Oxford Handbooks Online

Abstract and Keywords

This chapter argues for the use of creative and targeted strategies rather than a one-size-fits-all approach to capturing information on low-incidence populations. It first conceptualizes low-incidence populations and highlights the existing empirical literature on them and then frames the challenges of polling these populations with an overview of sampling, contacting, and analytical strategies. It highlights the inherent trade-offs of each approach and points to the factors to consider when determining which strategy is best suited to particular research questions. It then details polling efforts designed to capture attitudes and behaviors of American Muslims, Asian Americans, and nonelected political activists. The chapter concludes with a discussion of fruitful polling practices for conducting research on low-incidence U.S. populations, arguing that the approach to polling these populations must be equally informed by the unique characteristics of the target group and the analytical conclusions one seeks to draw.

Keywords: difficult-to-reach, low-incidence populations, American Muslims, Asian Americans, nonelected political activists, polling, religion, race, ethnicity, political behavior

Introduction

An increasingly diverse U.S. population presents survey researchers with new and multifaceted challenges. Those seeking to map American attitudes and behaviors with more precision and gradation can expect, for example, myriad difficulties attendant on surveying groups that constitute a relatively small portion of the populace. Such low-incidence populations can be characterized by recency of immigration, foreign-language dominance, racial and ethnic minority status, and geographic idiosyncrasies (i.e., whether the population of interest is relatively dispersed or concentrated in a given location). Thus, many of the characteristics related to higher unit and item nonresponse in polls are often more prevalent among these groups. Difficult as it is to identify and survey low-incidence populations, however, the descriptive and inferential findings gleaned from these efforts add valuable nuances to general population trends, allow for informative intra- and intergroup comparisons, and elaborate subgroups of particular political or theoretical importance.

This chapter outlines strategies for reaching beyond the "low-hanging fruit" of populations that are relatively easy to identify and survey. We argue for creative and targeted strategies rather than a one-size-fits-all approach to capturing information on low-incidence populations, beginning with consideration of the characteristics that make populations difficult to sample, interview, and analyze. To illuminate our approach, we utilize three cases of low-incidence populations in the United States characterized by religion, race and ethnicity, and political behavior. We begin by conceptualizing low-incidence populations and highlighting the existing empirical literature on these populations. We then turn our attention to framing the challenges of polling low-incidence populations, with an overview of sampling, contacting, and analytical strategies. In this section we highlight the inherent trade-offs of each approach and point to the factors that have to be considered when determining which strategy is best suited to particular research questions. Next we detail polling efforts designed to capture attitudes and behaviors of three low-incidence populations in the United States: (1) American Muslims, (2) Asian Americans, and (3) nonelected political activists. We conclude with a discussion of fruitful polling practices for conducting research on low-incidence populations in the United States. Ultimately, we argue that the approach to polling these populations must be equally informed by the unique characteristics of the target group and the analytical conclusions one seeks to draw.

Low-Incidence Populations

Often referred to in the polling literature as "rare" or "special," low-incidence populations can be thought of as a subset of difficult-to-reach populations. By *low incidence* we mean a group of individuals who share a common characteristic and make up a relatively small proportion of the broader population. Although difficult-to-reach populations may also have low incidence rates, these two traits are not necessarily intertwined. For example, corporate CEOs constitute a low-incidence population that is often difficult to reach. Alternatively, young adults between eighteen and twenty-nine years of age form a large segment of the population, but they can be quite difficult to reach, and when contacted are less likely to cooperate (De Leeuw et al. 2007; Curtin, Presser, and Singer 2005). Young adults are less likely to live in homes with landlines, reside in single-unit homes, remain in the same residence for an extended period of time, or be registered to vote, all of which makes it less likely that they will be adequately covered in a sampling frame that relies on data tied to these characteristics (Blumberg and Luke 2009).

Although empirical studies on low-incidence populations often focus on racial or ethnic minorities, this line of research also targets groups on the basis of, for example, specific types of illness, military service, or socioeconomic status. Studies based on samples of racial and ethnic low-incidence populations have been done about American Indians (Lavelle, Larsen, and Gundersen 2009), American Jews (Reed 1975; Lazerwitz 1978; Shor 2000), Afro-Caribbean blacks (Greer 2013), young black females (Ericksen 1976), non-English-speaking Chinese (Elliott et al. 2012), and Cambodian immigrants (Elliott et al. 2009). In addition, researchers have compiled national samples of minority populations, including the Pilot Asian American Political Survey (Lien, Conway, and Wong 2004), the National Asian American Survey (Wong et al. 2011), the National Survey of Black Americans (Jackson and Gurin 1987, 1999; Jackson and Neighbors 1997), the National Black Election Study (Jackson, Gurin, and Hatchett 1984; Tate 1997), the National Politics Study (Jackson et al. 2004), the Latino National Survey (Fraga et al. 2006), and the Latino National Political Survey (De la Garza et al. 1998). Multiple studies have also analyzed groups who suffer from a rare illness (Czaja et al. 1998; Sirken, Graubard, and McDaniel 1978), are at a greater risk of contracting an infectious disease (Watters and Biernacki 1989), and other at-risk populations (Biernacki and Waldorf 1981; O'Donnell et al. 1976; Rossi et al. 1987). Finally, research has investigated low-incidence populations on the basis of common military service (Rothbart, Fine, and Sudman 1982), and membership in an elite circle (Rossi and Crain 1968).

Each of the aforementioned studies focuses on low-incidence populations, but the particular characteristics of each population vary considerably. Some of the important differences include the extent to which the unifying rare characteristic is identifiable to the researcher, whether the group is geographically concentrated or dispersed, the level of preexisting research on the group, and finally the degree of uniformity among its members. The unique characteristics of one's population, coupled with the inferences one seeks to draw, ought to inform a study's approach to sampling, contacting, and analyzing

a target population. We identify three particular challenges to polling low-incidence populations and discuss each in turn below.

Sampling LowIncidence Populations

One of the central challenges of sampling low-incidence populations is identifying and locating individuals who share the characteristics in question. Low-incidence populations are often not characterized by either an observable trait or one that is recorded in official records. In our discussion of cases of religious and behaviorally defined groups below, we detail ways researchers have addressed the challenges of identifying and locating the low-incidence populations of American Muslims and political activists who do not serve in elective office. In these cases, a priori and robust measures of religious affiliation and political engagement are not available in official government data. Aside from certain historical spans when the U.S. racial taxonomy included categories for Jews and "Hindus," for example, religious affiliation has not been officially enumerated for the U.S. population (see Nobles 2000; Hattam 2007). Similarly, when interested in selecting a sample of politically active Americans, records of behavioral traits such as taking part in community-based political events are not readily available. In addition, and in partial contrast to religious affiliation (except for conversion and the sometimes fluid designation of a largely religious identity; see, e.g., Pew Research Center 2013), participatory behavior is a dynamic, moving target, changing with context, environment, and time.

Even when characteristics of low-incidence populations are observable and recorded, for example in racial enumeration records, identifying and locating groups that match a specific trait is complicated by geographic dispersion and heterogeneity within racial groups. Polling Asian Americans is a third case we examine in greater detail below. While grouped together racially in the official U.S. taxonomy, Asian Americans are a remarkably diverse set of people with a wide range of both immigrant trajectories and sending countries. Asian immigration to the U.S. is relatively recent, a function of pre-1965 federal immigration policies barring new entrants to the United States from Asian nations. As a result, Asian Americans today are a largely immigrant population, with nearly eight in ten adults born abroad. Immigration from Asia has not been dominated by a single nation, and Asian Americans come from a multitude of countries and speak a wide variety of languages. While family names may be distinctively "Asian" for East Asian nations such as China, Korea, and Japan, surnames for Asian Americans with colonial histories such as Filipino Americans and South Asian Americans are more difficult to distinguish from Americans whose racial and ethnic backgrounds are Latino or Arab American. The distinct surnames, coupled with the diversity of languages spoken, pose significant challenges to researchers who wish to poll this low-incidence population.

Recognizing these inherent difficulties in locating and identifying low-incidence populations, researchers have utilized three approaches to sampling individuals in these groups: (1) stratified designs, (2) list-based selection, and (3) density strategies. We provide an overview of each of the sampling strategies, weigh the associated trade-offs,

and highlight the factors to be considered when determining which approach is best suited for research. As we further illustrate in the case studies following this section, researchers must tailor their sampling strategies to the unique characteristics of their target populations and the type of inferences they seek to draw.

