Public Opinion Quarterly

ARE SURVEY RESPONDENTS LYING ABOUT THEIR SUPPORT FOR SAME-SEX MARRIAGE?
LESSONS FROM A LIST EXPERIMENT

JEFFREY R. LAX
JUSTIN H. PHILLIPS*
ALISSA F. STOLLWERK

Abstract Public opinion polls consistently show that a growing majority of Americans support same-sex marriage. Critics, however, raise the possibility that these polls are plagued by social desirability bias, and thereby may overstate public support for gay and lesbian rights. We test this proposition using a list experiment embedded in the 2013 Cooperative Congressional Election Study. List experiments afford respondents an anonymity that allows them to provide more truthful answers to potentially sensitive survey items. Our experiment finds no evidence that social desirability is affecting overall survey results. If there is social desirability in polling on same-sex marriage, it pushes in both directions. Indeed, our efforts provide new evidence that a national opinion majority favors same-sex marriage. To evaluate the robustness of our findings, we analyze a second list experiment, this one focusing on the inclusion of sexual orientation in employment nondiscrimination laws. Again, we find no overall evidence of bias.

Public support for gay and lesbian rights has risen dramatically over the past two decades. Nowhere is this more apparent than in responses to survey questions about same-sex marriage. While in the mid-1990s, fewer than one-third of Americans thought it should be legal for same-sex couples to marry, surveys now indicate that a growing majority support marriage equality. This sea

JEFFREY R. LAX is an associate professor of political science at Columbia University, New York, NY, USA. JUSTIN H. PHILLIPS is an associate professor of political science at Columbia University, New York, NY, USA. ALISSA F. STOLLWERK is a doctoral candidate in the Department of Political Science at Columbia University, New York, NY, USA. The authors thank Peter Aronow, Adam Berinsky, Graeme Blair, Alex Coppock, Dan Corstange, Albert Fang, Adam Glynn, and Don Green for their helpful advice on this project. *Address correspondence to Justin Phillips, Columbia University, Department of Political Science, 420 W. 118th St., Mail Code 3320, New York, NY 10027, USA; e-mail: jhp2121@columbia.edu.

doi:10.1093/poq/nfv056

change in public attitudes has received a great deal of attention in the academic literature and has been widely reported in the media. But is it real? How certain can we be that a majority of the public now supports same-sex marriage?

Concerns arise because survey responses to potentially sensitive questions (e.g., questions about prejudice, religious attendance, drug use) are subject to social desirability bias. That is, respondents may lie to pollsters when they believe that their true opinion runs counter to perceived societal norms. As messages from cultural and political elites have becoming increasingly supportive of gay and lesbian rights, survey respondents who oppose same-sex marriage may now feel psychological pressure to conceal from pollsters their true preferences. This possibility has been raised in academic work (Egan 2008; Emerson and Essenburg 2013; Powell 2013) as well as in some media outlets (Regnerus 2013). The presence of social desirability bias may be particularly plausible given that opponents of same-sex marriage are now sometimes portrayed as being on the "wrong side of history." If such a bias is present in polling, scholars, the media, the courts, and elected officials may have a false sense of the public's opinion on marriage equality.

While traditional public opinion polls are ill equipped to tease out the presence of these effects, a technique known as a list experiment can do so. List experiments, by design, afford survey respondents anonymity that allows them to provide truthful answers to sensitive questions. Indeed, this technique is commonly employed in the social sciences to study views or behaviors that may be difficult to measure with direct questions (Gilens, Sniderman, and Kuklinski 1998; Streb et al. 2008; Lyall, Blair, and Imai 2013). In our case, we embedded a list experiment in the 2013 Cooperative Congressional Election Study (CCES), a large online and nationally representative academic survey. The design of our list experiment enables us not only to test whether social desirability bias is skewing overall measures of public support for same-sex marriage, but also to consider the possibility that this bias is not unidirectional (i.e., it may lead some subgroups of the population to overreport their support for same-sex marriage, while leading others to underreport their support).

The results of our list experiment have important implications. First, they contribute to our understanding of the state of American public opinion on same-sex marriage. Because battles over legal recognition for such marriages are ongoing, knowing where the public stands is of crucial importance, especially given the long-established link between public preferences and policymaking (Page and Shapiro 1983; Burstein 2003; Brooks and Manza 2007; Lewis and Oh 2008). Second, our study speaks to the prevalence of social desirability bias in computer-based surveys. Research suggests that such surveys, because they are completed in private, are likely to elicit truthful answers (Holbrook and Krosnick 2010).[1] As computer-based surveys become more

1. A recent study by Ansolabehere and Schaffner (2014) suggests that there is no difference in social desirability bias between telephone and computer surveys.

common, it is important to evaluate the accuracy of the data they generate. Indeed, we find no evidence that social desirability bias is skewing overall survey results on same-sex marriage. If such bias in polling on this issue exists, it pushes in both directions. Furthermore, our efforts provide new evidence that a national opinion majority favors same-sex marriage and should increase confidence in the ability of computer-based surveys to produce opinion estimates free of social desirability bias.

This paper begins by documenting the dramatic rise in public support for same-sex marriage as well as the evidence suggesting that social desirability bias may exist in polling on this issue. We then discuss list experiments as an approach for generating estimates of public preferences that avoid this bias. Next, we present and evaluate our experimental design. After confirming that the assumptions for a successful list experiment have been met, we present our findings, looking for the presence of social desirability bias at the aggregate level and across specific subgroups of the population. To evaluate the robustness of our results, we analyze a second list experiment, one focusing on the inclusion of sexual orientation in employment nondiscrimination laws. We conclude by more fully discussing the implications of our findings.

Public Support for Same-Sex Marriage

In May 1993, the Hawaii Supreme Court issued a landmark decision in *Baehr v. Lewin*. In this ruling, the Court held that by denying marriage licenses to same-sex couples, Hawaii was discriminating on the basis of sex, and thereby violating the Equal Rights Amendment in the state's constitution. This made the Hawaii Supreme Court the first court of last resort in the United States to issue a ruling in favor of same-sex marriage. While the legislature and voters overturned *Baehr* via a constitutional amendment, the Hawaii decision placed the issue of same-sex marriage on the national political agenda. It also engendered a backlash, resulting in the passage of Defense of Marriage Acts (DOMAs) by Congress and more than thirty state legislatures (Pinello 2006).

In the years following *Baehr*, national polling firms began to sporadically measure public support for legalizing marriages between same-sex couples. The first Gallup poll on the subject was conducted in March 1996 and asked respondents, "Do you think marriages between homosexuals should or should not be recognized by the law as valid, with the same rights as traditional marriages?"[2] The results were not positive for advocates of gay and lesbian rights: only 27 percent thought such marriages should be valid. The results of this early Gallup poll were consistent with those of other reputable polling firms conducted around the same time.

2. The wording of this question has changed very little over time. The most notable change is that polling firms have replaced the word "homosexuals" with "same-sex couples" or "gay and lesbian couples."

