



Can Non-full-probability Internet Surveys Yield Useful Data? A Comparison with Full-probability Face-to-face Surveys in the Domain of Race and Social Inequality Attitudes

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

| | |
|-------------------|--|
| Citation | Simmons, A. D., and L. D. Bobo. 2015. "Can Non-Full-Probability Internet Surveys Yield Useful Data? A Comparison with Full-Probability Face-to-Face Surveys in the Domain of Race and Social Inequality Attitudes." <i>Sociological Methodology</i> 45 (1) (March 11): 357–387. doi:10.1177/0081175015570096. |
| Published Version | doi:10.1177/0081175015570096 |
| Citable link | http://nrs.harvard.edu/urn-3:HUL.InstRepos:22565792 |
| Terms of Use | This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP |

**Can Non-Full-Probability Internet Surveys Yield Useful Data?:
A Comparison to Full-Probability Face-to-Face Surveys
in the Domain of Race and Social Inequality Attitudes***

Alicia D. Simmons
Colgate University

Lawrence D. Bobo
Harvard University

Word Count (text, footnotes, and references) = 9,074

* We thank Adam Berinsky, Howard Schuman, Nicholas Valentino, and our anonymous reviewers for helpful comments on earlier drafts. Direct correspondence to Alicia D. Simmons, Department of Sociology and Anthropology, Colgate University, 13 Oak Drive, Hamilton, NY 13346; email: asimmons@colgate.edu.

**Can Non-Full-Probability Internet Surveys Yield Useful Data?:
A Comparison to Full-Probability Face-to-Face Surveys
in the Domain of Race and Social Inequality Attitudes**

We investigate the potential utility of web-based surveys of non-full-probabilistically sampled respondents for social science

research. Specifically, we compare demographic, attitude response, and multivariate model results produced by two distinct survey modalities: the traditional full-probability sample face-to-face survey and the non-full probability web survey. Using data from the 2009 Race Cues, Attitudes, and Punitiveness Survey (RCAPS), the 2008 General Social Survey (GSS), and the 2008 American National Election Study (ANES), we find that: a) the unweighted demographic differences between surveys tend to be slight, b) in comparison to the GSS and ANES, RCAPS respondents are more interested in politics and ideologically polarized, c) in comparison to the ANES, RCAPS respondents are more racially and socially conservative, often selecting the most extreme response option, d) when the dependent variable is a more general and abstract measure of social attitudes, the multivariate models generated by the ANES and RCAPS show several differences that are trivial in magnitude, however, when the dependent variable is a more specific and concrete measure, the models show remarkable similarity, and e) RCAPS multivariate models consistently explain more variance than ANES models. Overall, our findings show both substantial similarities across the two survey modalities as well as a few clear, reasonably well-specified differences.

KEYWORDS: Methodology; Survey; Mode; Sampling; Attitudes

**Can Non-Full-Probability Internet Surveys Yield Useful Data?:
A Comparison to Full-Probability Face-to-Face Surveys
in the Domain of Race and Social Inequality Attitudes**

1. INTRODUCTION

Approaches to survey research have evolved in important ways over the years. To cite just two examples: a heavy reliance on open ended-questions and loosely structured guidelines yielded to closed-ended questions with tightly structured questionnaires, and quota sampling designs gave way to full-probability sample designs (Converse 1987). The forces driving these changes, such as the desire for higher quality data and a push to collect data in more efficient and cost effective ways, continue to influence how researchers balance maintaining high-quality data standards against financial and temporal costs. One current horizon for survey methodology concerns modes of survey administration; today, scholars are increasingly interested in using web-based surveys for research purposes. In their examination of survey modes used in approximately 2,000 reports featured in issues of *Survey Research* from 1995-2009, Brown and Johnson (2011) found that 70.5 percent were by telephone, 21.7 percent were face-to-face, 17.4 percent were by mail, and 5.6 percent were web-based. While telephone surveys were dominant in the overall sample, trends revealed that the prevalence of telephone surveys declined over the 14-year span while the prevalence of web surveys increased. In 2009, web surveys were the primary or secondary mode of administration in 33.3 percent of studies. Not only are web surveys frequently used, they are used to conduct high-caliber research, appearing in leading journals such as the *American Sociological Review* (e.g., Schwartz and Schuman 2005;

Rosenfeld and Thomas 2012), the *American Journal of Political Science* (e.g., Prior 2005; Brooks and Geer 2007; Jerit 2009; Ansolabehere and Jones 2010), and the *Journal of Personality and Social Psychology* (e.g., Skitka et al. 2002; Seery, Holman, and Silver 2010).

In addition to investigating new survey modes, the field has also been exploring sampling techniques that differ from the consensual gold-standard of full-probability designs. In the context of web surveys, these non-full-probability designs rely on cultivating opt-in pools of respondents, and then applying sampling algorithms to the pool in order to construct representative samples for individual surveys (e.g., Rivers 2007).

This paper examines the potential utility of web surveys of non-full-probabilistically sampled respondents for social science research. Specifically, we compare demographic, attitude response, and multivariate model results produced by two distinct survey modalities: a non-full probability web survey and two of the most highly regarded face-to-face social science surveys using full-probability samples, the General Social Survey and the American National Election Study. Given its design, this study cannot fully untangle the unique effects of mode and sampling strategy; however, it can compare and contrast their intersection, highlighting the important similarities and differences in the resulting data.

While the bulk of contemporary research in this vein has focused on surveys of broad political attitudes (e.g., Malhotra and Krosnick 2007; Chang and Krosnick 2009; Stephenson and Crête 2011), this project advances the literature by focusing on social attitudes that underlie support for public policies: egalitarianism and racial resentment (Krysan 2000; Hutchings and Valentino 2004). This focus is an important extension for two primary reasons. First, it creates an opportunity to explore how respondents' expressions of social attitudes vary with the presence of human interviewers. Krysan (1998) links whites' expressions of racial attitudes to survey mode, finding a positive relationship between expressions of racial conservatism and the level of privacy provided by the survey context. Web surveys provide an ideal setting for respondent privacy, greatly benefiting researchers who are concerned about interviewers' effects on respondents' expressed opinions. Second, our work addresses researchers' critiques that experimental investigations of prejudice are overly-reliant on convenience samples of college students (e.g., Sears 1986; Henry 2008). If data about social attitudes collected from non-full-probability sampled (NPS) web surveys are substantially similar to that collected from full-probability sampled (FPS) face-to-face surveys, this strategy may provide a cost-effective and more efficient way to reach meaningful samples of respondents, while simultaneously providing greater external validity.

To foreshadow a bit, our results point to five broad conclusions. First, the demographic representativeness of the NPS web survey mode is roughly equivalent to the FPS face-to-face survey mode. Second, the NPS web survey respondents are more interested in politics and tend to be stronger political ideologues than the FPS face-to-face survey respondents. Third, the NPS web survey respondents tend to express more socially conservative opinions than the FPS face-to-face survey respondents, particularly when it comes to voicing the most extreme opinions. Fourth, the similarity of multivariate models produced by the

surveys appears to be contingent on how general and abstract versus specific and concrete respondents perceive the attitude domain to be. When examining more general and abstract attitude constructs, the models show several differences that are trivial in magnitude; in contrast, when the dependent variable is more specific and concrete, we see greater similarity across models. Fifth, the NPS web survey consistently explains more variance in the dependent variables than the FPS face-to-face surveys. Given these findings, and in light of the potential cost-savings and speed of data collection advantages that NPS web surveys provide, there are a variety of circumstances when this methodology may be used to study the empirical and theoretical relationships between variables, while remaining mindful of when and to what extent mode and sampling choices may affect results and the scope of legitimate inferences.

