interview Survey, January–December 2007.” In *National Health Statistics Reports*. Hyattsville, MD: National Center for Health Statistics.

Blumberg, Stephen J. and Julian V. Luke. 2007. "Coverage Bias in Traditional Telephone Surveys of Low-Income and Young Adults." *Public Opinion Quarterly* 71: 734-49.

Bouffard, Julie, and Jennifer Tancreto. 2006. “2005 National Census Test: Bilingual Form Analysis.” 2010 Census Test Memoranda Series, No. 22. Washington, DC: U.S. Census Bureau.

Boyle, John, Michael Bucuvalas, Linda Piekarski, and Andy Weiss. 2009. "Zero Banks: Coverage Error and Bias in RDD Samples Based on Banks with Listed Numbers." *Public Opinion Quarterly* 73 (Winter): 729-50.

Brennan, Mike, and Jan Charbonneau. 2009. "Improving Mail Survey Response Rates Using Chocolate and Replacement Questionnaires." *Public Opinion Quarterly* 73 (Summer): 368-78

Brick, J. Michael, and Douglas Williams. 2013. "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys." *The Annals of the American Academy of Political and Social Science* 645: 36 Thousand Oaks, CA: Sage Publications.

Brick, J. Michael, Douglas Williams, and Jill M. Motaquila. 2011. "Address-Based Sampling for Subpopulation Surveys." *Public Opinion Quarterly* 75 (Fall): 409-28.

Brick, J. Michael, Ismael Flores Cervantes, Sunghee Lee, and Greg Norman. 2011. "Nonsampling Errors in Dual-Frame Telephone Surveys." *Survey Methodology* 37(1):1–12.

Brick, J. Michael, W. Sherman Edwards, and Sunghee Lee. 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. Michael, Sarah Dipko, Stanley Presser, Clyde Tucker, and Yangyang Yuan. 2006. "Nonresponse Bias in a Dual-Frame Sample of Cell and Landline Number." *Public Opinion Quarterly* 70(5):780–93.

Brick, J. Michael, Jill M. Montaquila, Daifeng Han, and Douglas Williams. 2012. "Improving Response Rates for Spanish Speakers in Two-Phase Mail Surveys." *Public Opinion Quarterly* 76 (Winter): 721-32.

Callegargo, Mario, and Charles Disogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72: 1008-32.

Chang, Linchiat, and Jon A. Krosnick. 2010. "Comparing Oral Interviewing with Self-Administered Computerized Questionnaires: An Experiment." *Public Opinion Quarterly* 74 (Spring): 154-67.

Chang, Linchiat, and Jon A. Krosnick. 2009. "National Surveys Via RDD Telephone Interviewing Versus the Internet: Comparing Sample Representativeness And Response Quality." *Public Opinion Quarterly* 73 (Winter): 641-78.

Christian, Leah Melanni, Don A. Dillman, and Jolene D. Smyth. 2007. "Helping Respondents Get It Right The First Time: The Influence of Words, Symbols, and Graphics in Web Surveys." *Public Opinion Quarterly* 71 (Spring): 113-25.

Christian, Leah M., Michael Dimock, and Scott Keeter. 2009. “Accurately Locating Where Wireless Responses Live Requires More than a Phone Number.” Pew Research Center for People and the Press. Available at: http://www.pewresearch.org/2009/07/09/accurately-locating-where-wireless-respondents-live-requires-more-than-a-phone-number/

Cobben, F. and Schouten, B. 2007 An Empirical Evaluation of R-Indicators. Discussion Paper 08006. Statistics Netherlands. Voorburg/Heerlen. Retrieved from: http://www.risq-project.eu/papers/cobben-schouten-2008-a.pdf

Couper, Mick P. 1998. “Measuring Survey Quality in a CASIC Environment.” *Proceedings of the Survey Methodology Section, American Statistical Association*: 41–49.

Couper, Mick P., Frederick G. Conrad, and Roger Tourganeau. 2007. "Visual Context Effects in Web Surveys." *Public Opinion Quarterly* 71 (Winter): 623-34.

Couper, Mick P., and Lars Lyberg. 2005. “The Use of Paradata in Survey Research.” *Proceedings of the 55th Session of the International Statistical Institute*, Sydney, Australia.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. “Effects of Response Rate Changes on the Index of Consumer Sentiment.” *Public Opinion Quarterly* 64:413–28.

Curtin, R., Stanley Presser, and Eleanor Singer. 2005. “Changes in Telephone Survey Nonresponse over the Past Quarter Century.” *Public Opinion Quarterly* 69(1):87–98.