As the years passed, however, support for legalizing same-sex marriage began to rise.[3] Figure 1 plots the results of nearly 100 national surveys over a twenty-year period, along with the results from our own online survey conducted as part of the 2013 CCES. Average opinion rose steadily from the mid-1990s through the present, with recent polls showing a clear national majority in favor of legal recognition. Indeed, among those polls conducted in 2014, average support for same-sex marriage is just over 56 percent, with a Pew survey (from February 2014) placing support for marriage at 59 percent. This change in public opinion is correlated (though imperfectly) with changes in public policy (Lax and Phillips 2009; Krimmel, Lax, and Phillips 2012). Today, thirty-three states and the District of Columbia legally allow for same-sex marriages and the US Supreme Court has invalidated parts of the federal Defense of Marriage Act.

However, not everyone is convinced that public opinion on this matter has changed as much as polling suggests. A study by Powell (2013), for example, compares the accuracy of preelection polling on same-sex marriage ballot measures to similar polling on other statewide ballot issues, including taxes, bonds, and term limits. He finds that opposition to same-sex marriage is five to seven percentage points higher on Election Day than in preelection polling, but that corresponding inaccuracies are absent for other issues. For Powell, these results are consistent with the presence of social desirability bias in polling on same-sex marriage (14). In a similar study, Egan (2008) compares the outcomes of thirty-three state-level ballot measures on same-sex marriage to opinion polls, finding that preelection polls consistently underestimated opposition to same-sex marriage (i.e., support for constitutional bans) by an average of seven percentage points. Finally, a recent longitudinal study by Emerson and Essenburg (2013) tracks the opinions of nearly 1,300 Americans from 2006 to 2012, and finds little evidence of opinion change during this time. The red flags raised by these studies indicate that a careful, individual-level inquiry into the existence of social desirability bias in same-sex marriage polling is warranted.

Uncovering "True" Opinion

There is plenty of evidence that respondents sometimes provide socially acceptable (as opposed to true) responses to direct survey questions. For example, respondents have been shown to overreport voting in the most recent election (Silver, Anderson, and Abramson 1986; Presser 1990), church attendance (Hadaway, Marler, and Chaves 1993; Smith 1998), and their willingness to vote for black and female candidates (Finkel, Guterback, and Borg 1991; Berinsky 1999; Streb et al. 2008). While traditional public opinion polls are ill equipped to tease out the presence of these effects, a list experiment is

3. For a detailed discussion of the potential causes of changing public opinion on gay and lesbian rights, see Brewer (2008).

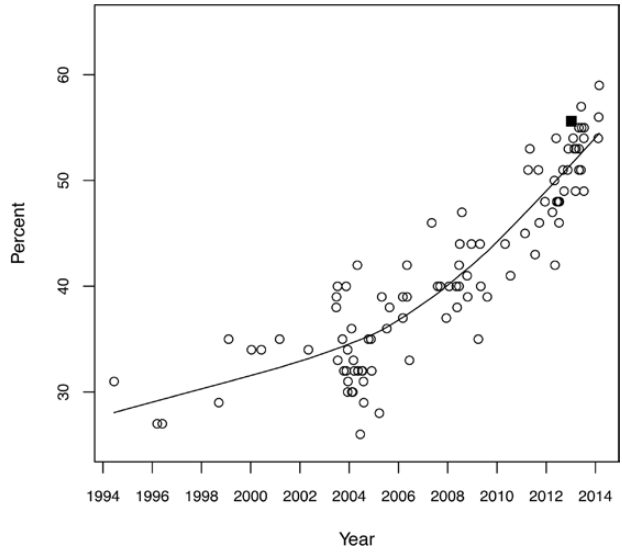

Figure 1. Support for Same-Sex Marriage (1994–2014). SOURCE.—iPoll, from the Roper Center for Public Opinion Research (January 1994 to March 2014). Each plotted circle represents a single poll result, the percentage of respondents who report (under direct questioning) support for the legalization of marriages between same-sex couples. The time trend is measured using a lowess curve. The solid square is the weighted percentage of untreated respondents from our CCES module who directly report supporting same-sex marriage.

one accepted method for doing so. This technique affords survey respondents an additional layer of anonymity that has been shown to elicit more truthful reports of behaviors and beliefs that are perceived as socially undesirable (Dalton, Wimbush, and Daily 1994; LaBrie and Earleywine 2000; Tsuchiya, Hirai, and Ono 2007; Tourangeau and Yan 2007).

In a list experiment, subjects are randomized into control and treatment groups. In the control group, subjects are given a list of J nonsensitive items and asked to report *how many*, not which ones, they support. Members of the treatment group are assigned the same task, but receive a list of $J + 1$ items. This list includes the same nonsensitive items given to the control group plus the sensitive item of interest to researchers. With a sufficiently large sample, researchers can estimate the population proportion that supports the sensitive item by taking the difference between the average response of the treatment group and the control group. By not directly asking respondents their views on the sensitive question, it is impossible for the researcher to infer a specific individual's response to the sensitive item. This veiled approach all but eliminates pressure to mislead the researcher.

List experiments, however, are not as straightforward to implement in practice as they may seem in theory. Problems can arise from a lack of statistical power, the construction of lists in which the sensitive item is obviously distinct from the control items, uneven implementation by enumerators, and the construction of lists for which a larger number of respondents support or oppose all of the nonsensitive items. In designing our experiment, we take great care to avoid these potential pitfalls, adopting best practices recommend by Glynn (2013) and others. In particular, we use a large sample size, employ nonsensitive list items that (like our sensitive item) are political in nature, and use pairs of nonsensitive items for which respondents' answers are likely to be negatively correlated.[4] For a more detailed discussion of the logic behind our design choices, see the section on experimental design.

USING LIST EXPERIMENTS TO STUDY ATTITUDES TOWARD GAYS AND LESBIANS

To the best of our knowledge, list experiments have been used in two prior instances to study the public's evolving attitudes toward gay and lesbian rights. In an unpublished paper, Goldman (2008) employs two list experiments embedded in the 2006 Cooperative Congressional Election Study to measure the public's anti-gay attitudes. Specifically, Goldman measures what proportion of the population is angered by "a gay or lesbian family moving next door" as well as "the growing acceptance of homosexuality." Goldman also asks these questions directly to measure social desirability bias, but finds no evidence of such bias in the aggregate. This paper unfortunately includes respondents from only sixteen states, and thus is not nationally representative. It also does not ask about the specific policies (e.g., same-sex marriage) that are at the center of debates over gay and lesbian rights and that have been the subject of most existing polling.[5]

In a National Bureau of Economic Research working paper, Coffman, Coffman, and Ericson (2013) find evidence of social desirability bias. Using a series of list experiments and direct questions, they show that respondents, when given the anonymity of a list experiment, are more likely to self-identify as gay, express disapproval of an openly gay manager at work, and support discrimination against gay, lesbian, or bisexual individuals. However, Coffman et al. found only a small (and statistically insignificant) difference in responses to direct and indirect questioning about support for same-sex marriage. The notable exception is self-identified Democrats, who are significantly more willing to admit to opposing same-sex marriage on a list experiment. While this study suggests the

4. Also, a list experiment assumes that people are willing to tell the truth when provided increased anonymity. If the norm to lie is so strong that respondents continue to do so even when afforded anonymity, the experiment will not reveal true opinion (for more discussion of the no-liars assumption, see Blair and Imai [2012]). We do not, however, expect the urge to lie to be so strong in our case; the sort of bias studied here is indeed the usual target for a list experiment.