2. BACKGROUND

An exhaustive comparison of the full constellation of survey modes and sampling strategies is beyond the scope of our capabilities given the available resources; thus, this paper illuminates one important set of comparisons, namely, how full-probability face-to-face surveys compare to non-full-probability website surveys (henceforth we use “survey modalities” to denote this particular set of survey modes and sampling strategies). To do so, we begin with a brief review of the differences across survey modes and how these characteristics influence data quality. Next, we discuss the promises and pitfalls of full-probability and non-full-probability sampling designs. Finally, we discuss how the intersection of survey modes and sampling strategies can influence data quality.

2.1. *Characteristics of Survey Modes*

Our first concern is with how differences in the face-to-face and web survey modes influence data quality. As outlined by de Leeuw (1992), survey modes can be differentiated along three broad dimensions: media factors, information transmission, and interviewer impact.

The most important media factor distinguishing the face-to-face and web modes is familiarity with the media. While the majority of respondents likely have extensive experience with face-to-face interactions, familiarity with website environments may be more variable. Thus, those respondents who are less familiar with web environments potentially face an added cognitive burden that could reduce data quality.

An information transmission factor differentiating the face-to-face and web modes is the presentation of stimuli, such that the modes differ in regards to their temporal ordering. Face-to-face modes are predominantly sequential, with little (perceived or actual) opportunity for backtracking. In contrast, temporal ordering is more variable in the web mode. At one end of the continuum, there might be a great deal of respondent-driven choice, with several different questions appearing on one page, or an explicit option may be available for revising previous answers (e.g., a “back” button). Alternatively, there might be no opportunity for respondent-driven choice, such that respondents are provided with one question at a time and not given the

option to revise previous responses. The ability of respondents to alter the order in which they answer questions or to revise previous answers may influence the resulting data via consistency or contrast effects occurring among questions involving identical or closely related issues (Schuman and Presser 1981).

Finally, face-to-face and web modes have differing opportunities for interviewer impact on respondents; while a human interviewer has a great potential to influence a respondent in the face-to-face mode, interviewer effects are largely absent in the web mode. The presence of a human interviewer may positively impact a respondent by providing motivation, clarifying the nature of questions, and clarifying the nature of respondents' answers. However, an interviewer may have negative impacts; most importantly, interviewers may reduce respondents' feelings of privacy, thus compromising the data they provide due to social desirability concerns. For example, Krysan (1998) argues that respondents holding racially conservative opinions view the basis of their beliefs as non-racial. In a public survey mode, "... a respondent cannot be certain that the interviewer will appreciate their "nonracist" but "conservative" response. Respondents... may feel pressure in the face-to-face conditions to give a "safe" and "liberal-sounding" response, lest their motives be misinterpreted by the interviewer" (532). Respondents can freely express racially conservative opinions in a private survey mode because in this mode, their non-racist self-image is not jeopardized.

2.2 *Sampling Strategies in the Web Mode*

Our second concern centers on how sampling strategy influences data quality. Here, the critical issue centers on how accurately the sample represents the target population. Couper (2000) argues that coverage error is one of the primary concerns facing researchers using web surveys. *Coverage error* arises when members of the target population under study are excluded from the *sampling frame*, or the enumerable and bounded group from which a sample will ultimately be drawn. For example, researchers interested in studying the target population of American adults might use a sampling frame of residential telephone numbers; in this instance, coverage error arises because American adults without residential telephone numbers are excluded.

In the case of web surveys, one factor influencing coverage error is the proportion of the target population that can be accessed via the Internet, or the *coverage rate* (Couper 2000). This may be affected in two ways. First, the online sampling universe is limited to those who have Internet access. While the number of American adults online has grown dramatically from 46 percent in 2000 to 85 percent in 2013, a non-trivial segment of the population remains uncovered; in addition, 9 percent of those who do use the Internet do not do so at home, where respondents may be more likely to participate in web-based surveys (Zickuhr 2013). Second, Zickuhr's (2013) work indicates that Internet access is not randomly distributed throughout the U.S. population. While adult men and women are equally represented online, Internet use is lower among Hispanics when compared to blacks and whites. Furthermore, Internet usage is negatively correlated with age, and positively correlated with education and income.

The second factor influencing coverage error in web surveys concerns constructing a sampling frame that accurately represents the general population. Web surveys using *full-probability samples* (FPS) employ techniques such as random digit dialing (RDD) to recruit individuals into their respondent pool, and then provide computers and Internet access to those who do not already possess them free of charge. Survey-firms then randomly select individuals from their pools for specific surveys. While this method is quite costly, it creates a sampling frame of respondents that is representative of the general population, not merely the population that uses the Internet on its own accord.

In contrast, web surveys using *non-full-probability samples* (NPS) use recruiting techniques such as solicitations sent out to email lists purchased from Internet marketing firms, banner advertisements appearing on popular websites, registration processes for sweepstakes, and RDD. Individuals are then sampled from these pools using an algorithm in order to obtain a sample that is representative of the target population.

Constructing a NPS is less costly than creating a pool by exclusively using probability methods. However, because the sampling frame is assembled in a non-full-probability fashion, the distribution of demographic variables and attitudes in the pool are more likely to systematically differ from the general population. Post-stratification weights can be used to more closely align the demographic characteristics of the sample and the overall population; however, if the variables of interest are weakly correlated with demographic variables, then weighting will not lead to alignment in the substantive answers of the two groups. Previous research by Chang and Krosnick (2009) and Malhotra and Krosnick (2007) indicates that before weighting, the demographic distribution of NPS web surveys differ from their FPS counterparts, however, these differences are significantly reduced once the appropriate weights are applied.

2.3 Response Quality Concerns in the Web Mode

Ultimately, we are interested in how the intersection of survey mode and sampling strategy influences response quality. One aspect of response quality is the degree of similarity in response distributions produced by various survey modalities. Smith's (2003) comparison of 32 items tapping attitudes about government spending in the General Social Survey and a FPS web survey found that the distribution of responses produced by the two datasets were typically within five percentage points. Despite the similarities, two important differences were noteworthy. First, the web survey produced more extreme responses on Likert-scale items. Second, discrepant response distributions were particularly notable on questions dealing with spending on problems largely associated with the lower class.

Smith (2003) attributes both of these divergences to the lack of interviewer effects in the web mode, which allows respondents the privacy and comfort to report more socially conservative responses. Supporting this interpretation, Chang and Krosnick's (2009) study about support for government aid to blacks found that data generated by web surveys show less evidence of social

desirability bias than telephone data. In addition, Kreuter, Presser, and Tourangeau (2008) surveyed university alumni via the web, telephone, or interactive voice recognition about desirable (e.g., GPA > 3.5; received honors) and undesirable characteristics (e.g., GPA < 2.5; received a D or F). When comparing respondents' self-reports to their university transcripts, they found that while the reporting of desirable information did not vary by mode, respondents were significantly more likely to report undesirable information in the web mode than in telephone mode. Furthermore, web respondents were significantly less likely to misreport their information in a socially desirable way as compared to telephone respondents.

Beyond interviewer effects, response quality may be influenced by respondents' engagement with the survey's topic. Chang and Krosnick (2009) find that NPS web respondents are more interested in and knowledgeable about the survey topic than FPS respondents, a finding they attribute to the former respondents' selectively completing surveys that they are highly interested in. Furthermore, Heerwegh (2009) finds that web respondents are no more likely than face-to-face respondents to engage in satisficing, a form of non-engagement with the survey instrument.

A second aspect of response quality is the degree to which the multivariate models produced by distinct survey modalities show similarities in the direction, magnitude, and significance of predictor variables. In their comparison of FPS telephone, FPS web, and NPS web surveys, Berrens et al. (2003) found that while differences in the distributions of response options and relational inferences did arise, overall, the multivariate results produced by the three survey types were quite similar.