De Leeuw, Edith, Mario Callegaro, Joop Hox, Elly Korendijk, and Gerty Lensvelt-Mulders. 2007. "The Influence Of Advance Letters on Response in Telephone Surveys: A Meta-Analysis." *Public Opinion Quarterly* 71 (Fall): 413-43.

De Leeuw, Edith D. & De Heer, W. 2002. “Trends in Household Survey Nonresponse: A Longitudinal and International Comparison.” In: R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J.A. Little, eds. *Survey Nonresponse*. New York: Wiley, pp. 41-54.

Dillman, Donald A. 1978. Mail and Telephone Surveys: The Total Design Method. New York: Wiley.

Dillman, Donald A. 1991. “The Design and Administration of a Mail Survey.” *Annual Review of Sociology* 17: 225-49.

Dillman, Donald A. 1999. Mail and Internet Surveys – The Tailored Design Method, 2nd ed. New York: Wiley.

Ehlen, John, and Patrick Ehlen. 2007. "Cellular-Only Substitution In the United States as Lifestyle Adoption: Implications for Telephone Survey Coverage." *Public Opinion Quarterly* 71: 717-33.

Fullerton, Andrew S., Jeffery C. Dixon, and Casey Borch. 2007. "Bringing Registration into Models of Vote Overreporting." *Public Opinion Quarterly* 71 (Winter): 649-60.

Fricker, Scott, and Roger Tourangeau. 2010. "Examining the Relationship Between Nonresponse Propensity and Data Quality in Two National Household Surveys." *Public Opinion Quarterly* 74: 934-55.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Galesic, Mirta, and Michael Bosnjak. 2009. "Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey." *Public Opinion Quarterly* 73 (Summer):349-60.

Green, Donald P., and Alan S. Gerber. 2006. "Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts?" *Public Opinion Quarterly* 70 (Summer): 197-223.

Greene, William. 2012. *Econometric Analys*is. 7th edition. New York: Prentice Hall

Groves, Robert. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70: 646-75.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley.

Groves, Robert M., and Mick Couper. 1998. *Nonresponse in Household Interview Surveys.* New York: Wiley.

Groves, Robert M., Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Giorgina Piani Acosta, and Lindsay Nelson. 2006. "Experiments in Producing Nonresponse Bias." *Public Opinion Quarterly* 70: 720-36.

Groves, Robert M., Stanley Presser, Roger Tourangeau, Brady T. West, Mick P. Couper, Eleanor Singer, and Christopher Toppe. 2012. “Support for the Survey Sponsor and Nonresponse Bias.” *Public Opinion Quarterly* 76 (Fall): 512-24.

Groves, Robert M., and Lars Lyberg. 2010. “Total Survey Error: Past, Present, and Future.” *Public Opinion Quarterly* 74 (5): 849-79.

Groves, Robert, and Emilia Peytcheva. 2008. “The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis.” *Public Opinion Quarterly* 72:167–89.

Groves, Robert, Stanley Presser, and Sarah Dipko. 2004. “The Role of Topic Interest in Survey Participation Decisions.” *Public Opinion Quarterly* 68: 2–31.

Groves, R.M., and Heeringa, S. 2006 Responsive Design for Household Surveys: Tools for Actively Controlling Survey Costs and Errors.” *Journal of the Royal Statistical Society* 169(3):439–457.

Guo, Shenyang, and Mark W. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Applications.* Thousand Oaks, CA: Sage.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 4**7** (1): 153–61.

Heerwegh, Dirk, and Geert Loosveldt. 2008. " Face-to-Face Versus Web Surveying in a High Internet-Coverage Population: Differences in Response Quality." *Public Opinion Quarterly* 72: 836-46.

Herbst, Susan. 1995. How Public Opinion Polling Has Shaped American Politics. Chicago: University of Chicago Press.

Holbrook, Allyson L., Jon A. Krosnick, David Moore, and Roger Tourangeau. 2009. "Response Order Effects in Dichotomous Categorical Questions Presented Orally: The Impact of Question and Respondent Attributes." *Public Opinion Quarterly* 71 (Fall): 325-48.

Hox, Joop J., Edith D. de Leeuw, and Hsuan-Tzu Chang. 2012. “Nonresponse versus Measurement Error: Are Reluctant Respondents Worth Pursuing?” *Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique* 113 (1): 5-19.

Iannacchione, Vincent G., Jennifer M. Staab, and David T. Redden. 2003. "Evaluating the Use of Residential Mailing Addresses in A Metropolitan Household Survey." *Public Opinion Quarterly* 67: 202-10.