5. Goldman does show that those with higher levels of education are somewhat more likely to censor expressions of anger, but only when questioned about the growing acceptance of homosexuality and not about a gay or lesbian family moving next door.

presence of social desirability bias on policy questions involving LGBT rights, some caution is warranted. In particular, Coffman et al. conduct their investigation using Amazon's Mechanical Turk (an online labor market) and therefore do not have a representative sample of the American public.[6]

Using List Experiments to Study Attitudes toward Gays and Lesbians

EXPERIMENTAL DESIGN

Our data come from a survey experiment embedded in the 2013 Cooperative Congressional Election Study (CCES).[7] Our sample, unlike those used in prior list experiments on gay and lesbian rights, is nationally representative. The list experiment was part of a survey module that asks respondents a series of direct questions about state government and their opinions on a variety of public policy matters, some of which are traditionally set at the state level and others at the national government level. Given the content of the survey module, the list experiment was unlikely to strike respondents as odd. All that differentiated the list experiment was that respondents were asked to report the number of policies they support from a list instead of responding to individual direct questions. The full text of the experiment is listed below (with the sensitive item last).

Same-Sex Marriage List:
Please take your time and tell us *how many* of the following you support. We do not need to know which ones, just *how many.*

- President Obama's healthcare reform ("ObamaCare")
- Making birth control illegal
- Cutting spending on food stamps

6. Mechanical Turk (MT) is an online labor market in which employers post solicitations for paid work (sometimes this work involves participating in social science surveys). Pay is typically substantially less than the minimum wage.

7. The CCES is a large-scale academic survey that is fielded annually by a consortium of academic researchers through an online survey conducted by YouGov. YouGov maintains a large panel of subjects who are recruited to join the panel via online advertisements. Respondents for each survey are selected from YouGov's ongoing panel of respondents, and the sample is stratified by state and congressional district. The target population for the CCES is the American public (ages 18 and up), and the survey is conducted in English. CCES surveys do not utilize true random samples (as would result from an RDD sampling methodology). However, to make the sample more representative, YouGov employs a sample-matching algorithm. In CCES surveys, all respondents answer a set of common questions and are then assigned a specific module created by one team of researchers. The 2013 CCES survey was in the field from November 6 through December 6, 2013. Of the YouGov panel enrollees who were asked to do the CCES survey, 37 percent completed the questionnaire. For more information about the CCES, see Vavreck and Rivers (2008).

- Laws that make drunk driving illegal
- *Allowing gays and lesbians to marry legally*

In designing our list experiment, we were careful to avoid potential pitfalls. First, list experiments require larger sample sizes than are typically necessary for a direct question, given the larger standard errors they produce. Corstange (2009) recommends researchers use samples of at least 1,000 respondents but suggests closer to 2,000 if possible. We follow this advice, obtaining a sample of 1,900. Second, to draw less attention to our sensitive item, each nonsensitive item in our list is also political in nature (Kuklinski, Cobb, and Gilens 1997; Glynn 2013; Aronow et al. 2015). Doing so also ensures that the list experiment blends with the rest of the survey. Third, we were careful to avoid the presence of ceiling and floor effects. That is, in selecting nonsensitive items for our lists, we ensure a low probability that any respondent would answer either "yes" or "no" to all nonsensitive items, since doing so can remove the anonymity that is essential to a list experiment. To guard against these effects, we do not include too many high- or low-prevalence items (Kuklinski, Cobb, and Gilens 1997; Tsuchiya, Hirai, and Ono 2007; Glynn 2013).

We also attempted to design a set of nonsensitive items for which the mean number of items supported is two (out of a possible four). To achieve this, we included one statement that, based on existing public opinion data, we expected almost all respondents to support—laws that make drunk driving illegal. Likewise, we included one statement that we expected almost everyone to reject—making birth control illegal. For our last two items, we chose statements expected to be negatively correlated: President Obama's healthcare reform ("ObamaCare") and cutting spending on food stamps. Glynn (2013) demonstrates that negative correlation within the list items and a modal response of support for two out of four control items will reduce variance. Given the inherent noisiness of list experiments (since they are indirect measures of preferences), it is important to use the design to lower variance wherever possible.

Participants were randomized into two groups. The control group received a list that included the first four items, while the treatment group received the full list (i.e., the control items plus the sensitive item). The order in which items appear in the lists was randomized across respondents. All respondents were also directly asked whether they support same-sex marriage ("Do you favor or oppose legally sanctioned marriages between gay and lesbian couples?"). The sensitive question was asked near the end of the survey module, well after respondents had completed the list experiment. The inclusion of the direct question provides us with the baseline estimate of public support for same-sex marriage in the survey and allows us to detect whether social desirability bias is present among some groups of respondents but not others.

Of course, in asking the direct question after the list experiment, we assume that the presence of the list experiment does not change answers to the direct question. Fortunately, this assumption can be tested by comparing the mean response to the direct question among those in the control group and those in the treatment group.

EVALUATING THE DESIGN

The evaluation of the list-experiment design, shown in table 1, provides evidence that potential problems have been avoided. First, the potential for floor and ceiling effects appears to be quite small: Out of the 899 survey respondents in the control group, only 4 percent said they supported zero policies and only 3 percent said they supported all four policies. Additionally, the modal response to our set of nonsensitive items was two, with just over 52 percent of the respondents in the control group providing that number.[8] This suggests that we achieved the desired negative correlation among our nonsensitive items.

Next, we consider whether design effects are present. A list experiment has a design effect when an individual's response to the nonsensitive items changes whether or not the sensitive item is present. For the list experiment difference-in-means estimate to be valid, the mean for support for the nonsensitive items must be the same on average across treatment and control (Imai 2011). Following Blair and Imai (2012), we use the *List* package in R to test for design effects. The *p*-value on this test is 0.63, so we cannot reject the null hypothesis of no design effect. Given this result, we move forward in the analysis of our list experiment under the assumption of no design effect.

Table 1. Observed Data

	Control group		Treatment group	
Response value	Frequency	Percentage	Frequency	Percentage
0	39	4.3	22	2.3
1	243	27.0	137	14.0
2	474	52.7	338	34.5
3	114	12.7	385	39.3
4	29	3.2	68	6.0
5			29	3.0
Total	899		979	

NOTE.—This table displays the number and percentage of respondents for each value of *Y*, the number of items that the respondent supports in the list experiment, for both the control and treatment groups. Percentages may not add up to 100 percent due to rounding.

8. Of the 1,900 respondents who were presented with either the control or treatment list, the nonresponse rate was 0.6 percent. We observe the same nonresponse rate for the direct question.