2.4 Hypotheses

In light of the results of previous research, we offer four hypotheses. Our first set of analyses focuses on the extent to which the two survey modalities accurately reflect the demography of the United States.

Hypothesis 1: Due to differences in coverage error, we expect that the demographic characteristics of the FPS face-to-face surveys will be more accurate than those of the NPS web survey. However, we further expect that the divergences across surveys will not be overwhelming (e.g. Malhotra and Krosnick 2007), primarily due to the efforts of the proprietary algorithms that NPS web survey uses to draw a representative sample from its pool.

We further expect to see some differences in the distribution of responses across the survey modalities.

Hypothesis 2: We expect that web respondents will report more extreme opinions than their face-to-face counterparts (Smith 2003). We expect that this is partially a sampling issue, as members of a survey pool exercise selectivity in choosing surveys to participate in (Chang and Krosnick 2009), and partially a mode issue, as respondents may be more comfortable expressing extreme opinions when a human interviewer is absent.

Hypothesis 3: Web respondents will report less socially desirable opinions than their face-to-face counterparts (Chang and Krosnick 2009; Kreuter et al. 2008). We expect this to occur due to the lack of a human interviewer in the web survey mode (Krysan 1998).

Finally, we consider the degree of similarity across multivariate models produced by the two survey modalities.

Hypothesis 4: Despite the slight differences in response distributions (occurring due to sampling and mode effects) we expect the multivariate models of the FPS face-to-face surveys and the NPS web survey to exhibit substantial similarity in the direction, magnitude, and significance of the coefficients of the predictor variables (Berrens et al. 2003).

3. DATA AND MEASURES

3.1. *Data*

Our analyses employ three datasets: the 2008 General Social Survey (GSS), the 2008 American National Election Study (ANES), and the 2009 Race Cues, Attitudes, and Punitiveness Survey (RCAPS). While an ideal research design would examine surveys administered in the same year, we find a one-year discrepancy to be acceptable, given previous research indicating that the marginal distributions of the patterns of relationships between racial attitudes and other variables very rarely undergo substantial shifts in short one-year time spans (e.g., Schuman, Steeh, Bobo, and Krysan 1997; Bobo, Charles, Krysan, and Simmons 2012).

General Social Survey. The 2008 GSS was conducted via face-to-face interviews. The sampling universe contained English- or Spanish-speaking U.S. household residents who were eighteen years-of-age or older. Sampling occurred across multiple stages, each focusing on a geographic area of decreasing size. From February through April 2008, 2,044 interviews were completed, with an AAPOR Response Rate 5 of 70.4 percent.

American National Election Survey. The 2008 ANES was conducted via face-to-face interviews. The sampling universe contained English- or Spanish-speaking U.S. citizens living in the 48 coterminous states or the District of Columbia who were eighteen years-of-age or older by October 31, 2008. As rates of citizenship vary across race/ethnicity, the ANES included oversamples of Latinos and African Americans. Sampling occurred across multiple stages focusing on areas of decreasing geographic size, and was stratified by metropolitan statistical area and race/ethnicity. Between September 2 and November 3, 2008, the pre-election survey interviewed 2,323 respondents, with an AAPOR Response Rate 1 of 59.5 percent. The post-election survey interviewed 2,102 respondents between November 5 and December 30, 2008, a re-interview rate of 90.5 percent.

Race Cues, Attitudes, and Punitiveness Survey. The 2009 RCAPS was conducted via the Internet, achieving an AAPOR Response Rate 1 of 27.3 percent. The sampling universe contained members of an opt-in survey pool, consisting of over one million U.S. residents recruited through Internet advertising, email, postal mail, and RDD. Respondents were sampled from the pool using the *sample matching* process. First, the sampling frame was created by drawing a sample from the 2006 U.S. Census American Community Survey, stratified by age, race, gender and education, with random sampling within the strata. The resulting data included information on respondents' age, race, gender, education, marital status, number of children under age 18, family income, employment status, citizenship, and state/metropolitan area of residence. Data on interest in politics and party identification were imputed into this sampling frame from the 2004 National Annenberg Election Study on the basis

of age, race, gender and education. The second step of sample matching was to create a profile of the target sample, known as a *target matrix*. This was accomplished by using a stratification scheme developed by cross-classifying gender, age (i.e., 18-24; 25-34; 35-54; 55+), race (i.e., black; Latino; white and all non-black/non-Hispanic), and education (i.e., high school graduate or less; some college; college graduate; post-graduate education). In the final step of sample matching, individuals were selected from the opt-in pool such that each individual selected from the pool was as similar as possible on the range of characteristics to a specific member of the target sample. The resulting sample included 1,500 respondents who were representative of the U.S. population, and an oversample of 600 African American respondents.

3.2 Measures

The variables used in this study fall into three categories: demographic characteristics, political orientations, and social attitudes about egalitarianism and African Americans. The demographic variables include sex, race/ethnicity, region of residence, age, education, and household income. The political orientation variables include interest in politics and political ideology. The latter variable is scored such that higher values indicate stronger affiliation with Conservative ideology. The social attitudes variables include affect toward whites and blacks (calculated as a difference score between a respondent's ratings of the two groups), stereotypes about the industriousness and intelligence of whites and blacks (summed into an index and then calculated as a difference score between a respondent's ratings of the two groups), a six-item egalitarianism index, and a four-item racial resentment index. The affect and stereotypes measures are scored such that higher values indicate rating whites more positively than blacks. The egalitarianism index is scored such that higher values indicate greater social liberalism. The racial resentment index is scored such that higher values indicate greater racial conservatism.

4. ANALYSIS AND RESULTS

4.1. Analytic Approach

First, we begin by assessing the extent to which the two survey modalities accurately represent the demography of the U.S. population. To do so, we compare the distributions shown by each survey to estimates provided by the 2008 U.S. Census Current Population Survey March Supplement (CPS). Second, we assess data quality across survey modalities in two ways. First, we investigate the extent to which the modalities produce similar descriptions of the empirical world by comparing their distributions of political and social attitudes. Second, we assess the extent to which the modalities produce similar models of social processes by modeling the egalitarianism and racial resentment indexes. Our approach unfolds in two steps: we examine the models independently generated by each survey and then we pool the surveys, including a dummy variable for survey type.

All analyses are conducted on unweighted data, allowing us to assess their raw capabilities as opposed to the survey firms' success at constructing sampling weights. In addition, the analyses are constrained to white respondents only. We adopt this strategy because the ANES oversamples Blacks and Latinos, and the RCAPS oversamples Blacks. Given that variables such

as education, income, and political ideology systematically vary with race/ethnicity, the unweighted data including varying oversamples are necessarily discrepant and would thus cloud the results.

4.2. *Demographic Representativeness*

Table 1 presents the unweighted distribution of demographic variables from the CPS, GSS, ANES, and RCAPS, and the average percentage-point error of each survey when compared to the CPS benchmark. We used t-tests to compare the means for each response option in the GSS, ANES, and RCAPS to the CPS benchmark.

**** TABLE 1 ABOUT HERE ****

Each survey shows notable similarities and differences to the CPS benchmark. The GSS is the most accurate in emulating the benchmark distributions, with an overall average error of 3.4 percentage-points. Its estimates are statistically equivalent to the benchmark in regards to sex. Its divergences from the benchmark occur in its estimates of region (under-representing those in the Northeast and over-representing those in the South), age (under-representing younger Americans and over-representing older Americans), education (over-representing those at the tail-ends of the distribution, and under-representing those with some college/bachelor's degree), and income (over-representing those with the lowest income and under-representing those with the highest income).