Stratified sampling (sometimes referred to as "purposive" or "multistage") is one probability technique available to surveyors of low-incidence groups. To construct a stratified sample, researcher first must identify the characteristics by which they wish to stratify along with the frequency at which these strata occur in the target populations, and subsequently sample individuals within these strata at random until the preset frequency is reached (Neyman 1934). Selection is often made based on demographic, socioeconomic, or geographic traits. This approach enables researchers to address the additional challenges associated with low incidence while still obtaining a representative sample of the target population. In addition, if a stratified sample is chosen at random, the researcher can better guard against potential selection threats. An additional benefit of this sampling strategy is that by setting a target sample size during the design phase, researchers can better ensure that their sample is large enough for the type of analysis they wish to conduct.

While this approach has a number of advantages, it has significant drawbacks and is not appropriate for many low-incidence populations. First and foremost, this sampling strategy can be costly, and the cost increases with the relative rarity of the low-incidence population. Researchers who lack sufficient financial resources are likely to find the costs of building an adequate size sample prohibitive. For example, the principal investigators of the 2008 National Asian American Survey attempted to utilize a telephone interviewing strategy through random digit dialing and yielded a very small number of successful contacts with Asian Americans from several thousand numbers. The relatively low incidence of the Asian American population (5%) and the high rate of English as a second language made this sampling strategy particularly inefficient.

Second, stratified sampling requires a "benchmark" survey, such as the U.S. Census, to ensure that the size and diversity of the low-incidence population is representative of the target population. As previously discussed, low-incidence populations are often classified by a shared characteristic—such as religion, immigration status, sexual preference, political activity, or illness—that is not accurately recorded in government data. Thus it may be difficult to ensure that one's stratified sample accurately represents the actual size and diversity of the target population.

Considering these drawbacks, stratified sampling may be better suited for intra- as opposed to intergroup analysis. If researchers seek only to ensure that their sample includes a subsample that is reflective of the group's low incidence within the larger population, stratified sampling may be an effective strategy. On the other hand, if instead they seek to better understand the low-incidence population itself, it may be best to employ an alternative sampling strategy that increases the sample's size and diversity. Since individuals who are contacted often vary considerably from those who are difficult

to contact—and often in theoretically significant ways—a small sample is unlikely to be representative of the target population. Researchers who lack the necessary financial resources, are interested in a particularly rare population, or are seeking to conduct intergroup analysis are likely to find stratified random sampling ill-suited for their research.

Another approach to studying low-incidence populations is the use of list sampling (Green and Gerber 2006; Sudman and Kalton 1986; Link et al. 2008; Gentry et al. 2010; Lazerwitz 1978; Brick, Williams, and Montaquila 2011). List sampling requires access to a record that provides enough information to identify and contact eligible members of the low-incidence population. In essence this catalog, which may be a combination of multiple lists, serves as a single sampling frame (Sirken 1972). Lists may be constructed to serve a particular public function, for instance, voter registration (Green and Gerber 2006) or delivery of mail via the U.S. Postal Service (Brick et al. 2011; Link et al. 2008; Gentry et al. 2010). Potential sampling frames of particular populations may also be constructed by civic organizations, unions, special interest groups, or commercial firms and may prove very useful for empirical work on low-incidence populations (Wong 2006; Greer 2013; Lazerwitz 1978; Shor 2000). Finally, if a list of the broader population includes information that enables one to identify eligible members of a low-incidence population, one may remove ineligible members and randomly sample individuals who remain. While this approach still requires interviewers to screen respondents on their initial contact, it nevertheless reduces the cost of screening and greatly increases contact rates.

Researchers often incorporate multiple lists to increase the coverage of their sampling frame (Kalton and Anderson 1986; Lohr and Rao 2006). One may also make use of samples from preexisting surveys (Reed 1975; Sudman and Kalton 1986) or may incorporate lists with a known high frequency of low-incidence population within a larger representative sample of the broader population (Kalton and Anderson 1986). List sampling can dramatically decrease the cost of sampling a low-incidence population, while at the same time enabling researchers to increase the size of their sample.

An additional advantage of list sampling is that if eligible members of a group are identified prior to contact, researchers may design the survey protocol in a way to maximize response rates. For instance, one may alter the description of the survey to cater to the interests or assuage the concerns of particular populations. Research has demonstrated that potential respondents are far more likely to respond if they have confidence in the sponsor of the survey, perceive the survey topic to be salient and relevant, or anticipate their participation in the survey will be rewarding and meaningful (Groves, Singer, and Corning 2000; Groves et al. 2006). Furthermore, one may match potential respondents to interviewers who share characteristics or language similar to those of potential respondents to further increases response rates. List samples provide researchers prior knowledge of the potential respondents, enabling them to design the survey and method of data collection in a way that can maximize the participation of the population they seek to analyze. The size of one's sample and the associated costs of

polling are not merely a function of the number of potential respondents one contacts, but also of the percentage of those who comply and complete the survey. List sampling may provide researchers with a more effective way to accomplish both.

While list sampling provides a cost-efficient and practical way to construct a sufficiently large sample of a low-incidence population, it presents a number of methodological trade-offs. One of the drawbacks to the list-sampling approach is that one cannot be sure that the frame completely covers the population, possibly introducing noncoverage bias. Second, there may be an increased risk that lists developed by civic organizations or special interest groups do not meet the requirement that respondents in a sample be independent of one another. This approach may result in over-coverage, meaning individuals have an unequal probability of being selected, making the construction of robust sample weights particularly challenging. This problem may be further compounded by the fact that multiple lists are often used to ensure broader coverage. Third, if one constructs a list from individuals who were sampled in preexisting surveys, in addition to facing the challenge of duplicates, each individual survey is likely to have distinct sampling protocols, again complicating the weighting methodology. Finally, due to issues pertaining to privacy or commercial concerns, organizations may not be willing to share lists or may only make them available at a considerable cost.

A final sampling method researchers may employ is density sampling, which is also referred to as "area" or "clustered" sampling (Waksberg 1978; Ericksen 1976; Hedges 1979; Waksberg, Judkins, and Massey 1997; Lien, Conway, and Wong 2004; Fraga et al. 2006; Blair and Czaja 1982). While low-incidence populations are by their very definition small in size, they may also be concentrated within a particular geographic area. This heavy concentration of a particular subgroup may be the result of segregation and isolation or of self-selection. And thus, while the targeted group may have a low incidence within the broader population, it may have a high incidence within a more narrowly restricted area. The density sampling approach seeks to take advantage of this concentration to increase contact rates and consequently lower the greater cost typically associated with surveying a low-incidence population.

Density sampling is a multistage process that is similar to stratified sampling. As previously discussed, stratified sampling begins by identifying characteristics that researchers believe are important indicators of the outcomes they seek to measure. The population is then divided into these strata and is sampled in a manner to reflect how the broader population is stratified along these lines (Neyman 1934). In density sampling, a researcher identifies particular geographic regions such as neighborhoods, census blocks, metropolitan statistical areas, states, or larger regions that have a higher concentration of a low-incidence population. Once these areas, or clusters, are identified —typically through the use of enumeration or previous reliable survey data—researchers may either randomly sample individuals from this primary sampling unit or further divide the area into smaller strata and randomly sample at a lower level of observation (Kalton and Anderson 1986; Hedges 1979; Waksberg 1978). If a low-incidence population is geographically concentrated within a defined area, density sampling can significantly

increase contact rates and consequently significantly reduce the associated costs of polling. Furthermore, if the vast majority of the target population is located within the area sampled, and the researchers have no a priori reason to suspect that those outside this defined area vary in theoretically significant ways, they may construct a sample that is both representative and of sufficient size to conduct analysis.

As do all forms of sampling, density sampling has its drawbacks, and researchers must determine if it is the most appropriate sampling approach for their research. First, the increased efficacy of density sampling, as well as the researchers' ability to construct weights that properly adjust for disproportionate sampling, are dependent on the researchers' ability to accurately estimate the prevalence of low-incidence populations at the appropriate level of observation (Kalton and Anderson 1986). This requirement may pose a significant hurdle because low-incidence populations tend to be underrepresented in larger surveys. This problem is not necessarily mitigated through a reliance on benchmark Census enumeration, because the unifying characteristic of the low-incidence population may not be recorded. Furthermore, given the Census's infrequent collection, it may not accurately represent the extant prevalence of a low-incidence population within a given geographic region.