It is also important to determine whether there is balance among demographic covariates across the treatment and control groups. Table 2 suggests that there is balance: the difference in nearly all demographic categories is quite small, with the exception of ideological conservatives (the control group is 6 percent more conservative than the treatment group). Regressing assignment to treatment on these covariates shows none to be statistically significant. This is true even for ideological conservatives.[9]

Results

Table 3 shows the mean number of items supported by the control and treatment groups, the difference between the two, and the mean response to our direct question. Since one should avoid making inferences about the general population using unweighted survey data, our discussion in the text focuses on the results of our weighted sample. (Note that none of our substantive findings would differ if we relied primarily on our unweighted survey data.) By subtracting the mean number of items the control group supports from the mean number of items the treatment group supports, we obtain the list experiment estimate of support for same-sex marriage. The weighted estimate is 58.6 percent, with a 95 percent confidence interval bounded by 46.7 and 70.4 percent.[10] Since the confidence interval is fairly wide, we cannot conclude with 95 percent certainty that a majority supports marriage equality. However, we can conclude majority support with a confidence level of 85 percent.[11]

Do our estimates suggest that there is an overall social desirability bias? To address this question, we compare our list experiment estimate of opinion to our direct question estimate. In doing so, however, we must keep in mind that the direct question is a post-treatment covariate; that is, the direct question was asked after respondents had completed the list experiment (albeit with many questions between the two). To test whether this interfered with our direct estimate, we consider whether respondents who received the marriage treatment in the list experiment answered the direct question systematically differently from those in the control group. Since respondents were randomly assigned

9. As an additional test, we use randomization inference to simulate the random assignment procedure 100,000 times and calculate the F-statistic of this regression for each hypothetical sample. This collection of F-statistics can be thought of as the sampling distribution of the F-statistic under the null hypothesis that no covariate has any effect on assignment to treatment (Gerber and Green 2012). We find the *p*-value of the F-statistic by finding its location within the simulated sampling distribution. With a *p*-value of 0.20, we cannot reject the null hypothesis that the covariates are unrelated to treatment assignment, further confirming the validity of the randomization procedure.

10. For the analysis applying weights, linearized standard errors were calculated using the svy command in Stata.

11. For our sample itself, we can conclude majority support for marriage equality. Our list experiment indicates that support for same-sex marriage among our respondents (before adjusting for population weights) is 60.2 percent, with a 95 percent confidence interval bounded by 50.4 and 66.8 percent.

Table 2. Covariate Balance

	Treatment mean	Control mean	Difference
Sex			
Male	0.45	0.47	–0.02
Female	0.55	0.53	0.02
Education			
< High school	0.03	0.04	–0.00
High school	0.31	0.29	0.02
Some college	0.31	0.34	–0.03
Graduate college	0.22	0.25	–0.02
Postgrad	0.12	0.09	0.03
Race			
Asian	0.02	0.01	0.00
Black	0.10	0.11	–0.01
Hispanic	0.06	0.06	0.00
Other race	0.04	0.03	0.01
White	0.78	0.78	–0.00
Age			
18–29	0.17	0.18	–0.01
30–44	0.17	0.18	–0.01
45–64	0.44	0.43	0.01
65+	0.22	0.20	0.02
Party identification			
Democrat	0.39	0.35	0.04
Independent	0.31	0.31	–0.00
Other party	0.07	0.07	–0.00
Republican	0.23	0.26	–0.03
Political ideology			
Conservative	0.31	0.37	–0.06
Liberal	0.26	0.25	0.01
Moderate	0.34	0.30	0.04
Not sure (ideology)	0.09	0.08	0.01

NOTE.—This table displays the proportion of respondents by demographic group in the control and treatment groups and the difference between them. Percentages may not add to 100% due to rounding.

to the control group or treatment condition, the two groups should answer the direct question the same way if the list experiment has no effect on the direct response. Using a well-powered difference-of-proportions test, we find no significant difference across these two groups (55.6 percent of respondents in the control group say they support same-sex marriage treatment, while 56.9 percent report doing so in the treatment group).[12]

12. The difference of 1.3 percent is small, with a p-value of 0.66. This difference, with Ns of 899 and 979, has a power of roughly 0.9, which exceeds the standard power level of .8. This power allows us to detect a difference of proportion as small as a percentage point.

Table 3. Estimated Support for Same-Sex Marriage

	Control mean	Treatment mean	Difference-in-means	Direct question
Sample	1.83 (.03)	2.44 (.03)	0.60 (.04)***	0.57 (.01)
Weighted mean	1.83 (.05)	2.42 (.04)	0.59 (.06)***	0.56 (.02)
N	899	979		1878

NOTE.—The numbers in the parentheses are the standard errors. The reported difference-in-means may not equal the difference between the control mean and the treatment mean due to rounding. ***$p < .001$

Given that there do not appear to be any order effects, we compare our indirect and direct estimates of public opinion. The weighted direct estimate of 56.3 percent is similar to the indirect estimate (indeed, the former is actually three percentage points lower than the latter). If only answers to the direct question from the control group are used, the estimate of weighted direct support for same-sex marriage is virtually unchanged at 55.6 percent. This indicates that respondents are *not* systematically lying about their support for same-sex marriage, at least not in the way that we would anticipate if social desirability bias were inflating direct estimates of opinion. Furthermore, when we compare our list experiment estimate of opinion to recent estimates of support for same-sex marriage obtained by polling firms that employ a direct question approach, we find them to be very similar (as can be seen using the data presented in figure 1).[13] This serves as additional evidence against the presence of social desirability bias.

The results we present here obviously stand in contrast to those of Powell (2013) and Egan (2008), both of whom find evidence that is consistent with the presence of social desirability effects. As noted previously, these authors look for social desirability by comparing preelection surveys on same-sex marriage to the results of ballot measure elections. That our results differ, then, is not necessarily surprising. The tests employed by Powell and Egan are indirect, using aggregate-level data to study what is fundamentally an individual-level phenomenon.[14] We believe that individual-level evaluations are ultimately stronger tests of the truthfulness of respondents on surveys.

CONSIDERING SOCIAL DESIRABILITY BY SUBGROUPS

Until now, we have assumed that social pressure may lead opponents of same-sex marriage to conceal their true preferences. However, it may also be that for

13. Mean public support for same-sex marriage across all the phone-based surveys reported in iPoll for 2013 is 53 percent.

14. Of particular concern is that the sample in preelection surveys may not be representative of the individuals who turn out to vote. Furthermore, as Egan notes, ballot measures on same-sex marriage are often written in a manner that is intentionally confusing, whereas the questions asked by pollsters tend to be much clearer. Both factors can lead to differences between preelection polls and election results.

some types of respondents, social pressure works in the opposite direction. That is, they may feel pressured to state they oppose same-sex marriage (due to norms or pressures of their community or reference groups), when, in fact, they actually support marriage equality. Indeed, if both types of social pressure exist, they may be offsetting at the aggregate level and therefore not appear in our overall analysis. Fortunately, the inclusion of the direct question allows us to conduct the nuanced investigation that is necessary to test for conflicting forms of bias.