The ANES has an overall average error of 3.9 percentage-points from the CPS benchmark, a 0.5 point increase from the GSS. Its divergences occur in its estimates of sex (under-representing men), region (under-representing those in the Northeast and Midwest and over-representing those in the South and West), age (under-representing younger Americans and over-representing older Americans), education (under-representing those with bachelor's degrees and over-representing those with graduate educations), and income (under-representing those at the tail-ends of the distribution, and over-representing those in the middle).

Finally, the RCAPS has an overall average error of 3.8 percentage-points, a 0.4 point increase from the GSS, and a 0.1 decrease from the ANES. Its estimates are statistically equivalent to the benchmark in regards to sex. Its divergences occur on measures of region (under-representing those in the Northeast and over-representing those in the South), age (under-representing the youngest Americans and over-representing those aged 50-64), education (under-representing those with less than high school and those with bachelor's degrees, while over-representing those with high school degrees), and income (over-representing those earning \$25,000-74,999 and under-representing those with the highest incomes).

4.3. *Response Quality I: Distributions*

Table 2 presents the unweighted distribution of the interest in politics and political ideology variables. We used t-tests to compare means of different survey sources with the benchmark.

**** TABLE 2 ABOUT HERE ****

First, there is a statistically significant difference in the degree of interest in politics in the ANES and RCAPS, such that RCAPS respondents report far more interest than ANES respondents. Second, there are a variety of significant differences in political ideology across the surveys, with a broad pattern of the GSS and ANES being more similar to each other than to the RCAPS (although notably, the ANES reports the smallest number of moderates). In comparison to the other surveys, the RCAPS features heavier tail-ends of the distribution.

Table 3 presents the unweighted distribution of the social attitude variables: affect toward whites and blacks, stereotypes about whites and blacks, egalitarianism, and racial resentment toward blacks. We used t-tests to compare the means for each response option across surveys. Recall the affect, stereotyping, and racial resentment variables are scored such that higher values indicate greater racial conservatism, while the egalitarianism variable is scored such that higher values indicate greater social liberalism.

**** TABLE 3 ABOUT HERE ****

First, there was a significantly larger gap in affect toward whites and blacks in the RCAP than in the ANES. Second, there was a significant difference in the gap in stereotyping of whites and blacks across all three surveys, with the ANES showing the largest gap and the GSS showing the smallest. Third, scores on the egalitarianism and racial resentment indexes significantly differ between the ANES and RCAPS, such that RCAPS respondents reported less egalitarianism and greater resentment than ANES respondents. In additional analyses not shown here, we find that these differences are attributable to RCAPS respondents disproportionately selecting the most conservative response option.

4.4 Response Quality II: Multivariate Models

The next portion of our analysis assesses the similarity of multivariate models produced by the two survey modalities. To do this, we use the unweighted data from white respondents to regress egalitarianism and racial resentment on the demographic, political orientation, and social attitude variables. The models proceed in several steps, first entering demographic variables, then political and social attitudes, and finally interest in politics, as this latter variable is a notable point of distinction between surveys, with the RCAPS respondents being particularly interested in politics.

Table 4 presents separate models from the ANES and RCAPS predicting egalitarianism. Overall, these models reveal similarities and differences between the surveys. In regards to similarities, both surveys consistently show that income and increasingly conservative ideology are negatively related to egalitarianism, and that interest in politics exhibits no significant relationship. After political ideology is entered in the model 2, the ANES and RCAPS both reveal that Southern residence has no significant relationship with egalitarianism. In regards to differences, first, while the ANES shows a negative relationship between age and egalitarianism, the RCAPS shows a positive relationship, a finding that runs counter to theoretical expectations. Predicted probabilities indicate that the magnitude of the age effect is trivial; a one-unit increase in age is

associated with a -0.001 change in egalitarianism in the ANES, and a 0.002 change in the RCAPS. Thus, the egalitarianism scores of a twenty-year-old and a forty-year-old would differ by 0.02 points in the ANES and 0.04 points in the RCAPS. Second, while the ANES reveals that education and being female have a positive relationship with egalitarianism, education never attains significance in the RCAPS models, and sex is rendered insignificant when political ideology is added to the RCAPS model. Predicted probabilities indicate that the magnitude of the education effect is trivial; a one-unit increase in education is associated with a 0.02 change in egalitarianism in the ANES, and a 0.01 change in the RCAPS. Thus, the egalitarianism scores of a high-school graduate and a college graduate would differ by 0.08 in the ANES and 0.04 in the RCAPS. Likewise, the effect of sex is small; the shift from male to female is associated with a 0.05 change in the ANES and a 0.07 change in the RCAPS. Third, it is noteworthy that models 2 and 3 of the RCAPS consistently explain a larger share of the variance than the ANES models.

**** TABLE 4 ABOUT HERE ****

Table 5 presents models from the pooled ANES and RCAPS data predicting egalitarianism. Here, we find that education and being female are positively related to egalitarianism, while income and increasingly conservative political ideology are negatively related. Importantly, we find that the survey modality variable is significant, such that being an RCAPS respondent is negatively associated with egalitarianism. However, the substantive impact is small; shifting from the ANES to the RCAPS is associated with a -0.03 change in egalitarianism. Finally, the interactions of survey with age, education, and income are not significant predictors of egalitarianism. These latter findings bolster our confidence that the substantive importance of the divergences illustrated in table 4 is trivial.

**** TABLE 5 ABOUT HERE ****

Next, table 6 presents separate models of the ANES and RCAPS predicting racial resentment. Here, we find greater similarity across the two surveys. In the final model, both surveys show that education and egalitarianism have negative relationships with racial resentment. Furthermore, they reveal that racial resentment is positively related to Southern residence, increasingly conservative political ideology, and increasingly large gaps in affective feelings and stereotyping of blacks and whites (such that whites are favored). The only difference between the models is that in model 1 the RCAPS shows a positive relationship between age and resentment; but this relationship becomes insignificant when the political and inter-group attitudes are added in model 2. Finally, as was the case in the egalitarianism models, the RCAPS explains a larger amount of variance than the ANES.

**** TABLE 6 ABOUT HERE ****

Table 7 presents models from the pooled ANES and RCAPS data predicting racial resentment. Here, we find that education and egalitarianism are negatively associated with racial resentment, while resentment is positively associated with Southern residence, increasingly conservative political ideology, and increasingly large gaps in affective feelings and stereotyping of

blacks and whites (such that whites are favored). Importantly, we find that the survey modality variable is not a significant predictor of racial resentment. Finally, while the interaction of survey with age is not a significant predictor of racial resentment, the interaction of survey and education is. Here, we find that the liberalizing effect of education is about 60 percent stronger among the RCAPS respondents than among the ANES respondents. Substantively, however, this is a small effect; predicted probabilities indicate that a one-unit increase in education is associated with a 0.027 change in racial resentment in the ANES, and a 0.042 change in the RCAPS. Thus, the resentment scores of a high-school graduate and a college graduate would differ by 0.11 in the ANES and 0.17 in the RCAPS.

**** TABLE 7 ABOUT HERE ****

5. DISCUSSION AND CONCLUSION

In support of hypothesis 1, we find that while a FPS face-to-face survey is the most accurate in emulating the CPS benchmark regarding social demographics, the average errors are comparable across the three surveys: the average error of the GSS is 3.4 percentage-points, the RCAPS is 3.8, and the ANES is 3.9. Out of the three surveys, the RCAPS is the most accurate at representing sex and region, and is the least accurate at representing age, education, and income. The RCAPS significantly under-represents respondents under age 49, while over-representing those aged 50-64. In regards to education, the RCAPS significantly under-represents those with less than a high school diploma and those with a bachelor's degree, while significantly over-representing those who are high school graduates. Finally, the RCAPS significantly over-represents those earning \$25,000-74,999 while significantly under-representing those earning \$75,000 or more. Given that previous research indicates that Internet adoption is negatively correlated with age and positively correlated with education and income (Zickuhr 2013), our findings concerning demographic representativeness are surprising. Thus, it appears that access to the Internet is not the only important factor in determining who opts into NPS web surveys. This finding resonates with research indicating that various demographic groups have divergent patterns in their *use* of the Internet once they are online. For example, social class has important implications for online activities; there are positive correlations between a user's income bracket and their engagement in activities such as researching products, getting news online, visiting government websites, and researching medical issues (Jansen 2010). In addition, research finds positive correlations between educational attainment and online activities such as email use and product purchasing, while finding negative correlations between age and activities such as social networking use and online banking (Zickuhr and Smith 2012).