Jerit, Jennifer, Jason Barabas, and Scott Clifford. 2013. “Comparing Contemporaneous Laboratory and Field Experiments on Media Effects.” *Public Opinion Quarterly*.

Keeter, Scott. 2012. “Presidential Address: Survey Research, Its New Frontiers, and Democracy.” *Public Opinion Quarterly* 76 (Fall): 600-08.

Keeter, Scott. 2006. "The Impact of Cell Phone Noncoverage Bias on Polling in the 2004 Presidential Election." *Public Opinion Quarterly* 70 (Spring): 88-8.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert Groves, and Stanley Presser. 2000.“Consequences of Reducing Nonresponse in a Large National Telephone Survey.” *Public Opinion Quarterly* 64:125–48.

Keeter, Scott, Courtney Kennedy, April Clark, Trevor Tompson, and Mike Mokrzycki. 2007. What's Missing from National Landline RDD Surveys?: The Impact of the Growing Cell-Only Population." *Public Opinion Quarterly* 71: 772-92.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. 2006. "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey." *Public Opinion Quarterly* 70: 759-79.

Kennedy, Courtney. 2007. "Evaluating the Effects of Screening for Telephone Service in Dual Frame RDD Surveys." *Public Opinion Quarterly* 71: 750-71.

Kennedy, Courtney, Scott Keeter, and Michael Dimock. 2008. "A 'Brute Force' Estimation of the Residency Rate for Undetermined Telephone Numbers in an RDD Survey." *Public Opinion Quarterly* 72: 28-9.

Kish, Leslie. 1965. *Survey Sampling*. New York: Wiley.

Kohut, Andrew, Scott Keeter, Carroll Doherty, Michael Dimock, and Leah Christian. 2012. "Assessing the Representativeness of Public Opinion Surveys." Washington, DC: Pew Research Center.

Lavrakas, Paul J., Charles D. Shuttles, Charlotte Steeh, and Howard Fienberg. 2007. "The State of Surveying Cell Phone Numbers in the United States: 2007 and Beyond." *Public Opinion Quarterly* 71: 840-54.

Link ,Michael W., Michael P. Battaglia, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2007. "Reaching the U.S. Cell Phone Generation: Comparison of Cell Phone Survey Results With an Ongoing Landline Telephone Survey." *Public Opinion Quarterly* 71:814-39.

Link ,Michael W., Michael P. Battaglia, Martin R. Frankel, and Larry Osborn. 2008. "A Comparison of Address-Based Sampling (ABS) Versus Random-Digit Dialing (RDD) For General Population Surveys." *Public Opinion Quarterly* 72: 6-27.

Link, Michael w., and Jennie E. Lai. 2011. "Cell -Phone-Only Households and Problems of Differential Nonresponse Using an Address-Based Sampling Design." *Public Opinion Quarterly* 75 (Winter): 613-35.

Lohr, Sharon L. 2010. *Sampling: Design and Analysis*, 2nd ed. Boston: Cengage.

Lynn, Peter, and Gerry Nicolaas. 2010. "Making Good Use of Survey Paradata." *Survey Practice* 3 (April).

Malhotra, Neil. 2008. "Completion Time and Response Order Effects in Web Surveys." *Public Opinion Quarterly* 72: 914-34.

Malhotra, Neil. 2009. "Order Effects in Complex and Simple Tasks." *Public Opinion Quarterly* 73 (Spring): 180-98.

Malhotra, Neil, Jon A. Krosnick, and Randall K. Thomas. 2009. "Optimal Design of Branching Questions to Measure Bipolar Constructs." *Public Opinion Quarterly* 73 (Summer): 304-24.

Martin, Elizabeth. 2009. "Can A Deadline and Compressed Mailing Schedule Improve Mail Response in the Decennial Census?" *Public Opinion Quarterly* 73 (Summer): 361-67.

Massey, Douglas S., and Roger Tourganeau. 2013a. “Introduction-New Challenges to Social Measurement.” *Annals of the American Academy of Political and Social Science* 645 (Jan.): 6-22.

Massey, Douglas S., and Roger Tourganeau. 2013b. “Where Do We Go From Here? Nonresponse and Social Measurement.” *Annals of the American Academy of Political and Social Science* 645 (Jan.): 222-36.

Maynard, Douglas W., Jeremy Freese, and Nora C. Schaeffer. 2010. “Calling for Participation: Requests, Blocking moves, and Rational (inter)action in Survey Introductions. *American Sociological Review* 75(5):791–814.