Table 4 compares the list experiment difference-in-means for two subgroups—those who said they support same-sex marriage when asked directly and those who did not. If there is no lying when answering the direct question, the list-experiment estimate should be 1 among those who directly report supporting same-sex marriage and 0 among those who directly report opposition.[15] This is not, however, what emerged. Among those who report that they oppose same-sex marriage, the difference of means is 0.15, and 0 is not included in a 95 percent confidence interval around the estimate. The opposite pattern emerges when considering those who report, under questioning, that they support marriage equality. Among this group of respondents, the difference of means estimate is 0.93 (while 1 falls within the 95 percent confidence interval, it does not fall within a 90 percent confidence interval). These results raise the possibility that a social desirability bias exists in polling on same-sex marriage, but that it pushes some respondents into overstating their support for marriage equality and others into underreporting their support. Indeed, the point estimates produced in the table indicate that, if anything, more respondents are *underreporting* than overreporting their support for same-sex marriage.

Table 4. Which Way Does Social Desirability Work?

	Sample				
Direct question	Control mean	Treatment mean	Expected diff.-in-means	Actual diff.-in-means	95% Conf. interval
For gay marriage	1.78 (.03)	2.71 (.04)	1.00	0.93 (.05)	[0.83, 1.03]
N	490	575			
Against gay marriage	1.90 (.04)	2.04 (.05)	0.00	0.15 (.07)	[0.02, 0.28]
N	409	404			

NOTE.—The numbers in the parentheses are the standard errors. The reported difference-in-means may not equal the difference between the control mean and the treatment mean due to rounding.

15. For examples of scholars who also utilize a direct question in the context of a list experiment, see Aronow et al. (2015) and Gonzalez-Ocantos et al. (2012).

These results prompt us to explore the possibility that social desirability operates in unique ways across subgroups. It may be that the direction in which social desirability bias works is predicted by a respondent's key reference groups. For example, a religious conservative who personally favors same-sex marriage but whose religious community is against marriage equality may conclude that the socially desirable answer is to say that one is against same-sex marriage; we might expect the reverse among someone who is a Democrat or who is not religious.

In keeping with these examples, our analysis considers whether social desirability effects differ by a respondent's partisan identification and religious affiliation. Indeed, prior work has found some evidence that self-identified Democrats are more willing to admit opposition to same-sex marriage in a list experiment than under direct questioning (Coffman, Coffman, and Ericson 2013). In addition, we consider the possibility that social desirability effects vary by geography (comparing respondents from the South to those from other regions of the country) and by educational attainment. These distinctions have been used in prior studies that explore attitudes toward minority groups, and Goldman (2008) found that highly educated respondents were more likely to censor expressions of anger about the growing acceptance of homosexuality than were the less educated.

Researchers who have previously explored cross-group differences in social desirability bias have typically done so by limiting their data only to the group of interest and then employing the difference-of-means estimator used earlier to produce the comparisons shown in table 3 (cf. Kuklinski, Cobb, and Gilens 1997). But conducting analyses in this fashion is less than ideal. When subgroups are small, tests will be underpowered. Furthermore, the basic difference-in-means approach does not allow researchers to adjust for multiple covariates at the same time.

Fortunately, Imai (2011) has proposed a regression methodology for list experiments as a solution to these problems (see also Blair and Imai [2012]). Imai's approach first estimates a multivariate model of support for same-sex marriage that uses only answers from the list experiment. To do so, he has developed a maximum likelihood estimator to estimate the joint distribution of $(Y_i(0), Z^*_{i,J+1})$, where $Y_{i(0)}$ is the number of control items supported by the *i*th respondent and $Z^*_{i,J+1}$ is the *i*th respondent's truthful answer to the sensitive item. This yields coefficients for predicting the count of nonsensitive items a respondent supports as well as the likelihood that a respondent will support the sensitive item. The multivariate regression allows the researcher to model the relationships between several respondent characteristics and their answers to the sensitive question (Blair and Imai 2012). By analyzing the treatment and control groups together, this approach relies on the fact that identical control items were asked to both groups to improve statistical efficiency.

Using the results of this regression, one then generates predicted probabilities of support for the sensitive item (in this case same-sex marriage) by the

respondent characteristics of interest. This first set of predicted probabilities can be thought of as being devoid of social desirability bias. The next step is to use a standard binary logistical regression and responses to only the direct question to generate a second multivariate model of support for the sensitive item. Again, results of the model are used to generate predicted probabilities by respondent characteristics. These estimates, because they do not rely on responses to the list experiment, can be thought of as being contaminated by social desirability bias, assuming such bias exists. The difference between the first and second set of estimates is the size of the bias, controlling for other demographic characteristics. These can be easily plotted with corresponding confidence intervals to evaluate statistical significance.

We employ the Imai approach here, and include in our models gender, education, race, age, partisan affiliation, geography (a dummy variable indicating whether the responded lives in a southern state), and a dummy variable for religious conservativism.[16] We use the *R* package designed by Blair and Imai called *list* to estimate both models.

The results of the first model (using only responses from the list experiment) are presented in the first two columns of table 5. The coefficients in the sensitive-item column predict whether someone will answer yes to the sensitive item in the list experiment (i.e., support same-sex marriage). Again, these can be thought of as showing the demographic correlates of support for same-sex marriage absent any potential social desirability effects. The coefficients in the control-items column predict the count of nonsensitive items the respondent will answer in the affirmative (note that the coefficients in columns one and two are on different scales). The last column presents the results of the standard logit model that uses only data from the direct question. The results in the table largely confirm findings in the existing public opinion literature (see Brewer [2008]).[17] Both younger respondents and respondents with higher levels of education are more likely to support marriage equality, while Republicans and religious conservatives are much less likely to support same-sex marriage.

Figure 2 displays the results of our subgroup analysis. The subgroups of interest are along the *x*-axis, and proportion support for same-sex marriage is along the *y*-axis. The bars around each estimate depict 95 percent confidence

16. We use the Census definition of the South and define a religious conservative as someone who self-identifies as either Mormon or a "born-again" Christian. We measure education using four dummy variables (the reference category is "less than a high school education"). To generate the results reported in figure 2, we re-estimate the model, replacing these four education dummy variables with a single dichotomous measure indicating whether a respondent has a college degree. This alternative specification is employed for ease of presentation, but its use does not change our findings.

17. The only odd result is that the coefficients on our measures of education fail to reach statistical significance in the direct question regression. That being said, they all have the anticipated sign and "Postgrad" would reach statistical significance at the 90 percent level, using a one-tailed test. We are not overly concerned by this result, since the education variables perform as expected in column one.