The second set of analyses concerned the quality of the more substantive data provided by each survey. First, we compared the distribution of the political orientation and social attitudes across the three samples. In support of hypothesis 2, we find that NPS web respondents report more extreme opinions than FPS face-to-face respondents. For example, in comparison to the GSS and ANES respondents, RCAPS respondents report far more interest in politics and are more likely to adopt "very liberal" and "very conservative" political ideologies. This finding may be based in respondent selectivity that occurs during the

sampling process (e.g., Chang and Krosnick 2009). Selectivity might operate at the beginning of the process, whereby people who have strongly held views may be more likely to join opt-in pools than those with weaker views. Alternatively, it might operate later in the process; because web respondents have a broad sense of the types of surveys that are available (due to their membership in a sample pool), they can exercise discretion about which surveys to complete, only choosing to participate in surveys that closely align with their interests. Despite the fact that individuals are more likely to cooperate with surveys on topics that they are personally interested in, this tendency does not necessarily influence data quality (Groves, Presser, and Dipko 2004).

In support of hypothesis 3, we find that NPS web respondents report more socially undesirable attitudes than FPS face-to-face respondents. In comparison to the ANES respondents, RCAPS respondents report less egalitarianism and more racial resentment. On one hand, these findings may occur because of NPS web respondents' differential selectivity during the sampling process, such that respondents with strongly held views were more likely to agree to participate in this survey. On the other hand, these findings may have roots in survey mode; without a human interviewer present, web respondents may feel more comfortable expressing socially undesirable opinions than face-to-face respondents (e.g., Krysan 1998).

As a second test of data quality, we compared multivariate models predicting egalitarianism and racial resentment. Here, we find mixed support for hypothesis 4, which asserted that the multivariate models would exhibit substantial similarity in the direction, magnitude, and significance of the coefficients of the predictor variables. Considering egalitarianism, we find several instances where the survey modalities differ. While the ANES shows that education and sex are significant predictors of egalitarianism, the RCAPS shows no such relationships. In addition, while the ANES reveals that age has a negative relationship with egalitarianism, the RCAPS reveals a positive relationship. Overall, we find the differences between the models to be substantively minor. First, the substantive importance of these differences is trivial; each one-unit increase in age is related to a 0.001-0.002 shift in egalitarianism, one-unit increases in education are associated with a 0.01 shift, and a transition from male to female is associated with a 0.05-0.07 shift. Second, the interactions of survey with age, education, and income are not significant predictors of egalitarianism. Furthermore, while in the pooled model we find that survey mode is a significant predictor of egalitarianism, such that RCAPS respondents are less egalitarian than ANES respondents, we find this to be a substantively small effect; shifting from the ANES to the RCAPS is associated with a -0.03 change in egalitarianism. Considering racial resentment, we find stronger support for hypothesis 4 in the form of remarkable similarity across survey modalities. In the separate regression models, we find only one instance of difference: when the demographic variables are entered in model 1, the RCAPS indicates age to be a significant predictor of resentment while the ANES does not. In all of the subsequent models, the surveys reveal identical estimates of which predictor variables are significant, their direction, and their relative magnitudes. In the pooled models, the survey mode variable is insignificant, indicating that ANES and RCAPS respondents are similarly resentful. Further analyses reveal that there is a significant interaction between survey and education,

such that RCAPS respondents experience a stronger liberalizing effect of education. This is, however, a small effect; predicted probabilities indicate that a one-unit increase in education is associated with a 0.027 change in racial resentment in the ANES, and a 0.042 change in the RCAPS.

To understand why the survey modalities produce similar multivariate models of racial resentment but appear more dissimilar in the multivariate models of egalitarianism, we suggest that one likely explanation involves differences in the concreteness of the underlying measures for most respondents. Research on question order and context effects has shown that abstract and general attitude items may be more susceptible to context effects than specific and concrete items (Schuman and Presser 1981; Schuman 2008). How respondents understand the meaning of the former is more open to influence by preceding items in a questionnaire than is the latter. The measure of egalitarianism we employ is aimed at tapping a highly general orientation toward inequality between a great variety social groups, across many situations (e.g., “If people were treated more equally in this country we would have many fewer problems”). As such, the egalitarianism scale aims at the abstract ideas of equality of opportunity, legal equality, and equality of rewards (Feldman 1999; Sidanius and Pratto 1999). Almost of necessity, therefore, respondents are prompted to draw upon a very broad range of considerations when forming judgments, and a variety of differences in questionnaire content across the two survey modalities may, therefore, come to influence how individuals react. In contrast, the racial resentment items, relatively speaking, constitute a much more common and ordinary aspect of routine social discourse on black-white relations in the U.S. (Edsall and Edsall 1991; Kinder and Sanders 1996; Lamont 2000; Bobo et al 2012).

Overall, our results point to five key findings. First, the ability to estimate the distribution of demographic characteristics in the general population is roughly equivalent across survey modalities. However, researchers must be aware that there are some instances where predictable demographic divergences, such as in the distribution of age, education, or income may have implications for their results. Second, NPS web respondents tend to be more politically interested and ideologically extreme than FPS face-to-face respondents. Third, NPS web respondents tend to produce more racially and socially conservative responses than respondents from traditional surveys, often choosing the most extreme response. Fourth, the extent to which FPS face-to-face and NPS web surveys produce similar multivariate models appears to be contingent upon the nature of the dependent variable. When the dependent variable is a more general and abstract measure of attitudes, the survey modalities show several differences, yet these tend to be minor in magnitude. In contrast, when the dependent variable is a more specific and concrete measure, the survey modalities show greater similarity. Fifth, the NPS web models explain noticeably more variance than traditional surveys. This likely is linked to the sampling process; since these respondents are more interested and politically engaged, there is more constraint and coherence in their opinions and thus more variance is explained. In addition, this may be an artifact of the disproportionate amount of the NPS web respondents selecting the most extreme response option.

To conclude, this project has illuminated how full-probability face-to-face surveys compare to non-full-probability web surveys. Given its design, this study cannot determine when differences found between surveys are the result of mode (i.e., face-to-face versus web), sampling strategy (i.e., non-full-probability versus full-probability), or a combination of the two; however, it can determine the degree to which the data collected from FPS face-to-face surveys is similar to from that collected from a NPS web survey. This is one important set of comparisons, a full research program that investigates the other intersections of survey modes and sampling strategies is a worthy next step.

With this limitation in mind, we do not wish to lose sight of the forest for having delved carefully into looking at each and every tree. A holistic read of the results leads us to conclude that the data generated by NPS web surveys is in many ways comparable to that from FPS face-to-face surveys, and thus has a place in the social science toolkit. Researchers must be mindful, however, of the instances where NPS web surveys are unique: they contain notable error in the unweighted estimates of some demographic groups, their respondents are more politically interested and ideologically extreme, their respondents are more likely to select extreme response options, and their higher variance may be more of an artifact of sample composition than a true instance of enhanced explanatory power. While weighting may correct for distortions in social demographic variables, it cannot smooth out all of the observed divergences in attitudes. Thus, given these differences, and the secretive nature of Internet survey-houses' proprietary sampling algorithms, we agree with AAPOR's recommendation that this type of data should not be used to generate precise estimates of population values (Baker et al. 2010). This is one task where the gold-standard of face-to-face surveys using full-probability samples excel. Furthermore, these types of surveys are invaluable because there are many instances where respondents can benefit from the presence of a human interviewer.