An additional drawback to this sampling approach is that there is no assurance that members of the subpopulation who live within these densely populated clusters do not differ systematically from those who do not. Although low-incidence populations present the challenge of detection, they equally present the challenge of inference: To what extent can the population of the sample be generalized to the subgroup as a whole (Venette, Moon, and Hutchison 2002)? Consequently, the use of density sampling is not well suited for all low-incidence populations. It is a more effective means of polling if the population surveyors seek to research is geographically clustered in a densely populated area, and they do not have a priori reason to believe that members of the population inside the clustered areas vary significantly from those outside.

Surveying and Gaining Cooperation with Low-Incidence Populations

In addition to the challenges associated with sampling, researchers polling low-incidence populations face an additional hurdle. While it is a concern for any form of survey research, gaining cooperation with individuals in relatively rare groups comes with specific challenges. Precisely because of their relatively low frequency in an overall population, and as a result of the more complex sampling strategies undertaken, surveying and gaining cooperation with low-incidence populations must be approached with additional care. To this end, Groves and coauthors tested a theory they refer to as "leverage-saliency theory" (Groves, Singer, and Corning 2000) They hypothesize that during the screening phase of a survey, individuals will evaluate the perceived costs and benefits of participating in the survey, which will impact their propensity to respond. For instance, potential respondents may assign differing levels of benefits to participation due to perceived legitimacy of the sponsor, material incentives, and the perceived saliency

and importance of the topic, as well as potential costs (e.g., the length of the survey; cognitive or language-based burden; questions that are deemed to be embarrassing, invasive, or socially undesirable). Thus, one must tailor the design of the study to maximize the perceived benefits and minimize the perceived costs of the particular population one seeks to poll. This theory becomes particularly relevant for researchers who seek to include the surveying of a low-incidence population within a study of the larger population. If the perceived benefits and/or costs for members of the low-incidence population vary considerably from those for the larger population, researchers may face a significant hurdle to maximizing the level of cooperation of respondents. Our case studies of American Muslims and Asian Americans further illustrate this point.

One of the most common, and meaningful, ways in which the associated costs of a survey may differ between the general population and smaller subpopulations is language. A respondent's low level of English fluency may pose a significant hurdle to completion of the survey, and a researcher's failure to adequately account for this difficulty may significantly reduce the representation of particular low-incidence populations. A broad literature that has sought to identity the factors that contribute to increased levels of unit nonresponse has identified the potential barriers that may reduce a respondent's propensity to cooperate. In addition to identifying the role that perceived burden plays in reducing cooperation rates—such as the length of a survey, the level of knowledge that is required, or the risk of being forced to answer embarrassing questions—scholars have also identified language as an important barrier to cooperation (Groves, Presser, and Dipko 2004). If significant portions of a low-incidence population are systematically eliminated from a sample due to their inability to complete the survey in the language in which it is offered—such as recent immigrant groups for whom English is their second language—the resulting sample may not be representative of the broader population, and nonresponse bias may result. Since nativity, length of time in the United States, levels of education, and racial and ethnic identity are correlated with both response propensity and many political outcomes of interest—for example, public opinion, voting, voter registration, civic engagement—there is an increased risk of nonresponse bias.

Furthermore, if steps are not taken to alter the selection mechanism—in this case, a match between the language of the potential respondent and the survey instrument—then neither oversampling nor back-end statistical adjustments are likely to reduce the level of bias. For instance, if non-English-speaking Asian Americans vary systematically from English-speaking Asian Americans, even if one constructs a sample that is proportional to the size of the group in the broader population, the respondents within the subgroup may not be representative of the subgroup as a whole. Failure to correct for the selection mechanism will not only potentially bias population estimates, but also prevent accurate subgroup analysis. Furthermore, statistical adjustments on the back end will be difficult, because researchers will be unable to place greater weight on those respondents with a low propensity on the basis of language, because they are likely to have been fully eliminated from a sample.

Recognizing the significant challenges associated with a bilingual population in the United States, researchers have increasingly conducted surveys in languages other than English. However, while researchers have increased the number of surveys conducted in Spanish, the population is increasingly multilingual with an ever-growing number of languages being spoken. According to the U.S. Census, the population of adults who speak a language other than English at home increased from 13.8% in 1990 to 17.8% in 2000. If we extend our analysis to the entire population in the most recent Census (five years or older), 20% of Americans speak a language other than English at home, and of this population, 22.4% speak English either "not well" or "not at all" (2010 U.S. Census).

Asian Americans, who now make up the largest share of immigrants, account for 15% of those who speak a language other than English, but represent a higher percentage of those who speak English "not well" or "not at all." Focusing on four of the largest co-ethnic Asian American groups, 29.6% of Chinese, 28.4% of Koreans, 33.1% of Vietnamese, and 15.1% of Japanese Americans speak English "not well" or "not at all." In addition to problems associated with nonresponse, the inclusion of respondents who complete the survey in a language in which they are not fully proficient may increase measurement error that may similarly bias results. For these reasons, it is essential that effective protocols be established to ensure that both questionnaires and surveyors are reflective of the target population of the study. While translation is both a costly and an arduous process, it is likely to reduce total survey error by increasing both contact and cooperation rates and reducing the degree of measurement error. Strategies that have been implemented to increase response rates, for instance advance letters or prescreening phone messages, will be ineffective if they do not reflect the diverse languages of the target population.

In an effort to combat these challenges, surveys that focus on low-incidence populations, as well as larger surveys seeking a nationally representative survey, typically have translators available in call centers. However, fielding a bilingual or multilingual poll can be both challenging and costly. Matching potential respondents with the correct foreign-language interviewer and conducting the survey with a translated instrument is a more costly and difficult process in multiple languages than it is when the survey is done only in English and Spanish. If the languages spoken by the translators do not represent the diversity of language spoken by the survey population, it may not eliminate the potential for nonresponse bias. Furthermore, if screening calls are still conducted in English, there is an increased risk that the potential respondent may terminate the interview before the interviewer is able to match the respondent with an interviewer who can conduct the survey in the appropriate language. While the percentage of respondents who are lost during the transition to the translator, and the associated bias that transition may induce, are unknown, evidence in an analogous area suggests it may pose a problem. More specifically, a significant number of interviews are terminated during the time that interviews are transferred from a live interviewer to the interview and data collection system (Tourangeau, Groves, and Redline 2010).

Reaching Beyond Low-Hanging Fruit: Surveying Low-Incidence Populations

While maximizing response rates with an unbiased sample of respondents is ideal, it is not always possible within the constraints of budget and time. When conducting a baseline study of a set of behaviors or attitudes of the population in question (such as the National Asian American Survey, discussed below) for which earlier systematic data are not available, incurring the time and expense of maximum coverage of a low-incidence population is imperative. Subsequent studies and other efforts, however, can utilize baseline studies conducted with maximum-coverage designs to provide some measure of comparison when full coverage of a low-incidence population is not feasible. Nevertheless, comparisons to baseline data should be conducted with care given the dynamism of low-incidence populations such as Asian Americans.

Drawing Inferences from Data Collected from Low-Incidence Populations

After clearing the multiple hurdles associated with sampling and surveying low-incidence populations, researchers face additional challenges to analyzing the data. On the back end of a survey, a survey methodologist may take additional steps to adjust for the potential of nonresponse bias. The distinct statistical strategies are informed by the logic of response propensity. One commonly used statistical strategy is post-stratification. Akin to stratified sampling, in this strategy researchers attempt to make the sample more representative of the target population. Researchers identify characteristics they believe are likely to correlate with the outcome measurements of interest—typically demographic, socioeconomic, or geographic in nature—and make statistical adjustments so that the sample matches the characteristics of a "bench-mark survey," such as the U.S. Census, or those of the sampling frame (Brehm 1993; Olson 2006). These adjustments are typically accomplished by increasing the weight of responses from individuals whose characteristics match those of a subgroup population that responded at lower rates than their population proportion.

Another strategy employed by survey researchers is the use of propensity score weights (Groves et al. 2006; Peytchev, Peytcheva, and Groves 2010; Groves 2006; Olson 2006; Tourangeau, Groves, and Redline 2010). This back-end adjustment technique is analogous to post-stratification. However, rather than matching respondents to the general population along theoretically important strata, one is attempting to match respondents in the sample to nonrespondents based on their shared low propensity to respond. In employing propensity scores one is attempting to limit potential nonresponse bias by modeling the response process. If researchers can identify the particular predispositions that increase or decrease an individual's propensity, they can assign every respondent within their sample a propensity score ranging from 0 to 1. If the propensity scores are accurate, the survey researchers can place greater weight on respondents who have a relatively low propensity to respond. By modeling response propensity, researchers can adjust weights to account for unequal selection rates, as well as unequal response rates that may bias their results.

