HOW ACCURATE ARE SURVEYED PREFERENCES FOR PUBLIC POLICIES? EVIDENCE FROM A UNIQUE INSTITUTIONAL SETUP

Patricia Funk*

Abstract—Opinion polls and other surveys are used to capture public sentiments on a variety of issues. If citizens are unwilling to reveal certain policy preferences to others, surveys may fail to characterize population preferences accurately. The innovation of this paper is to use unique data that allow one to measure biases in surveyed preferences for a broad range of public policies. I combine data on 184 referenda held in Switzerland between 1987 and 2007 with postballot surveys that ask how the citizens voted for each proposal. The difference between stated preferences in the survey and revealed preferences at the ballot box provides a direct measure of survey bias. I find that these biases vary by policy area, with the largest occurring in policies on immigration, international integration, and votes involving liberal or conservative attitudes. Also, citizens show a tendency to respond in accordance with the majority.

I. Introduction

Researchers interested in eliciting citizens’ values, attitudes, or policy preferences often rely on surveys. Likewise, opinion polls are the main vehicle for extracting information on voters’ preferences in representative democracies (Besley & Coate, 2008). But how accurate are surveyed preferences for public policies? Unfortunately, the only information available is that expressed in the surveys (privately held opinions are unobservable), which makes it impossible to assess the truthfulness of its content. This lack of knowledge has led researchers, particularly in economics, to be wary about the use of subjective survey data (Bertrand & Mullainathan, 2001; Olken, 2009). Nevertheless, there are important questions on attitudes and preferences for specific policies that can only be answered with the help of survey data (Alesina & Fuchs-Schündeln, 2007; Alesina, Giuliano, & Nunn, 2013). In this light, it seems particularly useful if we could somehow characterize the accuracy of surveyed preferences.

The innovation of this paper is to use unique data that allow measuring the accuracy of opinion polls for various types of public policies. The data come from Switzerland, the world leader in the use of direct democracy. In Switzerland, citizens vote on all major policies through referendum. The results of these ballots provide a measure of revealed preferences for policies with which survey responses can be compared. Starting in 1987, telephone surveys have been conducted after each federal vote, covering samples of roughly 1,000 eligible voters (these surveys are called VOX survey).1 The survey is conducted two or three weeks after the vote and aims to gather information about the respondents’ voting behavior, with questions ranging from whether and how the respondent voted, information on political views, knowledge about the ballot, the perceived importance of the ballot, and various sociodemographic characteristics.

The key idea of the paper is to compare, for each vote, the approval stated in the survey (= share yes of all respondents who indicate they voted and reveal their result) with true approval as given by the voting result (real share yes per ballot).2 This difference between stated and revealed voter preferences provides a direct measure of survey bias for a broad range of policy issues. Since Swiss citizens are asked how they voted shortly after the vote, it gives rise to a clean experiment; in contrast to pre-polls, there is no incentive for respondents to answer strategically (unless the goal is to hide their true policy preferences), and the question does not allow for changing preferences over time.3 Therefore, any difference in approval between the vote and the postvote survey must be caused by either differences in the population of voters and survey respondents who declare they have voted or citizens misreporting their preferences in surveys.

Information on revealed preferences for public policies is absent in representative democracies because citizens do not vote on public policies directly. Even in representative democracies with direct democratic elements at the subfederal level (e.g., California in the United States), the number of votes held is typically low, covering a small set of policies. Furthermore, exit-poll data are not publicly available on a large scale, which makes a more refined analysis impossible. In Switzerland, all individual-level postvote survey data are publicly available. This will allow me to shed some light on the determinants of the survey bias.

Received for publication August 10, 2015. Editor: Amitabh Chandra.

* Università della Svizzera Italiana and Universitat Pompeu Fabra.

I thank the editor and one very helpful anonymous referee for excellent comments. Further thanks go to Larbi Alaoui, Antonio Ciccone, Rajeev Dehejia, Gabrielle Fack, Claudio Ferraz, Stephan Litschig, Torsten Persson, Alessandro Tarozzi, Joachim Voth, conference participants at the EEA meeting 2011, and seminar participants at Universitat Pompeu Fabra, University of Zurich, University of Oslo, BI Norwegian Business School, DIW Berlin, VATT Helsinki, and University of Basel for useful comments and discussions. I also thank Lukas Golder from the Gfs for his cooperation. Financial support from the Barcelona GSE Research Network and the Swiss National Science Foundation (Sinergia grant 130648) is gratefully acknowledged.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00585.


© 2016 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology
doi:10.1162/REST_a_00585

1The samples are selected by random sampling based on the telephone book. Response rates to the survey fluctuate between 28% and 48%, and are slightly higher than the average response rate of similarly conducted telephone surveys by major news media in the United States. Holbrook, Kronsick, and Pfent (2007) analyze 114 telephone studies that were conducted by major American news media (e.g., ABC News, New York Times, Gallup, Washington Post) between 1996 and 2005 and find an average response rate of 30%.