We also believe, however, that it is in the best interest of science for there to be an abundance of data available to researchers; this allows us to answer new questions and to replicate the findings of previous scholars. In order to meet these goals, it is not in the best interest of the scholarly community to hold everyone to the financially costly and time-consuming gold-standard of FPS face-to-face surveys, for as the work of Keeter et al. (2000) indicates, there are instances where the empirical quality of data are unrelated to the amount of time and financial resources dedicated to collecting it. Here, we have shown that the similarities of NPS web surveys and FPS face-to-face surveys often outweigh their differences. Furthermore, we have specified the conditions under which these differences arise. With these specified limitations in mind, researchers should move forward in using NPS web surveys for the types of projects that they are well-suited for, exploring the relationships between variables and testing theories, while avoiding attempts to make precise population estimates.

NOTES

REFERENCES

- Ansolabehere, Stephen and Philip E. Jones. 2010. "Constituents' Responses to Congressional Roll-Call Voting." *American Journal of Political Science* 54(3):583-597.
- Baker, Reg, Stephen Blumberg, J. Michael Brick, Mick P. Couper, Melanie Courtright, Mike Dennis, Don Dillman, Martin R. Frankel, Philip Garland, Robert M. Groves, Courtney Kennedy, Jon Krosnick, Sunghee Lee, Paul J. Lavrakas, Michael Link, Linda Piekarski, Kumar Rao, Douglas Rivers, Randall K. Thomas, and Dan Zahs. 2010. "AAPOR Report on Online Panels." Retrieved January 3, 2013. (http://www.aapor.org/AM/Template.cfm?Section=AAPOR_Committee_and_Task_Force_Reports&Template=/CM/ContentDisplay.cfm&ContentID=2223).
- Berrens, Robert P., Alok K. Bohara, Hank Jenkins-Smith, Carol Silva, and David Weiman. 2003. "The Advent of Internet Surveys for Political Research: A Comparison on Telephone and Internet Samples." *Political Analysis* 11(1): 1-22.
- Bobo, Lawrence D. 1991. "Social Responsibility, Individualism, and Redistributive Policies." *Sociological Forum* 12(1):147-176.
- Bobo, Lawrence D., Camille Z. Charles, Maria Krysan, and Alicia D. Simmons. 2012. "The Real Record on Racial Attitudes." Pp. 38-83 in *Social Trends in American Life: Findings from the General Social Survey Since 1972*, edited by P. V. Marsden. Princeton, NJ: Princeton University Press.
- Brooks, Deborah Jordan and John G. Geer. 2007. "Beyond Negativity: The Effects of Incivility on the Electorate." *American Journal of Political Science* 51(1):1-16.
- Brown, Ethan and Timothy P. Johnson. 2011. "Diffusion of Web Survey Methodology, 1995-2009." *Survey Research* 42(1):1-3.
- 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(4):641-678.
- Converse, Philip E. 1987. "Changing Conceptions of Public Opinion in the Political Process." *Public Opinion Quarterly* 51(50th Anniversary Issue):S21-S24.
- Couper, Mick P. 2000. "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly* 64(4):464-494.
- de Leeuw, Edith Desiree. 1992. *Data Quality in Mail, Telephone, and Face to Face Surveys*. Amsterdam: TT-Publikaties.
- Edsall, Thomas B. and Mary D. Edsall. 1991. *Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics*. New York: W. W. Norton & Company.
- Feldman, Stanley. 1999. "Economic Values and Inequality." in *Measures of Political Attitudes*, edited by J. P. Robinson, P. R. Shaver, and L. S. Wrightsman. San Diego, CA: Academic Press.

- Groves, Robert M., Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68(1):2-31.
- Heerwegh, Dirk. 2009. "Mode Differences between Face-to-Face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects." *International Journal of Public Opinion Research* 21(1):111-121.
- Henry, P. J. 2008. "College Sophomores in the Laboratory Redux: Influences of a Narrow Data Base on Social Psychology's View of the Nature of Prejudice." *Psychological Inquiry* 19(2):49-71.
- Hutchings, Vincent L. 2009. "Change or More of the Same?: Evaluating Racial Attitudes in the Obama Era." *Public Opinion Quarterly* 73(5):917-942.
- Hutchings, Vincent L. and Nicholas A. Valentino. 2004. "The Centrality of Race in American Politics." *Annual Review of Political Science* 7:383-408.
- Jansen, Jim. 2010. "Use of the Internet in Higher-Income Households." Washington, D.C.: Pew Research Center's Internet & American Life Project. Retrieved September 15, 2014. (<http://www.pewinternet.org/files/old-media/Files/Reports/2010/PIP-Better-off-households-final.pdf>).
- Jerit, Jennifer. 2009. "How Predictive Appeals Affect Policy Opinions." *American Journal of Political Science* 53(2): 411-426.
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64(2):125-148.
- Kinder, Donald R. and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic Ideals*. Chicago: University Of Chicago Press.
- Kluegel, James R. and Eliot R. Smith. 1986. *Beliefs About Inequality: Americans' Views of What Is and What Ought to Be*. Piscataway, NJ: Transaction Publishers, Rutgers - The State University.
- Kreuter, Frauke, Stanley Presser, and Roger Tourangeau. 2008. "Social Desirability Bias in CATI, IVR, and Web Surveys." *Public Opinion Quarterly* 72(5):847-865.
- Krysan, Maria. 1998. "Privacy and the Expression of White Racial Attitudes: A Comparison Across Three Contexts." *Public Opinion Quarterly* 62(4):506-544.
- . 2000. "Prejudice, Politics, and Public Opinion: Understanding the Sources of Racial Policy Attitudes." *Annual Review of Sociology* 26(1):135-168.
- Krysan, Maria and Mick P. Couper. 2003. "Race in the Live and Virtual Interview: Racial Difference, Social

- Desirability, and Activation Effects in Attitude Surveys." *Social Psychology Quarterly* 66(4):364-383.
- Lamont, Michele. 2000. *The Dignity of Working Men: Morality and the Boundaries of Race, Class, and Immigration*. New York: Russell Sage Foundation.
- Malhotra, Neil and Jon A. Krosnick. 2007. "The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to Internet Surveys with Nonprobability Samples." *Political Analysis* 15:286-323.
- McDermott, Monica. 2006. *Working Class White: The Making and Unmaking of Race Relations*. Berkeley, CA: University of California Press.
- Prior, Markus. 2005. "News vs. Entertainment: How Increasing Media Choice Widens Gaps in Political Knowledge and Turnout." *American Journal of Political Science* 49(3):577-592.
- Rivers, Douglas. 2007. "Sampling for Web Surveys." in *Joint Statistical Meetings*. Salt Lake City, UT.
- Rosenfeld, Michael J. and Ruben J. Thomas. 2012. "Searching for a Mate: The Rise of the Internet as a Social Intermediary." *American Sociological Review* 77(4):523-547.
- Schuman, Howard. 2008. *Method and Meaning in Polls and Surveys*. Cambridge, MA: Harvard University Press.
- Schuman, Howard, Charlotte Steeh, Lawrence D. Bobo, and Maria Krysan. 1997. *Racial Attitudes in America: Trends and Interpretations*. Cambridge, MA: Harvard University Press.
- Schuman, Howard and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.
- Schwartz, Barry and Howard Schuman. 2005. "History, Commemoration, and Belief: Abraham Lincoln in American Memory, 1945-2001." *American Sociological Review* 70(2):183-203.
- Sears, David O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature." *Journal of Personality and Social Psychology* 51(3):515-530.
- Seery, Mark D., E. Alison Holman, and Roxane Cohen Silver. 2010. "Whatever Does Not Kill Us: Cumulative Lifetime Adversity, Vulnerability, and Resilience." *Journal of Personality and Social Psychology* 99(6):1025-1041.
- Sidanius, Jim and Felicia Pratto. 1999. *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression*. New York: Cambridge University Press.
- Skitka, Linda J., Elizabeth Mullen, Thomas Griffin, Susan Hutchinson, and Brian Chamberlin. 2002. "Dispositions, Scripts, or Motivated Correction?: Understanding Ideological Differences in Explanations for Social Problems." *Journal of Personality and Social Psychology* 83(2):470-487.
- Smith, Tom W. 2003. "An Experimental Comparison of Knowledge Networks and the GSS." *International Journal of*

Public Opinion Research 15(2):167-179.