Nevertheless, the effectiveness of propensity scores, like that of post-stratification, depends on survey researchers' knowledge of which characteristics best explain response propensity. Thus, postsurvey adjustments are based on assumptions about the relationship between response propensity and the survey estimate in question. Some survey methodologists argue that nonresponse is a complex and interactive process that includes a multitude of individual- and survey-level characteristics, which are likely to vary across distinct survey items, and thus caution against overreliance on back-end statistical adjustments (Brehm 1993; Brick 2011; Olson 2006; Groves et al. 2006).

A similar technique involves the identification of "high-effort cases" (Curtin, Presser, and Singer 2005; Keeter et al. 2000; Keeter et al. 2006; Teitler, Reichman, and Sprachman 2003; Stinchcombe, Jones, and Sheatsley 1981; Sakshaug, Yan, and Tourangeau 2010). The theory is that if researchers identify respondents in the sample who required multiple attempts to contact and/or were initially unwilling to cooperate, those people can serve as proxies for those who did not respond. If these high-effort cases share no unifying characteristic, then nonresponse may be random, thereby minimizing the threat of bias from this source. On the other hand, if they do share a unifying characteristic, researchers can account for it.

However, a body of research suggests that high-effort cases do not resemble nonresponders along key demographic lines and thus may not be effective in correcting for nonresponse bias (Brick 2011; Lin and Schaeffer 1995; Teitler, Reichman, and Sprachman 2003; Olson 2006; Groves and Couper 1998). These concerns have led survey researchers to suggest that bias resulting from nonresponse may be more successfully dealt with at the design phase and during the process of data collection (Olson 2006; Brehm 1993; Brick 2011; Sakshaug, Yan, and Tourangeau 2010; De Leeuw et al. 2007). Low-incidence populations, due to a combination of their low prevalence and the many challenges of researching them, are typically underrepresented in the empirical literature. As a result, we often lack the empirical data and associated theory required to make accurate post-data collection adjustments. It is for this reason that we argue that such collection adjustments are unlikely to overcome deficiencies in the survey design and the protocol employed during the phase of data collection.

Taken together, these three challenges facing researchers interested in low-incidence populations—drawing a sample, gaining cooperation, and analyzing the data—present a high bar indeed for successful polling. At the same time, the benefits of gathering systematic and high-quality data for low-incidence populations is well worth the effort. In the next section we detail attempts by researchers to survey three specific types of low-incidence populations: the religious group of Muslim Americans, the racial and ethnic group of Asian Americans, and a behaviorally distinguished group of political activists who do not hold official elective office. Our discussion of these three case studies is meant to serve as a reference for those who have a substantive interest in these three groups, as well as to elucidate the various factors one must consider when designing survey research on low-incidence populations. The combination of strategies undertaken to sample, contact, and poll members of low-incidence groups must equally reflect the

unique characteristics of the groups, the resources at the researchers' disposal, and the type of inferences they seek to draw. Each low-incidence population presents unique challenges and opportunities, and researchers must tailor their survey research accordingly.

Successful Studies of Low-Incidence Populations

Although there are standard difficulties that come with designing and implementing a survey targeting low-incidence groups, each population will naturally pose its own unique challenges. Researchers therefore often require a priori knowledge of the population of interest to anticipate any sampling or analytical hurdles they will need to clear, or at least sidestep. Yet, this universal prescription varies in its ease of applicability. Surveyors of low-incidence populations must therefore implement a tailored approach that accounts for the trade-offs accompanying key strategic decisions (Dillman, Smyth, and Christian 2014). The following assessment of efforts to sample low-incidence groups begins with the relatively rare population of American Muslims.

American Muslims

Researchers compiling original data on American Muslim attitudes and behaviors face difficult decisions when formulating a sampling strategy. Like all surveyors, they aim to minimize survey error while working within time and financial constraints. However, the calculus undergirding these core considerations can shift dramatically when targeting low-incidence groups. In seeking a national probability sample, for instance, the total number of contacts needed to secure an adequate number of respondents can quickly grow as incidence rate decreases, putting a substantial strain on resources. Yet as principal investigators move away from an ideal research design to relieve cost burdens, the already elevated risks of sampling error and myriad biases can become too difficult to manage or fully account for. Striking a balance between these competing interests is made all the more challenging in light of complications particular to American Muslims.

Beyond a very low incidence rate,[1] researchers face additional legal, demographic, and social challenges in surveying American Muslims. The chief obstacle to sampling this community stems from a lack of any official data, as both the Census Bureau and Immigration and Naturalization Service are legally barred from compiling statistics on religious affiliation. This limitation naturally puts those researching any religious group at a disadvantage compared to those surveying ethnic or racial communities, for whom population data are readily available. Yet American Muslims' linguistic and ethnic diversity makes sampling them even more complex than, say, American Jews or Mormons.

Indeed, the most reliable estimates peg Muslims as the most ethnically and racially diverse religious minority in America (Pew Research Center 2007, 2011; Gallup 2009), with a linguistic variety perhaps rivaled only by Asian Americans (Junn et al. 2011).

Moreover, while 80% of American Jews reside in five states (Pew Research Center 2013), and over one-half of Mormons live in the Mountain West (Pew Research Center 2012a), Muslim congregations are found across all major geographic regions, with Islam constituting the second largest religion in as many as twenty states (Grammich et al. 2012). There are certainly areas that are home to comparatively large Muslim populations (the New York metro region; Dearborn, Michigan; and Southern California, to name a few); on the whole, however, this community is not nearly as geographically concentrated as other religious minority groups. Such multidimensional heterogeneity and wide distribution means that even well-resourced surveys of American Muslims will face acute design trade-offs. Some of these bargains, whether made consciously or accepted ex post, are highlighted below.

The most comprehensive studies on American Muslims to date are those carried out by Pew (2007, 2011) and Gallup (2009, 2011)—though the two organizations took notably distinct methodological tacks. More specifically, Pew's (2007) original study used two sampling frames—a random digit dial (RDD) sample gleaned from geographic and list strata, which was coupled with a recontact frame drawn from Muslim respondents to previous national surveys—to produce the first national probability sample of American Muslims. Its representativeness, on the one hand, is bolstered by the interviews being conducted in four different languages (English, Arabic, Urdu, and Farsi), but on the other hand, is somewhat undermined by the RDD frame not including a cell phone component despite estimates at the time that 13.5% of U.S. households with telephones were wireless only (Blumberg and Luke 2009).[2]

The focus of the Pew (2007) study was likewise a double-edged sword. More specifically, concentrating the survey on Muslims in America allowed the researchers to field a questionnaire partially tailored to this community. That is, in addition to obtaining data on standard demographics, civic participation, political affiliation, and so forth, the survey also asked about respondents' experience with discrimination following the September 11, 2011, attacks; belief in official accounts of this terrorist plot; matters of religiosity particular to Islam; and other such issues that are especially informative for this population. Yet this greater flexibility in questionnaire design is mitigated by the heightened sensitivity that respondents may have had to the questions and the consequent threat such a posture poses to measurement validity. In turn, Pew researchers took several steps to preemptively alleviate nonresponse and social-desirability bias. These tactics included extensive pretesting of the questionnaire and an uncommon disclosure of the study's focus early in each interview. This latter decision, however, potentially traded one form of bias for another, further emphasizing the calibrations that belie a one-size-fits-all model for surveying low-incidence groups.

Gallup's (2009, 2011) survey methodology differed from Pew's (2007, 2011) in several key respects. Rather than targeting American Muslims, Gallup selected self-identified Muslim respondents from the Gallup Poll Daily survey, which tracks a general sample of American households. That is, rather than an oversample, Gallup aggregated the responses of 946 Muslims drawn from a database of nearly 320,000 adults across the United States. One of the more significant analytical advantages of this strategy is the ability to organically compare the opinions of American Muslims to other segments of the broader public, given the identical questionnaires and prompts used across all respondents. In addition, the extensive coverage of this technique is reinforced through a dual-mode RDD frame that included both landline and cellular numbers. While this methodology may have produced the "first-ever nationwide representative random sample of Muslim Americans" (Gallup 2009, 16), there were nonetheless several limitations inherent in the design.