Table 5. Multivariate Analysis of the List Experiment and the Direct Question

	Sensitive item		Control items		Direct question	
Variables	Est.	(SE)	Est.	(SE)	Est.	(SE)
(Intercept)	1.78	(1.24)	–0.17	(0.17)	2.74	(0.37)***
Male	0.28	(0.62)	0.04	(0.06)	–0.41	(0.12)***
High school	2.81	(1.51)	–0.16	(0.17)	–0.08	(0.34)
Some college	4.18	(1.51)**	–0.19	(0.16)	0.08	(0.34)
Graduated coll.	3.72	(1.48)*	–0.08	(0.16)	0.08	(0.35)
Postgrad	5.70	(2.24)*	–0.21	(0.19	0.59	(0.38)
Black	0.24	(0.88)	0.16	(0.10)	–0.99	(0.20)***
Hispanic	–2.18	(0.97)*	0.21	(0.11)	–0.42	(0.24)
Age 30–44	–1.69	(1.27)	0.13	(0.10)	–0.43	(0.21)*
Age 45–64	–3.13	(1.35)*	0.22	(0.10)*	–0.76	(0.17)***
Age 65+	–2.02	(1.27)	0.22	(0.10)*	–1.25	(0.20)***
Republican	2.91	(0.90)**	–0.10	(0.09)	–2.14	(0.16)***
Independent	–1.27	(0.80)	–0.18	(0.07)*	–0.95	(0.14)***
South	–0.93	(0.59)	0.04	(0.06)	–0.27	(0.12)
Relig. conserv.	–2.67	(0.69)***	0.07	(0.08)	–1.62	(0.13)***

*$p < .05$; **$p < .01$; ***$p < .001$

intervals. We use 10,000 Monte Carlo draws to estimate confidence intervals on effects and differences in effects.[18] The top two panels report estimates of social desirability by party and region (south vs. non-South). The next set of panels report estimates by religion (religious conservative vs. other) and education level (a college degree or higher vs. less than a college degree). The final set unpacks the education results a bit further.

In the top two panels, considering all subgroups, the difference between the list experiment estimate and the direct question estimate of support for same-sex marriage is greater than zero. This implies that in each of the key subgroups shown above, there is some understating of support for same-sex marriage when respondents are asked directly. Though such a result is somewhat surprising, it is consistent with the population-level point predictions in table 4. However, in all of these cases, given the 95 percent confidence intervals, none are statistically significant.

We come closest to finding statistical significance among respondents with at least a college education. These individuals are, on average, 10 percentage points more likely to report that they support same-sex marriage in a list experiment than on a direct question. The direction and magnitude of this result are

18. As in Blair and Imai (2012), our confidence intervals are calculated by first "sampling parameters from the multivariate normal distribution with mean set to the vector of parameter estimates and the variance set to the estimated covariance matrices." Next, we "calculate each quantity of interest…and average over the empirical distribution of covariates for the entire data" (Blair and Imai 2012, 61).

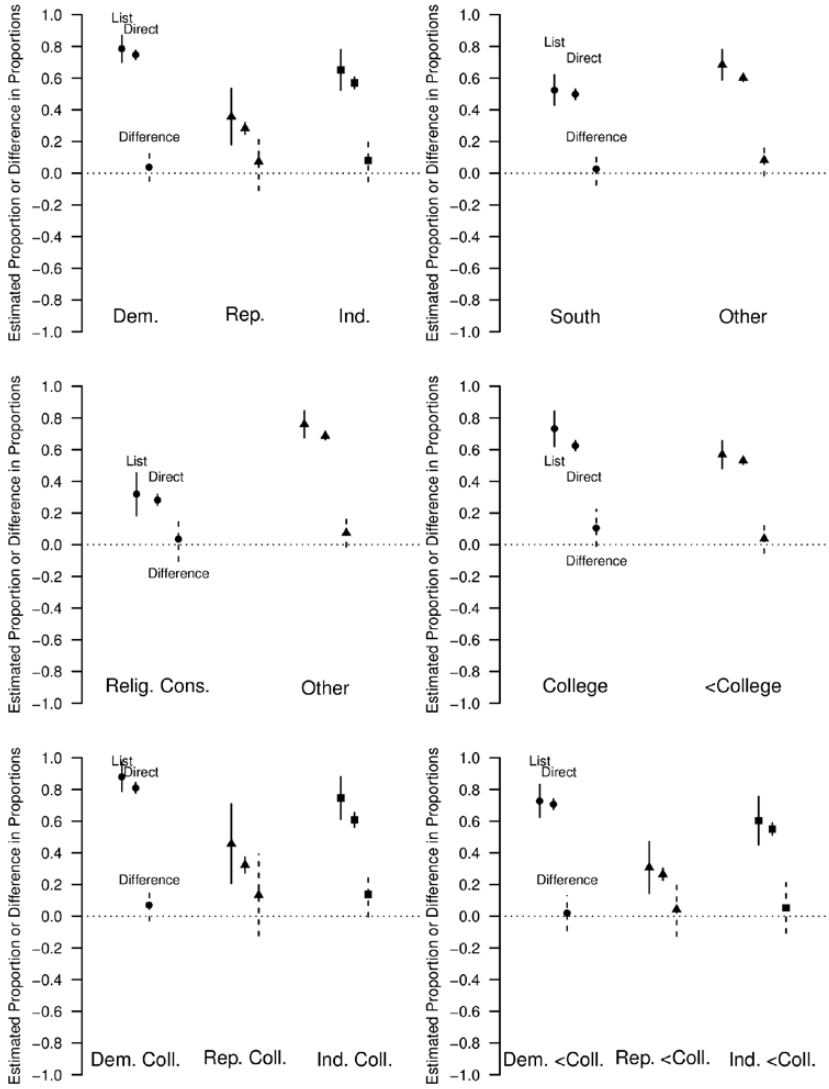

Figure 2. Multivariate Subgroup Analyses. The lines are the 95 percent confidence interval generated by Monte Carlo simulations.

not expected. Because support for marriage equality is strongly correlated to education levels, we would predict (if anything) that among educated respondents there would be social pressure to state that they support same-sex marriage when asked directly, but that when given the anonymity of a list experiment, that pressure would disappear (potentially leading to evidence of greater opposition).

To explore this further, we analyze the education results by both education and party, again using predictions based on the results in table 5. These analyses indicate that the result in the second panel is largely (though not exclusively) being driven by Republican respondents—it is well-educated Republicans who are more likely to report on a list experiment that they support marriage equality. This is a more predictable result. Still, however, the finding is not statistically significant.

Given this, what (if anything) does this subgroup analysis reveal about social desirability bias in polling on same-sex marriage? At best, the analysis presented in this section indicates that social desirability bias exists, but is not unidirectional. The results in table 4 are consistent with a world in which some respondents feel pressured to overreport their support for marriage equality while others feel pressure to underreport their support. These competing pressures are largely offsetting, and have little effect on national-level estimates of opinion.

However, it is also possible that table 4 is simply picking up noise in the data. List experiments are computationally more demanding than direct questions, which may lead some respondents to provide seemingly inconsistent answers when confronted with both types of questions. Furthermore, in neither the list experiment nor the direct question were respondents given the opportunity to provide a "don't know" answer. This means that respondents with weak or unclear preferences may be switching answers across questions.[19] Finally, that we do not uncover statistically meaningful evidence of social desirability bias in a more nuanced analysis of subgroups provides additional evidence that such a bias is simply not a factor in polling on same-sex marriage.

A FURTHER INQUIRY

While we find little to no evidence of a social desirability bias in polling on same-sex marriage, one might argue that it is too soon for such an effect to have emerged. Might we find evidence of social desirability in areas where opposition to gay rights may more clearly go against perceived societal norms of tolerance?