Stephenson, Laura B. and Jean Crête. 2011. "Studying Political Behavior: A Comparison of Internet and Telephone Surveys." *International Journal of Public Opinion Research* 23(1):24-55.

Zickuhr, Kathryn. 2013. "Who's Not Online and Why." Washington, D.C.: Pew Research Center's Internet & American Life Project. Retrieved (http://www.pewinternet.org/~media/Files/Reports/2013/PIP_Offline

[%20adults_092513_PDF.pdf](http://www.pewinternet.org/~media/Files/Reports/2013/PIP_Offline)). Zickuhr, Kathryn and Aaron Smith. 2012. "Digital Differences." Washington, D.C.: Pew Research Center's Internet & American Life Project. Retrieved September 15, 2014. (http://www.pewinternet.org/files/oldmedia/Files/Reports/2012/PIP_Digital_differences_041312.pdf).

| TABLE 1 | | | | |
|--|------|---------|----------|---------|
| Comparison of Unweighted Demographic Variable Distributions to the CPS Benchmark | | | | |
| | CPS | GSS | ANES | RCAPS |
| <i>Sex</i> | | | | |
| Male | 47.9 | -2.1 | -3.8** | -0.3 |
| Female | 52.1 | 2.1 | 3.8** | 0.3 |
| Average error | | 2.1 | 3.8 | 0.3 |
| <i>Region</i> | | | | |
| Northeast | 22.8 | -4.6*** | -10.6*** | -3.4** |
| Midwest | 26.9 | -1.9 | -5.7*** | -1.7 |
| South | 28.5 | 4.5*** | 11.3*** | 5.5*** |
| West | 21.8 | 2.0 | 5.1*** | -0.4 |
| Average error | | 3.3 | 8.2 | 2.2 |
| <i>Age</i> | | | | |
| 18-29 | 18.5 | -4.2*** | -3.0** | -4.3*** |
| 30-39 | 18.3 | -2.7** | -1.5 | -4.0*** |
| 40-49 | 22.7 | -4.6*** | -3.5** | -5.0*** |
| 50-64 | 24.7 | 4.7*** | 3.3* | 11.1*** |
| 65+ | 15.9 | 7.9*** | 4.6*** | 2.0 |

| | | | | |
|----------------------------|------|---------|---------|----------|
| Average error | | 4.8 | 3.2 | 5.3 |
| <i>Education</i> | | | | |
| > high school diploma | 9.0 | 2.9*** | 0.6 | -4.7*** |
| High school graduate | 30.7 | -0.5 | -0.5 | 13.4*** |
| Some college | 29.7 | -5.1*** | 0.2 | -1.9 |
| Bachelor's degree | 20.1 | -2.4* | -3.6*** | -5.6*** |
| Graduate school | 10.5 | 5.1*** | 3.3*** | -1.0 |
| Average error | | 3.2 | 1.6 | 5.3 |
| <i>Household income</i> | | | | |
| ≤\$24,999 | 17.7 | 5.8*** | -3.2* | 1.5 |
| \$25-49,999 | 21.5 | 0.9 | 1.2 | 5.0*** |
| \$50-74,999 | 18.6 | 2.9* | 4.4** | 6.7*** |
| \$75-99,999 | 14.5 | -6.0*** | 1.1 | -5.4*** |
| ≥\$100,000 | 27.6 | -3.6** | -3.4* | -10.8*** |
| Average error | | 3.8 | 2.7 | 5.9 |
| <i>Total average error</i> | | 3.4 | 3.9 | 3.8 |

Note: The analysis contains white respondents only. CPS figures represent the percentage of the response option's distribution in the population. Other values indicate the magnitude of the target survey's divergence from the CPS benchmark. Negative values indicate estimates that are lower than the benchmark while positive values indicate estimates that are higher. Significance is determined by t-tests comparing each response option to the benchmark. *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$

| TABLE 2 | | | |
|--|-----|-------------------|------------------|
| Comparison of Unweighted Political Variable Distributions across Surveys | | | |
| | GSS | ANES | RCAPS |
| <i>Interest in politics</i> | | | |
| Hardly at all | - | 14.3 ^a | 4.0 ^b |

| | | | |
|---------------------------|-------------------|-------------------|-------------------|
| Only now and then | - | 22.9 ^a | 11.0 ^b |
| Some of the time | - | 32.2 ^a | 24.1 ^b |
| Most of the time | - | 30.7 ^a | 60.9 ^b |
| <i>Political ideology</i> | | | |
| Very liberal | 3.1 ^a | 3.2 ^a | 6.1 ^b |
| Liberal | 21.3 ^a | 24.6 ^a | 15.7 ^b |
| Moderate | 36.3 ^a | 26.5 ^b | 29.9 ^a |
| Conservative | 35.7 ^a | 41.0 ^b | 28.5 ^c |
| Very conservative | 3.6 ^a | 4.6 ^a | 19.7 ^b |

Note: The analysis contains white respondents only. Values represent the percentage of the response option's distribution in the population. Significance is determined by t-tests. Superscripts denote differences that are significant at least at the $p \leq 0.05$ level. Items that do not have a superscript, or that share a common superscript, do not differ significantly.

| | GSS | ANES | RCAPS |
|-------------------------|-------------------|-------------------|-------------------|
| Affect score difference | - | 0.07 ^a | 0.10 ^b |
| Stereotyping difference | 0.07 ^a | 0.11 ^b | 0.09 ^c |
| Egalitarianism index | - | 0.60 ^a | 0.50 ^b |
| Racial resentment index | - | 0.65 ^a | 0.68 ^b |

Note: The analysis contains white respondents only. Values represent the mean response. The feeling thermometer and stereotype indexes represent the gap between ratings of white and black targets, with higher values indicating a greater distance between evaluations. Significance is determined by t-tests. Superscripts denote differences that are significant at least at the p

≤ 0.05 level. Items that do not have a superscript, or that share a common superscript, do not differ significantly.