Given that the Gallup Poll Daily targets a general population, the survey in turn had to be generally applicable. As such, many questions specific or more relevant to an American Muslim sample—arguably the very questions that scholars and policymakers most often seek answers to—were not included in the questionnaire. This broad scope also meant that there was no incentive to offer interviews in languages other than English and Spanish, which is especially problematic given that Arabic, Urdu, and Farsi interviews constituted 17% of Pew's (2007) sample. Certainly, however, a survey that does not specifically focus on American Muslim opinions may increase the response rate among this wary population. Yet a high response rate in itself does not guard against nonresponse bias (Groves and Peytcheva 2008), and Gallup's (2009) report, given the expansive sample it is drawn from, does not provide the same analysis of survey nonresponse as Pew's (2007). Ultimately, while a random national sample of American Muslims may be a significant achievement, it is no panacea for addressing the difficulties of low-incidence sampling.

If well-funded organizations are nonetheless forced to make certain concessions from a theoretically ideal sampling design, then naturally academic researchers and smaller institutes working within significantly tighter resource constraints will fare no better. Indeed, due to the numerous challenges discussed above, the vast majority of studies featuring original survey data on American Muslims, whether academic (Patterson, Gasim, and Choi 2011; Muedini 2009; Sharif et al. 2011) or institutional (Council on American Islamic Relations 2006, 2012; Muslim Public Affairs Council 2005), are effectively drawn from convenience samples or, at best, are representative of a local population (Bagby 2004). A number of projects with far more modest budgets than either Pew or Gallup have, however, sought (with varying degrees of success) to obtain a nationally representative sample of American Muslims.

Zogby International (2001, 2004) compiled arguably the most extensive accounts of the American Muslim population prior to the Pew (2007) study. The methodology employed across both of Zogby's surveys targeted Muslim respondents by randomly selecting 300 Islamic centers and drawing from a listing of Muslim surnames in the surrounding area to

populate an RDD frame. Additional in-person interviews sought out African American Muslims in New York, New York, Washington, D.C., Atlanta, Georgia, and Detroit, Michigan, to compensate for this subgroup's likely underrepresentation in the RDD sample. The reports offer no details, however, on how the sampling for the in-person interviews was undertaken, nor do they provide a rationale for not including Philadelphia, Pennsylvania, among the cities visited, given its high concentration of African American Muslims. Thus, despite conducting more interviews (n ≈ 1,800) than either the Pew or Gallup polls discussed above, the lack of methodological clarity (in addition, there is no mention of whether the interviews were carried out in languages other than English) makes it difficult to take the reports' claims of representativeness at face value (Zogby 2001, 2004).

Another project that cast a wide net was the Muslim American Public Opinion Survey (MAPOS) (2010). For this study, researchers recruited local Muslim enumerators in twenty-two locations across the United States (eleven cities; two locations in each city) to hand out two-page "exit-poll style" surveys following weekly Friday services and semiannual Eid celebrations. The benefits of this strategy include the ability to employ multilingual, self-administered surveys, which cut down on nonresponse and social desirability bias. Reliance on a single sampling mode and the exclusion of cities with large Muslim centers are among the study's limitations; but despite these drawbacks, the authors' contention that their sample is nonetheless representative is not without merit. The validity of this claim stems from the study's central questions, which gauge how religiosity influences political participation within this minority population. As such, the more informative opinions for the authors' purposes are those of the more religiously inclined that this sampling strategy targets. This method of using a study's motivating questions as a reference for calibrating resource allocation, as the concluding section of this chapter discusses, constitutes another rare universal prescription for pollsters targeting low-incidence populations.

Asian Americans

In comparison to Muslim Americans—who can be of any racial or ethnic background—the official American racial taxonomy classifies and enumerates Asian American races as a function of national origin. In 1960, and prior to the reopening of the United States to immigration from Asia in 1965 with the Immigration and Nationality Act, the size of the Asian American population was fewer than one million and represented only a fraction of the entire U.S. population. Subsequent years of increasing immigration to the United States from Asia have driven the size of the Asian American population to more than 5% of all Americans. Until the 1990s, Asian Americans were heavily of East Asian national origins, particularly Chinese and Japanese. But in subsequent decades, immigration from Asia to the United States has expanded to include large numbers of new Chinese, South Asian Indians, Filipinos, Vietnamese, Korean, and Southeast Asians. Because the vast majority of immigrants are recent, they speak native languages including Mandarin, Cantonese, Hindi, Bengali, Tagalog, Vietnamese, Korean, and Thai, among others. This degree of variation makes matching targeted Asian Americans to a language of interview a complex process requiring expertise in Asian culture. Similarly, and because the Asian American population is more heavily concentrated in some states, their geographic settlement patterns create challenges for researchers attempting to survey this low-incidence population.

Two recent national studies of Asian American opinion and political behavior provide excellent guidance for researchers interested in polling Asian Americans. The National Asian American Survey (NAAS) of 2008 was conducted over the telephone with 5,159 respondents (Wong et al. 2011). The largest national origin groups—Chinese, South Asian, Filipino, Vietnamese, Korean, and Japanese—were interviewed in the language of their choice. Selection of the sample was accomplished by a combination of techniques including utilizing lists, RDD, and stratified design, as well as density sampling. Because the researchers were interested in drawing inferences about Asian Americans in the United States overall as well as national origin groups and also particular states that were considered political "battleground" states in 2008, the principal investigators began the process of conducting the sample at the county level by selecting locations classified as high and low immigration as well as new and old immigrant destinations.

Identifying Asian Americans was accomplished primarily through a national list based primarily on surnames, but the NAAS researchers supplemented the known universes with both RDD (to test the frequency of incidence and resulting cost of attempting to screen from a random sample to capture Asian Americans) and lists constructed specifically to capture Filipino Americans. Many Filipino surnames have origins in the Spanish colonial experience and therefore are often conflated with Latino and Hispanic ethnic origin. In addition, NAAS researchers utilized a proprietary name-matching database to predict the ethnic origin and therefore the language preference of potential subjects. As discussed previously, if the initial contact with a respondent is not made in that person's language, the likelihood of completing the interview is substantially

reduced. Therefore, the selected sample was coded for likely national origin and assigned separately to bilingual interviewers who spoke Mandarin and English, or Tagalog and English, for example. All interviewers were bilingual in English and an Asian language.

The resulting data collection from more than five thousand Asian Americans in the NAAS represented the first national sample of Asian American political attitudes and behavior conducted in the United States. Benchmarks for partisan affiliation, voting turnout, and vote choice are established in these data not only for Asian Americans nationally, but also for particular national origin groups. Subsequent studies of Asian Americans have followed this multi-approach research design pioneered by the NAAS.

A second important national survey of Asian Americans was conducted by the Pew Research Center in 2012b. The Pew Asian-American Survey completed interviews with 3,511 respondents identified as Asian Americans. Similar to the NAAS, the Pew study conducted telephone interviews with respondents with bilingual interviewers and asked a series of questions about political activity, opinion, attitudes about politics, and socio-demographics. Many of the results are similar to the findings from the 2008 survey, even though the emphasis of the Pew study was on an overall portrait of Asian American social attitudes rather than on political activism and public opinion as it was in the NAAS. The Pew study utilized samples from its previous national studies to locate potential Asian American respondents in addition to existing lists and targeted density sampling. This study is another example of a creative use of multiple strategies of identifying, locating, sampling, analyzing, and surveying a low-incidence population. It is important to note, however, that this sampling strategy is conditioned on successful contact and willingness to cooperate in the previous study, rendering the eventual sample of respondents a conglomerate panel of earlier sampling frames. As a result, the survey sample in the Pew study is comprised of a subset of the larger population interviewed successfully once before, and this underlying bias should be taken into account when conducting analysis.

Political Activists

A final case of a low-incidence population in the United States is defined by a high degree of activism in politics. Despite a range of levels of government in which to participate—from the local community council and school board, to city hall, to the state house, to the federal government—most Americans take part in politics by engaging in one political act every four years, and that is voting in a presidential election. While voter turnout increased in the high-interest election of 2008, less than 60% of the eligible population of voters cast a ballot in the 2012 U.S. presidential election. Other forms of activity in the electoral arena, including working for campaigns or attending rallies, are even lower frequency, though larger proportions of Americans report having made a contribution to a candidate or political cause. Despite the old adage that "all politics is local," many fewer Americans vote in municipal or statewide elections than in federal elections, and a relatively small proportion report taking part in activities at the local level. Even the self-

interested act of contacting an elected official for help in solving a problem has low incidence in the U.S. population.