McDonald, Michael P. 2007. "The True Electorate: A Cross-Validation of Voter Registration Files and Election Survey Demographics." *Public Opinion Quarterly* 71 (Winter): 588-602.

Medway, Rebecca, and Jenna Fulton. 2012. “When More Gets You Less: A Meta-Analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates.” *Public Opinion Quarterly* 76 (Winter): 733-46.

Merkle, Daniel, and Murray Edelman. 2002. “Nonresponse in Exit Polls: A Comprehensive Analysis.” In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, pp. 243–57. New York: Wiley.

Merkle, Daniel M., Gary Langer, Jon Cohen, Linda B. Piekarski, Robert Benford, and David Lambert. 2009. "Unintended Consequences: The Cost of Purging Business Numbers in RDD Surveys." *Public Opinion Quarterly* 73 (Fall): 484-96.

Messer, Benjamin L., and Don A. Dillman. 2011. "Surveying the General Public Over the Internet Using Address-Based Sampling and Mail Contact Procedures." *Public Opinion Quarterly* 75 (Fall): 429-57.

Millar, Morgan M., and Don A. Dillman. 2011. "Improving Response to Web and Mixed-Mode Surveys." *Public Opinion Quarterly* 75 (Summer): 249-69.

Mokrzycki, Michael, Scott Keeter, and Courtey Kennedy. 2009. "Cell-Phone-Only Voters in the 2008 Exit Poll and Implications for Future Noncoverage Bias." *Public Opinion Quarterly* 73: 845-65.

Montaquila, Jill M., J. Michael Brick, Douglas Williams, Kwang Kim, and Daifeng Han. 2013. “A Study of Two-Phase Mail Survey Data Collection Methods.” *Journal of Survey Statistics and Methodology* 1: 66-87.

Montgomery, R., X. Tao, K. Wolter, S. Greby, and E. Kennedy. 2011. “Accuracy of Geographic Stratification in a Cell-Phone Survey.” Paper presented at the American Association for Public Opinion Research Conference, Phoenix, AZ, May 2011.

National Research Council. 2013*. Nonresponse in Social Science Surveys: A Research Agenda.* Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Nickerson, David. 2007. “Quality if Job One: Professional and Volunteer Voter Mobilization Calls.” *American Journal of Political Science* 51 (April): 269-82.

Olson, Kristen. 2013. “Paradata for Nonresponse Adjustment.” *The Annals of the American Academy of Political and Social* Science 645 (Jan.): 142-170.

Olson, Kristen. 2006. "Survey Participation, Nonresponse Bias, Measurement Error Bias, ant Total Bias." *Public Opinion Quarterly* 70: 737-58.

Olson, Kristen, Jolene D. Smyth, and Heather M. Wood. 2012. “Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination.” *Public Opinion Quarterly* 76 (Winter): 611-35.

Petrolia, Daniel R., and Sanjoy Bhattacharjee. 2009. "Revisiting Incentive Effects: Evidence From a Random-Sample Mail Survey on Consumer Preferences for Fuel Ethanol." *Public Opinion Quarterly* 73 (Fall): 537-50.

Peytchev, Andy. 2013. “Consequences of Survey Nonresponse.” *The Annals of the American Academy of Political and Social Science.* 645: 88-111.

Peytchev, Andy. 2012. “Multiple Imputation for Unit Nonresponse and Measurement Error.” *Public Opinion Quarterly* 76 (June): 214-37.

Peytchev, Andy, Rodney K. Baxter, and Lisa R. Carley-Baxter. 2009. "Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error." *Public Opinion Quarterly* 73 (Winter): 785-806.

Peytchev, Andy, and Benjamin Neely. 2013. “RDD Telephone Surveys: Toward a Single-Frame Cell-Phone Design.” *Public Opinion Quarterly* 77 (2): 1-22.

Porter, Stephen R., and Michael E. Whitcomb. 2007. "Mixed-Mode Contacts in Web Surveys: Paper is Not Necessarily Better." *Public Opinion Quarterly* 71 (Winter): 635-48.

Presser, Stanley, and Susan McCulloch. 2011. “The Growth of Survey Research in the United States: Government-Sponsored Surveys, 1984–2004.” *Social Science Research* 40:1019–1024.

Presser, Stanley, and Eleanor Singer. 2007. “Privacy, Confidentiality, and Respondent Burden as Factors in Telephone Surveys.” In J. Lepkowski et al. (eds.). *Advances in Telephone Survey Methodology*. New York: Wiley, pp. 449-70.