To test for this possibility, we analyze a second list experiment, this one focusing on employment nondiscrimination, which has been on the policy agendas of LGBT rights organizations for decades (much longer than same-sex marriage) and appears to be significantly less controversial with the American public. Figure 3 plots polls on this topic for the prior twenty years: Support

19. We do not believe that noise from the absence of a "don't know" option would bias our results in any particular direction. Indeed, there are reasons not to provide respondents with such an option when trying to get opinions devoid of social desirability bias. Berinsky (2004) finds that individuals who hold socially unacceptable opinions may hide their opinions behind a "don't know" response.

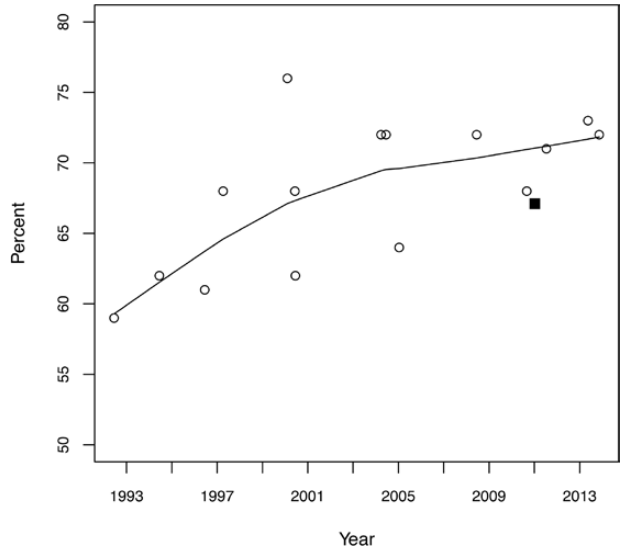

Figure 3. Support for Nondiscrimination Laws (1992–2014). SOURCE.—iPoll, from the Roper Center for Public Opinion Research (January 1992 to March 2014). Each plotted circle represents a single poll result, the percentage of respondents who report (under direct questioning) supporting laws that protect gays and lesbians against employment discrimination. The time trend is measured using a lowess curve. The solid square is the weighted percentage of untreated respondents from our CCES module who directly report supporting employment nondiscrimination.

has been quite high throughout, with the most recent surveys indicating that a large supermajority—over 70percent—favors such laws. The high level of support suggests greater social pressure to conform to the pro-gay policy position. This is reflected in the rhetoric of elites—mainstream elected officials and candidates for office rarely suggest that individuals should be fired on the basis of their sexual orientation. Indeed, as Brewer (2008, 37) suggests in his book about public opinion and gay rights, "On some policies, such as employment nondiscrimination and gays in the military, support for gay rights has approached the near-consensus levels attained by support for the principle of racial equality." This makes employment nondiscrimination an ideal area for evaluating the robustness of our findings.

The employment list experiment was embedded in the 2011 CCES. Unlike our same-sex marriage experiment, we employ a design in which participants were randomly divided into three (as opposed to two) groups: (1) the control group, consisting of 592 respondents, each of whom received a list that included only the first four (i.e., the nonsensitive) items; (2) the treatment group, consisting of 595 individuals, each of whom received the full list; and (3) a group of

608 respondents who were not given either list but were simply asked directly whether they favor or oppose such laws.[20] The three-group approach is similar to that of Gilens, Sniderman, and Kuklinski (1998) and was the convention at the time our survey went into the field. Unfortunately, this design limits our ability to directly test for the presence of social desirability bias among subgroups of respondents (as we did above) and reduces our overall sample size. That being said, the experiment still provides us with the necessary leverage to test for the presence of social desirability in the overall population.

Once again we use a nationally representative sample and unobtrusively embed our list experiment within a larger survey module that asks respondents their opinions on a variety of public policy matters. To draw less attention to our sensitive item, we include some nonsensitive items that are also political in nature. We were careful not to include too many high- or low-prevalence items to avoid ceiling and floor effects. The full text is as follows:

Employment Non-Discrimination List:
Please take your time and tell us *how many* of the following you support. We do not need to know which ones, just *how many.*

- A law requiring seat belts while driving
- Professional athletes getting million dollar-plus salaries
- Keeping a large number of troops in Afghanistan
- An amendment to the federal constitution requiring a balanced budget
- *A law protecting gays and lesbian against employment discrimination*

Table 6 summarizes the data. The distribution of responses indicates that there is unlikely to be much in the way of either large ceiling or floor effects. The

Table 6. Observed Data

	Control group		Treatment group	
Response value	Frequency	Percentage	Frequency	Percentage
0	89	15.03	23	3.87
1	197	33.28	128	21.51
2	217	36.66	222	37.31
3	77	13.01	159	26.72
4	12	2.03	46	7.73
5			17	2.86
Total	592		595	

NOTE.—This table displays the number and percentage of respondents for each value of *Y*, the number of items that the respondent supports in the list experiment, for both the control and treatment groups. Percentages may not add up to 100 due to rounding.

20. The nonresponse rate was 0.03 percent.

modal response in the control group is 2 (as desired), with just over 2 percent of respondents supporting all of the items and approximately 15 percent supporting none. The Blair and Imai (2012) test for design effects produces a p-value of 0.18, meaning that we cannot reject the null hypothesis of no design effect. Finally, there is balance among demographic covariates across treatment and control.

Table 7 presents the mean number of items supported by the control and treatment groups, the difference between the two, and the mean response to our direct question (note that the direct question was not asked to individuals in either the control or treatment group). The weighted estimate is 69.5 percent, with a 95 percent confidence interval bounded by 51.9 and 87.1 percent.[21] Even with the large confidence interval, we can safely conclude that a majority supports employment nondiscrimination laws for gays and lesbians.[22]

These estimates show no overall social desirability bias. The list experiment estimate of opinion is similar to that obtained from the direct question, and both direct and indirect estimates of support are consistent with those from recent national polls on the topic (see figure 3). This serves as additional evidence against the presence of social desirability bias.

Conclusion

It is natural to ask whether some of the purported increase in support for same-sex marriage is due to the presence of social desirability bias. Our list experiment in the 2013 Cooperative Congressional Study reveals no evidence of social desirability bias at the population level. Our list experiment measure of support for same-sex marriage (59 percent) is almost identical to the estimate we obtain from direct survey response. These estimates also match those returned in other recent national surveys. If social desirability in polling on

Table 7. Estimated Support for Employment Nondiscrimination

	Control mean	Treatment mean	Difference-in-means	Direct question
Sample	1.54 (.04)	2.22 (.04)	0.68 (.06)***	0.66 (.02)
Weighted mean	1.51 (.07)	2.21 (.06)	0.70 (.09)***	0.67 (.03)
N	592	595		608

NOTE.—The numbers in the parentheses are the standard errors. The reported difference-in-means may not equal the difference between the control mean and the treatment mean due to rounding. ***$p < .001$

21. For the analysis applying weights, linearized standard errors were calculated using the *svy* command in Stata.

22. The unweighted list experiment estimate of support for laws that protect gays and lesbians against employment discrimination is 67.8 percent.

same-sex marriage exists, our results indicate that it pushes in both directions. This experiment also provides new evidence that there exists majority national support for extending marriage rights to same-sex couples.