Thus political activists are rare in the United States, and finding those who are engaged in politics without being elected officials is a task that can only be accomplished by embracing the dynamism of politics and the fluidity of political behavior. While there are multiple examples of researchers surveying political activists—from political party convention delegates to citizens who attend local town hall meetings—the most substantial and comprehensive effort to assess the motivations and attitudes of ordinary Americans who are actively engaged in politics is a study conducted by the political scientists Sidney Verba, Kay Schlozman, Henry Brady, and Norman Nie (Verba, Schlozman and Brady 1995; Nie, Junn, and Stehlik-Barry 1996). The Citizen Participation Study began with a large RDD "screener" of 15,053 respondents. The screener sample was nationally representative, and interviews were conducted by telephone. Based on analysis of the screener data, which asked a range of questions on political and civic engagement in voting; electoral politics; community-based activity; contacting officials, local boards, and councils; and protest activity, a smaller set of respondents was selected for reinterview in a subsequent study.

The follow-up survey was conducted with 2,517 respondents in person and asked respondents about the specific political activities that they engaged in and the reasons they took part, along with a wide range of attitudinal and demographic questions. Oversamples of activists specializing in specific types of activities, such as protesters or campaign workers, were drawn in addition to a sample of ordinary Americans who were active in politics in multiple ways. This stratified design allowed researchers to analyze a randomly selected sample of different types of activists as well as view the U.S. population as a whole by employing post-stratification weights in analysis. The creative design employed by Verba and colleagues in their study of this particular low-incidence population has continued to pay dividends for researchers interested in understanding the dynamics of political participation. While difficult, expensive, and time-consuming, the Citizen Participation Study has yielded important insights into why and why not Americans take part in the politics of their nation.

Discussion and Conclusion

Surveys provide an indispensable tool to describe and explain characteristics of a broader public. While it is typically not feasible to contact every individual one seeks to describe, advances in probability theory, increased computing power, and continual advancements in modes of communication have enabled survey researchers to generate valid and reliable measures of a population from the observations of a sample. In the annual presidential address to AAPOR, Cliff Zukin made the claim, "our methodology is built on the notion—and science—of sampling. That is, we select and interview a small group of people to represent an underlying population" (2006, 428). Sidney Verba goes even further in extolling the virtues of sampling, comparing the empirical approach with the normative goals we seek to measure. Verba contends, "Surveys produce just what democracy is supposed to produce—equal representation of all citizens. The sample survey is rigorously egalitarian; it is designed so that each citizen has an equal chance to participate and an equal voice when participating" (1995, 3).

However, the validity and reliability of one's inferences depend on the extent to which the sample one observes is, in fact, representative of the broader population one seeks to describe. Gathering systematic and high-quality data from low-incidence populations presents substantial if not insurmountable challenges to pollsters. Low-incidence groups are characterized not only by their relative rarity in the population, but also by the accompanying fact that these individuals are both more difficult to identify and contact. Yet despite the difficulties low-incidence populations present, it is essential for the surveyor to develop effective strategies to meet these challenges.

Similar to difficult-to-reach populations more generally, if there is a theoretical reason to believe that subgroups differ significantly from the broader population along outcomes of interest, then their omission may bias a study's estimates. The extent of bias depends on the relative size of the low-incidence population to the total population, as well as the extent to which the low-incidence population differs from the total population on the measures of interest. Thus, one might argue that due to their inherent small size, low-incidence groups are unlikely to bias estimates of the general population. However, as smaller segments of the public grow in size (e.g., immigrant groups; cell-phone users), the omission of these increasingly prevalent individuals raises the risk of bias. This challenge is further complicated by the fact that low-incidence populations tend to be underrepresented in most survey samples, and thus we often lack the empirical evidence to assess the extent to which they differ. Furthermore, in order to conduct meaningful subgroup analysis, researchers must have adequate sample sizes. In addition to learning more about the particular subgroup, intergroup comparison will enable us to test the generalizability of theories. How do groups differ? What explains these differences? Answers to these questions will enable us to develop conditional theories that more accurately depict our diverse population.

As this chapter has highlighted time and again, however, researchers face tough decisions when it comes to resource allocation. Limited time and funding necessarily entail compromises. Although the difficulties particular to surveying low-incidence populations by and large belie one-size-fits-all prescriptions, two broad considerations should factor into any sampling design. First, mixed-mode data collection techniques offer researchers a way to potentially reduce costs and/or reduce nonresponse (Dillman, Smyth, and Christian 2014). For example, maximizing the number of responses attained through a relatively cheap mode (say, a mail-in or some other self-administered survey) before moving on to a more expensive mode (usually one requiring trained enumerators) is a generally optimal practice that is particularly beneficial to pollsters surveying rare groups, where the costs associated with coverage can be particularly burdensome. Moreover, when collecting data on populations that include a large portion of members whose first language is not English, the coverage advantages of face-to-face or telephone surveys can be easily outweighed by the nonresponse attendant on English-only interviews. In this scenario, adding a self-administered mode with several translations of the questionnaire could be far more cost effective than training multilingual interviewers. Indeed, a mixed-mode strategy is all the more advantageous given that cultural and linguistic minority groups may be more suspicious of interviewers, particularly if they are not members of their community (Harkness et al. 2014), yet if given the opportunity to share their opinions in the language of their choosing, such minorities may be willing to participate just as often as the majority population (Feskens et al. 2007).

Second, the inevitable trade-offs should be optimized with regard to the study's core research questions. This tailored approach is, again, applicable to polling generally, although its advantages are more acute in the case of low-incidence populations. For example, researchers with the MAPOS (2010) study aimed to elaborate the role of mosque attendance in social and civic participation; thus they opted for a sampling strategy—polling congregants as they left communal prayers—that likely skewed any resultant bias in representation toward the particular subgroup of interest within their target population. As obtaining a national probability sample of American Muslims would have been prohibitively expensive, the coordinators for this project maximized their resources by focusing on the sources of error they could best guard against: first by providing self-administered questionnaires in multiple languages, handed out by Muslim enumerators—which cut down on nonresponse and social desirability bias—and second, and more to the point, by tailoring their coverage priorities to the research questions motivating the study.

Beyond these front-end calibrations in research design, there are two meaningful actions researchers of low-incidence populations can take to improve the back-end data analysis. First, a meta-analysis of national probability surveys featuring meaningful numbers of the group of interest can provide researchers with more reliable demographic baselines. These more valid metrics would help researchers design more effective sampling strategies and apply more accurate post-stratification weighting. This approach has

successfully been utilized by pollsters studying American Jews (Tighe et al. 2010) and can effectively curb overreliance on the demographic picture painted by a single survey.[3]

Second, researchers and surveyors should be more forthcoming with detailed appraisals of their methodology. This goes beyond a general ethos of transparency to acknowledge that, as has been shown, nuanced decisions can have quite meaningful effects. One concrete measure that this can translate into is asking in-person enumerators, such as those of the MAPOS (2010) survey, to keep track of—and report—descriptive data on those individuals who opt not to take the survey, in order to paint a fuller picture of nonresponse error. These reports should include objective traits—such as age range, sex, location of contact, an so forth—but even more subjective inferences regarding the reasons behind their refusal to participate could prove useful (for example, whether it was because they were too busy or merely felt suspicious of the enumerators' motives). Noting those respondents who required extra cajoling to participate would similarly be of benefit to this end.

Since it is typically impractical (often close to impossible) to observe every unit of interest, scholars carefully attempt to construct a sample that is generally representative of the target group. In turn, the validity and reliability of one's inferences depend on the extent to which the resultant sample meets this criterion. This chapter discussed the heightened obstacles that researchers of low-incidence populations face in this regard and the possible paths they may take in meeting these added challenges. While there is no methodological silver bullet, each conscientious contribution helps to fill gaps and advance a more holistic understanding of not just rare populations, but society at large.

References

Bagby, I. 2004. *A Portrait of Detroit Mosques: Muslim Views on Policy, Politics, and Religion*. Detroit, MI: Institute of Social Policy and Understanding.