| | ANES | | | | RCAPS | | |
|----------------------|--------------------|--------------------|--------------------|--|--------------------|--------------------|--------------------|
| | Model 1 | Model 2 | Model 3 | | Model 1 | Model 2 | Model 3 |
| Constant | 0.49*** (0.06) | 0.68*** (0.06) | 0.67*** (0.06) | | 0.43*** (0.07) | 0.76*** (0.05) | 0.77*** (0.06) |
| Age | -0.12** (0.04) | -0.07 (0.04) | -0.08* (0.04) | | -0.07 (0.04) | 0.08* (0.03) | 0.09* (0.03) |
| Education | 0.26*** (0.07) | 0.18** (0.07) | 0.17** (0.06) | | 0.15 (0.08) | 0.05 (0.07) | 0.06 (0.07) |
| Income | -0.14*** (0.04) | -0.08** (0.04) | -0.09** (0.03) | | -0.13*** (0.04) | -0.08** (0.03) | -0.08** (0.03) |
| Southern | -0.06*** (0.02) | -0.03 (0.02) | -0.03 (0.02) | | -0.02 (0.02) | 0.004 (0.02) | 0.01 (0.02) |
| Female | 0.05** (0.02) | 0.03* (0.02) | 0.04* (0.02) | | 0.07*** (0.02) | 0.03 (0.01) | 0.02 (0.01) |
| Political ideology | | -0.32*** (0.03) | -0.32*** (0.03) | | | -0.55*** (0.03) | -0.54*** (0.03) |
| Interest in politics | | | 0.03 (0.03) | | | | -0.02 (0.03) |
| R ² | 0.10 | 0.23 | 0.24 | | 0.04 | 0.37 | 0.37 |
| N | 537 | | | | 867 | | |

Note: The analysis contains white respondents only. *** $p \leq 0.001$, ** $p \leq 0.01$,
 * $p \leq 0.05$

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Constant | 0.40*** (0.05) | 0.59*** (0.04) | 0.62*** (0.04) | 0.69*** (0.04) | 0.65*** (0.06) |
| Age | -0.09** (0.03) | 0.01 (0.03) | 0.03 (0.03) | 0.02 (0.03) | -0.02 (0.04) |
| Education | 0.25*** (0.06) | 0.16*** (0.05) | 0.18*** (0.05) | 0.12* (0.05) | 0.17* (0.07) |
| Income | -0.12*** (0.03) | -0.06** (0.02) | -0.06* (0.02) | -0.08*** (0.02) | -0.05 (0.04) |
| Southern | -0.03* (0.01) | 0.000 (0.01) | -0.001 (0.01) | -0.01 (0.01) | -0.01 (0.01) |
| Female | 0.06*** (0.01) | 0.04*** (0.01) | 0.03** (0.01) | 0.03** (0.01) | 0.03** (0.01) |
| Political ideology | | -0.30*** (0.01) | -0.30*** (0.01) | -0.30*** (0.01) | -0.30*** (0.01) |
| Interest in politics | | | -0.06** (0.02) | -0.01 (0.02) | -0.01 (0.02) |
| RCAPS | | | | -0.09*** (0.01) | -0.02 (0.08) |
| Age x RCAPS | | | | | 0.06 |

| | | | | | |
|-------------------|-------|------|------|------|--------|
| | | | | | (0.05) |
| Education x RCAPS | | | | | -0.10 |
| | | | | | (0.10) |
| Income x RCAPS | | | | | -0.05 |
| | | | | | (0.05) |
| R ² | 0.05 | 0.29 | 0.29 | 0.32 | 0.32 |
| N | 1,404 | | | | |

Note: The analysis contains white respondents only. *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$

| TABLE 6 | | | | | | | | | |
|--|----------|----------|----------|----------|--|----------|----------|----------|----------|
| Regression of Racial Resentment across Surveys | | | | | | | | | |
| | ANES | | | | | RCAPS | | | |
| | Model 1 | Model 2 | Model 3 | Model 4 | | Model 1 | Model 2 | Model 3 | Model 4 |
| Constant | 0.93*** | 0.66*** | 0.98*** | 0.98*** | | 1.04*** | 0.69*** | 1.06*** | 1.07*** |
| | (0.06) | (0.05) | (0.06) | (0.06) | | (0.06) | (0.05) | (0.05) | (0.05) |
| Age | 0.04 | -0.04 | -0.06 | -0.06 | | 0.12** | -0.05 | -0.01 | -0.002 |
| | (0.04) | (0.03) | (0.03) | (0.03) | | (0.04) | (0.03) | (0.03) | (0.03) |
| Education | -0.42*** | -0.29*** | -0.26*** | -0.27*** | | -0.55*** | -0.40*** | -0.41*** | -0.39*** |
| | (0.06) | (0.06) | (0.05) | (0.05) | | (0.07) | (0.06) | (0.05) | (0.05) |
| Southern | 0.10*** | 0.05*** | 0.05*** | 0.05*** | | 0.06** | 0.03* | 0.03* | 0.03* |
| | (0.02) | (0.02) | (0.01) | (0.01) | | (0.02) | (0.01) | (0.01) | (0.01) |
| Political ideology | | 0.31*** | 0.18*** | 0.18*** | | | 0.45*** | 0.21*** | 0.22*** |
| | | (0.03) | (0.03) | (0.03) | | | (0.02) | (0.03) | (0.03) |

| | | | | | | | | | |
|-----------------------|------|----------|----------|----------|--|------|---------|----------|----------|
| Affect difference | | 0.12** | 0.09* | 0.09* | | | 0.20*** | 0.13** | 0.13** |
| | | (0.05) | (0.04) | (0.04) | | | (0.05) | (0.04) | (0.04) |
| Stereotype difference | | 0.34**** | 0.27*** | 0.27*** | | | 0.27*** | 0.25*** | 0.50*** |
| | | (0.05) | (0.05) | (0.05) | | | (0.05) | (0.04) | (0.04) |
| Egalitarianism | | | -0.42*** | -0.42*** | | | | -0.44*** | -0.44*** |
| | | | (0.04) | (0.04) | | | | (0.03) | (0.03) |
| Interest in politics | | | | 0.002 | | | | | -0.04 |
| | | | | (0.03) | | | | | (0.03) |
| R ² | 0.09 | 0.28 | 0.37 | 0.37 | | 0.08 | 0.43 | 0.54 | 0.54 |
| N | 790 | | | | | 942 | | | |

Note: The analysis contains white respondents only. *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$

| TABLE 7 | | | | | | |
|---|----------|----------|----------|----------|----------|----------|
| Regression of Racial Resentment across Pooled Surveys | | | | | | |
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Constant | 1.00*** | 0.67*** | 1.00*** | 1.01*** | 1.03*** | 0.98*** |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) |
| Age | 0.07** | -0.05* | -0.04 | -0.03 | -0.03 | -0.06* |
| | (0.03) | (0.02) | (0.02) | (0.02) | (0.02) | (0.03) |
| Education | -0.50*** | -0.34*** | -0.32*** | -0.31*** | -0.32*** | -0.25*** |
| | (0.05) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) |
| Southern | 0.07*** | 0.04*** | 0.04*** | 0.04*** | 0.04*** | 0.04*** |

| | | | | | | |
|-----------------------|--------|---------|----------|----------|----------|----------|
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Political ideology | | 0.40*** | 0.20*** | 0.21*** | 0.20*** | 0.20*** |
| | | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| Affect difference | | 0.17*** | 0.10*** | 0.11*** | 0.11*** | 0.11*** |
| | | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) |
| Stereotype difference | | 0.30*** | 0.27*** | 0.27*** | 0.26*** | 0.26*** |
| | | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) |
| Egalitarianism | | | -0.42*** | -0.43*** | -0.43*** | -0.44*** |
| | | | (0.02) | (0.02) | (0.02) | (0.02) |
| Interest in politics | | | | -0.03 | -0.02 | -0.02 |
| | | | | (0.02) | (0.02) | (0.02) |
| RCAPS | | | | | -0.02 | 0.08 |
| | | | | | (0.01) | (0.06) |
| Age x RCAPS | | | | | | 0.06 |
| | | | | | | (0.04) |
| Education x RCAPS | | | | | | -0.16* |
| | | | | | | (0.07) |
| | | | | | | |
| R ² | 0.09 | 0.37 | 0.47 | 0.47 | 0.47 | 0.47 |
| N | 1,732 | | | | | |

Note: The analysis contains white respondents only. *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$