In our second list experiment, on the adoption of laws that make employment discrimination against gays and lesbians illegal, we again find no evidence of bias at the population level. In tandem, these list experiments cast serious doubt on claims that social desirability bias is plaguing estimates of supports of gay rights policies. This finding should strengthen social science work that has relied upon national surveys to study the opinion-policy relationship in this issue area (Lax and Phillips 2009; Krimmel, Lax, and Phillips 2012).

It is important to note that our results directly speak to only one mode of polling—surveys that are taken over the computer in the privacy of a respondent's home or office, where one might expect to observe the lowest levels of social desirability bias. Indeed, our results are consistent with work suggesting that the prevalence of social desirability bias in computer surveys should be low.[23] Furthermore, because there do not appear to be many differences between measures of public support for gay rights obtained via Internet and telephone polling,[24] we suspect that there is little if any social desirability bias in telephone polling on same–sex marriage as well.[25] This comports with recent research that suggests that differences in survey mode may not produce different levels of social desirability bias (Ansolabehere and Schaffner 2014).

References

Ansolabehere, Stephen, and Brian F. Schaffner. 2014. "Does Survey Mode Still Matter? Findings from a 2010 Multi-Mode Comparison." *Political Analysis* 22:285–303.

Aronow, Peter M., Alexander Coppock, Forrest W. Crawford, and Donald P. Green. 2015. "Combining List Experiment and Direct Question Estimates of Sensitive Behavior Prevalence." *Journal of Survey Statistics and Methodology* 3:43–66.

Berinsky, Adam J. 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* 43:1209–1230.

———. 2004. *Silent Voices: Opinion Polls and Political Participation in America*. Princeton, NJ: Princeton University Press.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20:47–77.

Brewer, Paul R. 2008. *Value War: Public Opinion and the Politics of Gay Rights*. Lanham, MD: Rowman and Littlefield Publishers.

Brooks, Clem, and Jeff Manza. 2007. *Why Welfare States Persist: The Importance of Public Opinion in Democracies*. Chicago: University of Chicago Press.

23. Though such bias has previously been uncovered in these types of surveys (Goldman 2008; Coffman, Coffman, and Ericson 2013).

24. Note that our CCES estimate of support for same-sex marriage (obtained via a direct question) falls nicely in line with the results of recent phone surveys (see figure 1).

25. If there is social desirability bias in telephone polling but not Internet polling, we would expect the levels of support for same-sex marriage and employment nondiscrimination to be notably higher in telephone polls.

Burstein, Paul. 2003. "The Impact of Public Opinion on Public Policy: A Review and an Agenda." *Political Research Quarterly* 56:29–40.

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson. 2013. "The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment Are Substantially Underestimated." NBER Working Paper No. 19508.

Corstange, Daniel. 2009. "Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT." *Political Analysis* 17:45–63.

Dalton, Dan R., James C. Wimbush, and Catherine M. Daily. 1994. "Using the Unmatched Count Technique (UCT) to Estimate Base Rates for Sensitive Behavior." *Personnel Psychology* 47:817–29.

Egan, Patrick. 2008. "Findings from a Decade of Polling on Ballot Measures Regarding the Legal Status of Same-Sex Couples." Working Paper, New York University.

Emerson, Michael O., and Laura J. Essenburg. 2013. *"What Is Marriage? Americans Dividing."* Kinder Institute for Urban Research.

Finkel, Steven E., Thomas M. Guterbock, and Marian J. Borg. 1991. "Race-of-Interviewer Effects in Pre-Election Polls: Virginia 1989." *Public Opinion Quarterly* 55:313–30.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W. W. Norton and Company.

Gilens, Martin, Paul M. Sniderman, and James H. Kuklinski. 1998. "Affirmative Action and the Politics of Realignment." *British Journal of Political Science* 28:159–83.

Glynn, Adam N. 2013. "What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment." *Public Opinion Quarterly* 77:159–72.

Goldman, Seth. 2008. "Effects of Social Desirability on Expressions of Anti-Gay Attitudes: Revealing the Role of Education in Reducing Prejudice." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL, USA.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56:202–17.

Hadaway, C. Kirk, Penny Long Marler, and Mark Chaves. 1993. "What the Polls Don't Show: A Closer Look at US Church Attendance." *American Sociological Review* 58:741–52.

Holbrook, Allyson L., and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74:37–67.

Imai, Kosuke. 2011. "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association* 106:407–16.

Krimmel, Katherine L., Jeffrey R. Lax, and Justin H. Phillips. 2012. "Public Opinion and Gay Rights: Do Members of Congress Follow Their Constituents' Preferences?" Working Paper.

Kuklinski, James H., Michael D. Cobb, and Martin Gilens. 1997. "Racial Attitudes and the 'New South.'" *Journal of Politics* 59:323–49.

LaBrie, Joseph W., and Mitchell Earleywine. 2000. "Sexual Risk Behaviors and Alcohol: Higher Base Rates Revealed Using the Unmatched-Count Technique." *Journal of Sex Research* 37:321–26.

Lax, Jeffrey R., and Justin H. Phillips. 2009. "Gay Rights in the States: Public Opinion and Policy Responsiveness." *American Political Science Review* 103:367–86.

Lewis, Gregory B., and Seong Soo Oh. 2008. "Public Opinion and State Action on Same-Sex Marriage." *State and Local Government Review* 40:42–53.

Lyall, Jason, Graeme Blair, and Kosuke Imai. 2013. "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107:679–705.

Page, Benjamin I., and Robert Y. Shapiro. 1983. "Effects of Public Opinion on Policy." *American Political Science Review* 77:175–90.

Pinello, Daniel R. 2006. *America's Struggle for Same-Sex Marriage*. New York: Cambridge University Press.

Powell, Richard J. 2013. "Social Desirability Bias in Polling on Same-Sex Marriage Ballot Measures." *American Politics Research* 41:1052–1070.

Presser, Stanley. 1990. "Can Changes in Context Reduce Vote Overreporting in Surveys?" *Public Opinion Quarterly* 54:586–93.

Regnerus, Mark. 2013. "Right Side of History, or Primed to Say Yes? When Asking about Same-Sex Marriage, Polling Methods Matter." *National Review Online*. http://www.nationalreview.com/article/356220/right-side-history-or-primed-say-yes-mark-regnerus.

Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80:613–24.

Smith, Tom W. 1998. "A Review of Church Attendance Measures." *American Sociological Review* 63:131–36.

Streb, Matthew J., Barbara Burrell, Brian Frederick, and Michael A. Genovese. 2008. "Social Desirability Effects and Support for a Female American President." *Public Opinion Quarterly* 72:76–89.

Tourangeau, Roger, and Ting Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133:859–83.

Tsuchiya, Takahiro, Yoko Hirai, and Shigeru Ono. 2007. "A Study of the Properties of the Item Count Technique." *Public Opinion Quarterly* 71:253–72.

Vavreck, Lynn, and Douglas Rivers. 2008. "The 2006 Cooperative Congressional Election Study." *Journal of Elections, Public Opinion, and Parties* 18:355–66.