Biernacki, P., and D. Waldorf. 1981. "Snowball Sampling: Problems and Techniques in Chain-referral Sampling." *Social Methods and Research* 10: 141–163.

Blair, J., and R. Czaja. 1982. "Locating a Special Population Using Random Digit Dialing." *Public Opinion Quarterly* 46 (4): 585–590.

Blumberg, S. J., and J. V. Luke. 2009. "Reevaluating the Need for Concern Regarding Noncoverage Bias in Landline Surveys." *American Journal of Public Health* 99 (10): 1806–1810.

Brehm, J. 1993. *The Phantom Respondents*. Ann Arbor: University of Michigan Press.

Brick, J. M. 2011. "The Future of Survey Sampling." *Public Opinion Quarterly* 75: 872–878.

Brick, J. M., D. Williams, and J. M. Montaquila. 2011. "Address-Based Sampling for Subpopulation Surveys." *Public Opinion Quarterly* 75 (3): 409–428.

Council on American-Islamic Relations. 2006. *American Muslim Voters: A Demographic Profile and Survey of Attitudes*. Washington, DC: Council on American-Islamic Relations.

Council on American-Islamic Relations. 2012. *American Muslim Voters and the 2012 Elections: A Demographic Profile and Survey of Attitudes*. Washington, DC: Council on American-Islamic Relations.

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

Czaja, A. J., G. L. Davis, J. Ludwig, and H. F. Taswell. 1998. "Complete Resolution of Inflammatory Activity Following Corticosteroid Treatment of HBsAg-Negative Chronic Active Hepatitis." *Hepatology* 4 (4): 622–627.

Dana, K., M. A. Barreto, and K. A. R. Oskooii. 2011. "Mosques as American Institutions: Mosque Attendance, Religiosity and Integration into the Political System among American Muslims." *Religions* 2 (4): 504–524.

De la Garza, R., A. Falcon, F. C. Garcia, and J. A. Garcia. 1998. *Latino National Political Survey, 1989–1990*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

De Leeuw, E., M. Callegaro, J. Hox, E. Korendijk, and G. Lensvelt-Mulders. 2007. "The Influence of Advance Letters on Response in Telephone Surveys a Meta-Analysis." *Public Opinion Quarterly* 71 (3): 413–443.

Dillman, D. A., J. D. Smyth, and L. M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. Hoboken, NJ: Wiley.

Djupe, P. A., and B. R. Calfano. 2012. "American Muslim Investment in Civil Society Political Discussion, Disagreement, and Tolerance." *Political Research Quarterly* 65 (3): 516–528.

Elliott, M. N., W. S. Edwards, D. J. Klein, and A. Heller. 2012. "Differences by Survey Language and Mode among Chinese Respondents to a CAHPS Health Plan Survey." *Public Opinion Quarterly* 76 (2): 238–264.

Elliott, M. N., D. McCaffrey, J. Perlman, G. N. Marshall, and K. Hambarsoomians. 2009. "Use of Expert Ratings as Sampling Strata for a More Cost-Effective Probability Sample of a Rare Population." *Public Opinion Quarterly* 73 (1): 56–73.

Ericksen, E. P. 1976. "Sampling a Rare Population: A Case Study." *Journal of American Statistical Association* 71: 816–822.

Feskens, R., J. Hox, G. Lensvelt-Mulders, and H. Schmeets. 2007. "Nonresponse Among Ethnic Minorities: A Multivariate Analysis." *Journal of Official Statistics* 23 (3): 387–408.

Fraga, L. R., J. A. Garcia, R. Hero, M. Jones-Correa, V. Martinez-Ebers, and G. M. Segura. 2006. *Latino National Survey (LNS), 2006.* ICPSR 20862. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-06-05. **http://doi.org/10.3886/ICPSR20862.v6**.

Gallup. 2009. *Muslim Americans: A National Portrait.* Washington, DC: Gallup.

Gallup. 2011. *Muslim Americans: Faith, Freedom, and the Future*. Abu Dhabi: Abu Dhabi Gallup Center.

Gentry, R., M. Cantave, N. Wasikowski, and Y. Pens. 2010. "To Mail or to Call: How to Reach the Hard-to-Reach." Paper presented at the 65th Annual Meeting of the American Association for Public Opinion Research, Chicago.

Grammich, C., Hadaway, K., Houseal, R., Jones, D. E., Krindatch, A., Stanley, R., and Taylor, R. H. 2012. 2012 U.S. Religion Census: Religious Congregations and Membership Study. Association of Statisticians of American Religious Bodies. Nazarene Publishing House. **www.nph.com/nphweb/html/nph/itempage.jsp?itemid=9780615623443**.

Green, D. P., and A. S. Gerber. 2006. "Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts?" *Public Opinion Quarterly* 70 (2): 197–223.

Greer, C. 2013. *Black Ethnics: Race, Immigration, and the Pursuit of the American Dream*. New York: Oxford University Press.

Groves, R. M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys: What Do We Know about the Linkage between Nonresponse Rates and Nonresponse Bias?" *Public Opinion Quarterly* 70 (5): 646–675.

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

Groves, R. M., M. P. Couper, S. Presser, E. Singer, R. Tourangeau, G. P. Acosta, and L. Nelson. 2006. "Experiments in Producing Nonresponse Bias." *Public Opinion Quarterly* 70 (5): 720–736.

Groves, R. M., and E. Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias A Meta-Analysis." *Public Opinion Quarterly* 72 (2): 167–189.

Groves, R. M., S. Presser, and S. Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68: 2–31.

Groves, R. M., E. Singer, and A. Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 64 (3): 299–308.

Harkness, J., M. Stange, K. I. Cibelli, P. Mohler, and B. E. Pennell. 2014. "Surveying Cultural and Linguistic Minorities." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T. P. Johnson, K. M. Wolter, and N. Bates, 245–269. Cambridge, UK: Cambridge University Press.

Hattam, V. 2007. *In the Shadow of Race*. Chicago: University of Chicago Press.

Hedges, B. M. 1979. "Sampling Minority Populations." In *Social and Educational Research in Action,* edited by M. J. Wilson, 244–261. London: Longman.

Jackson, J. S., and G. Gurin. 1987. *National Survey of Black Americans, 1979–1980*. Vol. 8512. Ann Arbor, MI: Inter-University Consortium for Political & Social Research.

Jackson, J. S., and G. Gurin. 1999. "*National Survey of Black Americans, 1979–1980 [computer file]. ICPSR08512-v1*." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Jackson, J. S., P. Gurin, and S. J. Hatchett. 1989. *National Black Election Study, 1984. ICPSR08938-v1*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. **http://doi.org/10.3886/ICPSR08938.v1**.

Jackson, J. S., V. L. Hutchings, R. Brown, and C. Wong. 2004. *National Politics Study. ICPSR24483-v1*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-03-23. **http://doi.org/10.3886/ICPSR24483.v1**.

Jackson, J. S., and H. G. Neighbors. 1997. *National Survey of Black Americans, Waves 1–4, 1979–1980, 1987–1988, 1988–1989, 1992. ICPSR06668-v1*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. **http://doi.org/10.3886/ICPSR06668.v1**.

Junn, J., T. S. Lee, K. Ramakrishnan, and J. Wong. 2011. "Asian-American Public Opinion." In *The Oxford Handbook of American Public Opinion and the Media*, edited by Robert Y. Shapiro and Lawrence R. Jacobs, 520–534. Oxford, New York: Oxford University Press.

Kalton, G., and D. W. Anderson. 1986. "Sampling Rare Populations." *Journal of the Royal Statistical Society* 149 (1): 65–82.

Keeter, S., C. Miller, A. Kohut, R. Groves, and S. Presser. 2000. "Consequences of Reducing Nonresponse in a Large National Telephone Survey." *Public Opinion Quarterly*, 64: 125–148.

Keeter, S., C. Kennedy, M. Dimock, J. Best, and P. Craighill. 2006. "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey." *Public Opinion Quarterly* 70 (5): 737–758.

Lavelle, B., M. D. Larsen, and C. Gundersen. 2009. "Strategies for Surveys of American Indians." *Public Opinion Quarterly* 73 (2): 385–403.

Lazerwitz, B. 1978. "An Estimate of a Rare Population Group: The U.S. Jewish Population." *Demography* 15 (3): 389–394.

Lien, P., M. M. Conway, and J. S. Wong. 2004. *The Politics of Asian Americans: Diversity and Community*. New York: Routledge.