Presser, Stanley, Johnny Blair, and Timothy Triplett. 1992. “Survey Sponsorship, Response Rates, and Response Effects.” *Social Science Quarterly* 73:699–702.

Putnam, Robert. 2001. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Putnam, Robert. 1995. “Bowling Alone: America's Declining Social Capital.” *Journal of Democracy* 6(1):65–78.

Rao, Kumar, Olena Kaminska, and Allan L. McCutcheon. 2010. "Recruiting Probability Samples for a Multi-Mode Research Panel with Internet and Mail Components." *Public Opinion Quarterly* 74 (Spring): 68-84.

Sakshaug, Joseph W., Ting Yan, and Roger Tourangeau. 2010. "Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items." *Public Opinion Quarterly* 74: 907-33.

Sarndal, C. 2011. “The 2010 Morris Hansen Lecture Dealing with Survey Nonresponse in Data Collection, in Estimation. *Journal of Official Statistics* 27(1):1–21.

Scherpenzeel, Annette, and Vera Toepoel. 2012. “Recruiting a Probability Sample for an Online Panel: Effects of Contact Mode, Incentives, and Information.” *Public Opinion Quarterly* 76 (Fall): 470-90.

Schwartz, Doug, and Clay F Richards. 2003. “The Big East Two: A Comparison of RBS and RDD polls in the 2002 elections in New York and Pennsylvania.” Presented at the Annual Meeting of the American Political Science Association, Philadelphia.

Singer, Eleanor. 2006. Introduction—Nonresponse Bias in Household Surveys. *Public Opinion Quarterly* 70(5):637-645. Special Issue 2006.

Singer, Eleanor, and Cong Ye. 2013. “The Use and Effects of Incentives in Surveys.” *The Annals of the American Academy of Political and Social* Science 645 (Jan.): 112-141.

Skalland, Benjamin, and Meena Khare. 2013. “Geographic Inaccuracy of Cell Phone Samples and the Effect on Telephone Survey Bias, Variance, and Cost.” *Journal of Survey Statistics and Methodology* 1: 45-65.

Smyth, Jolene D., Don A. Dillman, Leah Melani Christian, and Michael J. Stern. 2006. "Comparing Check-All and Forced-Choice Question Formats in Web Surveys." *Public Opinion Quarterly* 70 (Spring): 66-77.

Stoop, Ineke. 2005. *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Hague: Social and Cultural Planning Office.

Stoop, Ineke, Billiet, J., Koch, A., and Fitzgerald, R. 2010. *Improving Survey Response: Lessons Learned from the European Social Survey*. New York: Wiley.

Toepoel, Vera, and Mick P. Couper. 2011. "Can Verbal Instructions Counteract Visual Context Effects in Web Surveys?" *Public Opinion Quarterly* 75 (Spring): 1-18.

Toepoel, Vera, Marcel Das, and Arthur Van Soest. 2008. "Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents." *Public Opinion Quarterly* 72: 985-1007

Tourangeau, Roger, Frauke Kreuter, and Stephanie Eckman. 2012. “Motivated Underreporting in Screening Interviews.” *Public Opinion Quarterly* 76 (Fall): 453-69.

Tucker, Clyde, J. Michael Brick, and Brian Meekins. 2007. "Household Telephone Service and Usage Patterns in the United States in 2004: Implications for Telephone Samples." *Public Opinion Quarterly* 71 (Spring): 3-22.

Van der Grijn, F., Schouten, B.,and Cobben, F. 2006. Balancing Representativity, Costs and Response Rates in a Call Scheduling Strategy. Paper presented at the 17th International Workshop on Household Survey Nonresponse. Omaha, NE., August 28-30

Wagner, James. 2012. “Research Synthesis: A Comparison of Alternative Indicators for the Risk of Nonresponse Bias.” *Public Opinion Quarterly* 76 (Fall): 555-75.

Wagner, James. 2013. “Alternative Measures for the Risk of Nonresponse Bias. Presentation to the Panel on a Research Agenda for the Future of Social Science

Data Collection," February 18, Washington, DC. Retrieved from: http://www7.nationalacademies.org/cnstat/1Wagner%20Pres.pdf

Weisberg, Herbert L. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press.

Yeager, David S., Jon A. Krosnick, Linchiat Chang, Harold S. Javitz, Matthew S. Levendusky, Alberto Simpser, and Rui Wang. 2011. "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples." *Public Opinion Quarterly* 75 (Winter): 709-47.