Lin, I., and N. Schaeffer. 1995. "Using Survey Participants to Estimate the Impact of Non-participation." *Public Opinion Quarterly* 59: 236–258.

Link, M. W., M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad. 2008. "A Comparison of Address-Based Sampling (ABS) versus Random Digit Dialing (RDD) for General Population Surveys." *Public Opinion Quarterly* 72: 6–27.

Lohr, S., and J. N. K. Rao. 2006. "Estimation in Multiple-Frame Surveys." *Journal of the American Statistical Association* 101 (475): 1019–1030.

Muedini, F. 2009. "Muslim American College Youth: Attitudes and Responses Five Years After 9/11." *The Muslim World* 99 (1): 39–59.

Muslim American Public Opinion Survey (MAPOS). 2010. **http://www.muslimamericansurvey.org/**.

Muslim Public Affairs Council. 2005. *Religion & Identity of Muslim American Youth Post-London Attacks*. Washington, DC: Muslim Public Affairs Council.

Neyman, J. 1934. "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection." *Journal of the Royal Statistical Society* 97 (4): 558–625.

Nie, N. H., J. Junn, and K. Stehlik-Barry. 1996. *Education and Democratic Citizenship in America*. Chicago: University of Chicago Press.

Nobles, M. 2000. *Shades of Citizenship: Race and the Census in Modern Politics*. Palo Alto, CA: Stanford University Press.

O'Donnell, J. A., H. L. Voss, R. R. Clayton, G. T. Slatin, and R. G. Room. 1976. *Young Men and Drugs: A Nationwide Survey; National Institute on Drug Abuse Research Monograph*. Washington, DC: US Department of Health and Human Services.

Olson, K. 2006. "Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias." *Public Opinion Quarterly* 70 (5): 737–758.

Patterson, D., G. Gasim, and J. Choi. 2011. "Identity, Attitudes, and the Voting Behavior of Mosque-Attending Muslim-Americans in the 2000 and 2004 Presidential Elections." *Politics and Religion* 4 (2): 289–311.

Pew Research Center. 2007. *Muslim Americans: Middle Class and Mostly Mainstream*. Washington, DC: Pew Research Center.

Pew Research Center for the People & the Press. 2011. *Muslim Americans: No Signs of Growth in Alienation or Support for Extremism*. Washington, DC: Pew Research Center.

Pew Research Center. 2012a. *Mormons in America: Certain in Their Beliefs, Uncertain of Their Place in Society*. Washington, DC: Pew Research Center.

Pew Research Center. 2012b. *The Rise of Asian Americans*. Washington, DC: Pew Research Center.

Pew Research Center. 2013. *A Portrait of Jewish Americans*. Washington, DC: Pew Research Center.

Peytchev, A., E. Peytcheva, and R. M. Groves. 2010. "Measurement Error, Unit Nonresponse, and Self-Reports of Abortion Experiences." *Public Opinion Quarterly* 74 (2): 319–327.

Reed, J. S. 1975. "Needles in Haystacks: Studying 'Rare' Populations by Secondary Analysis of National Sample Surveys." *Public Opinion Quarterly* 39 (4): 514–522.

Rossi, P. H., and R. Crain. 1968. "The NORC Permanent Community Sample." *Public Opinion Quarterly* 32 (2): 261–272.

Rossi, P. H., J. D. Wright, G. A. Fisher, and G. Willis. 1987. "The Urban Homeless: Estimating Composition and Size." *Science* 235: 1136–1141.

Rothbart, G. S., M. Fine, and S. Sudman. 1982. "On Finding and Interviewing the Needles in the Haystack: The Use of Multiplicity Sampling." *Public Opinion Quarterly* 46 (3): 408–421.

Sakshaug, J. W., T. Yan, and R. 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 (5): 907–933.

Sharif, A., H. Jawad, P. Nightingale, J. Hodson, G. Lipkin, P. Cockwell, S. Ball, and R. Borrows. 2011. "A Quantitative Survey of Western Muslim Attitudes to Solid Organ Donation." *Transplantation* 9 (10): 1108–1114.

Shor, R. 2000. "Jewish Immigrant Parents from the Former Soviet Union: A Method for Studying their Views of How to Respond to Children's Misbehavior." *Child Abuse & Neglect* 24 (3): 353–362.

Sirken, M. G. 1972. "Stratified Sample Surveys with Multiplicity." *Journal of the American Statistical Society* 67 (3): 224–227.

Sirken, M. G., B. I. Graubard, and M. J. McDaniel. 1978. "National Network Surveys of Diabetes." *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 631–635.

Smith, T. W. 2002. "Review: The Muslim Population of the United States; The Methodology of Estimates." *The Public Opinion Quarterly* 66 (3): 404–417.

Stinchcombe, A., L. C. Jones, and P. B. Sheatsley. 1981. "Nonresponse Bias for Attitude Questions." *Public Opinion Quarterly* 45 (3): 359–375.

Sudman, S., and G. Kalton. 1986. "New Developments in the Sampling of Special Populations." *Annual Review of Sociology* 12: 401–429.

Tate, K. 1997. *National Black Election Study, 1996. ICPSR version. Columbus, OH: Ohio State University [producer];* Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004. **http://doi.org/10.3886/ICPSR02029.v1**.

Teitler, J. O., N. E. Reichman, and S. Sprachman. 2003. "Costs and Benefits of Improving Response Rates for a Hard-to- Reach Population." *Public Opinion Quarterly* 67: 126–138.

Tighe, E., D. Livert, M. Barnett, and L. Saxe. 2010. "Cross-SurveyAnalysis to Estimate Low- Incidence Religious Groups." *Sociological Methods & Research* 39 (1): 56–82.

Tourangeau, R., R. M. Groves, and C. D. Redline. 2010. "Sensitive Topics and Reluctant Respondents: Demonstrating a Link between Nonresponse Bias and Measurement Error." *Public Opinion Quarterly* 74 (3): 413–432.

Venette, R. C., R. D. Moon, and W. D. Hutchison. 2002. "Strategies and Statistics of Sampling for Rare Individuals." *Annual Review of Entomology* 47: 143–174.

Verba, S. 1995. "The Citizen as Respondent: Sample Surveys and American Democracy Presidential Address, American Political Science Association." *American Political Science Review* 90 (1): 1–7.

Verba, S., K. L. Schlozman, and H. E. Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, MA: Harvard University Press.

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

Waksberg, J., D. Judkins, and J. T. Massey. 1997. "Geographic-based Oversampling in Demographic Surveys of the United States." *Survey Methodology* 23: 61–72.

Watters, J. K., and P. Biernacki. 1989. "Targeted Sampling: Options for the Study of Hidden Populations." *Social Problems* 36 (4): 416–430.

Wong, J. S. 2006. *Democracy's Promise: Immigrants and American Civic Institutions*. Ann Arbor: University of Michigan Press.

Wong, J. S., K. Ramakrishnan, T. Lee, and J. Junn. 2011. *Asian American Political Participation: Emerging Constituents and Their Political Identities*. New York: Russell Sage Foundation.

Zogby, J. 2001. *Muslims in the American Public Square*. Washington, DC: Zogby International.

Zogby, J. 2004. *Muslims in the American Public Square: Shifting Political Winds and Fallout from 9/11, Afghanistan, and Iraq*. Washington, DC: Zogby International.

Zukin, C. 2006. "Presidential Address: The Future Is Here! Where Are We Now? and How Do We Get There?" *Public Opinion Quarterly* 70 (3): 426–442.

Notes:

(¹) The proportion of adults in America who are Muslim is a contested matter (see Smith 2002 for a review), although Pew Research Center (2007, 2011) places the share at about .5 percent.

(²) This latter concern eventually led Pew methodologists to alter their sampling strategy in a follow-up survey of American Muslims (2011), amending the RDD frame to include both cellular and landline numbers.

(³) See, for example, Dana, Barreto, and Oskooii (2011); Djupe and Calfano (2012); and Patterson, Gasim, and Choi (2011); all comparing original data on American Muslims to Pew's (2007) sample to gauge representativeness.

Justin Berry

Political Science, Kalamazoo College

Assistant Professor, Department of Political Science, Kalamazoo College

Youssef Chouhoud

Political Science, University of Southern California

Youssef Chouhoud, Professor of Political Science, University of Southern California

Jane Junn

Political Science, University of Southern California

Professor of Political Science, University of Southern California

Subscriber: University of Southern California; date: 26 October 2017