2The survey gives the options “yes,” “no,” “no answer,” and “don’t know” for the voting result. The “share yes” is calculated as the sum of yes votes over the sum of yes and no votes, to match the definition of the ballot results.

3The survey asks clearly: “How did you vote in the ballot on topic X?” Therefore, even if preferences change, it will not affect the answer. In contrast, a difference between the official voting result and answers in pre-polls can arise because citizens strategically misrepresent their preferences in order to make other citizens change their votes or citizens, after learning the polling results, may change their minds on how to vote.
The main interest of the paper lies in quantifying the extent of the survey bias and relating it to the policy area of the vote. In Switzerland, ballots are regularly held covering all policy areas relevant in a mature democracy. Swiss citizens have voted on issues such as immigration, environmental protection, health, unemployment benefits, agriculture, the military, and various regulatory measures. I hypothesize that the survey accuracy depends on the topic, as citizens may be unwilling to reveal politically incorrect views (Morris, 2001; Loury, 2004). Especially on socially sensitive issues (e.g., immigration), citizens with politically incorrect views (e.g., anti-immigration) may choose not to respond to surveys or to respond but lie in their responses. Both channels lead to a gap between stated and revealed voter preferences. Testing for potential differences in survey accuracy with respect to the policy area is the main contribution of this paper.

The Swiss data allow measurement of survey accuracy for 184 federal votes. The survey bias (defined as the vote-specific difference between the reported and effective share yes) is 4.7 percentage points on average. For roughly one-third of the votes, this survey bias is statistically significant at conventional levels. For the rest of the votes, there is no statistically significant survey bias at all. The amount of survey bias varies quite substantially across the votes. For instance, the vote with the biggest difference between stated and true preferences concerned a proposed law change to improve gay rights, pro-environment. I explicitly test the hypothesis that votes supported by the left-wing party have a higher share yes in the survey (relative to the ballot approval) compared to the votes where the left-wing party recommended a no. The data strongly support the existence of a “liberal bias.”

Historically, politically correct views had a clear left-wing connotation (e.g., pro–gender equality, against racism, pro–gay rights, pro–environment). I explicitly test the hypothesis that votes supported by the left-wing party have a higher expressed share yes in the survey (relative to the ballot result) compared to the votes where the left-wing party recommended a no. The data strongly support the existence of a “liberal bias.”

Votes supported by the left-wing party had too high a share of yes votes in the survey, whereas votes where the left-wing party recommended a no displayed too high a share of no votes in the survey. The difference in survey bias (= stated approval in the survey minus real approval at the ballot) between votes that were and were not supported by the left-wing party is 5 percentage points and statistically significant. This liberal bias persists when the individual survey data are reweighted to correct for oversampling of observable individual characteristics (including self-reported party affiliation). Therefore, selection on observables is unlikely to be the driver behind this bias. Either citizens choose to respond to the survey based on unobservables such as their privately held policy preferences or they falsify their preferences in the survey. From a policy perspective, disentangling these two channels is not that essential, since in either case, there is not much the researcher can do to eliminate the bias in polls. Nevertheless, the data provide evidence that a portion of the people falsify their preferences in surveys. I compare votes held on the same day (leaving the respondent pool fixed), which were perceived to be of either high or low importance for the Swiss nation as a whole. I hypothesize that for votes of low importance, there is little pressure to lie. If so, the liberal bias should be larger in votes that were salient and regarded as being of high importance. In line with this intuition, the data show that the liberal bias is present only for votes of high importance. This again reconfirms the view that surveys are sometimes biased.

Next to a desire to be politically correct, citizens may want to conform to others (Bernheim, 1994; Kuran, 1995). Since the voting result is known at the time of the survey, a natural guess is that some of the respondents answer in line with the majority view. In fact, comparing how the same individual reports having voted on various (simultaneously held) votes for which his or her favored party recommended a no, the answer is more likely to be a yes when the proposal was accepted (this result holds for proposals accepted or rejected by a narrow margin, as well as a restricted set of proposals where parties’ ex ante support resembles voters’ support at the ballot). This suggests that some of the survey-respondents misreport their preferences due to either conformity or a desire to be on the winning side.

Clearly, I expect these behavioral biases (a desire to appear politically correct or to be on the winning side) to hold true outside the Swiss context. When it comes to responding to surveys or opinion polls, the first type of bias is most relevant. The following complementary evidence for the United States is at least consistent with the existence of a political correctness bias in other settings: pre-polls on gay rights have been particularly inaccurate in predicting later voting outcomes (with support in the pre-polls being too high; Egan, 2010, and Powell, 2013), and support for black candidates has also been overstated in surveys in the past (Hopkins, 2009).

The data at hand also allow me to investigate whether the accuracy of surveys differs according to culture, religion, or economic development of a geographic unit. Switzerland is a diverse country, with differences in languages (German-,

---

4 I call this left-wing bias “liberal bias” and refer to various societal attitudes rather than preferences for economic liberalization.

5 Note that the data rule out lying on turnout as the driver behind the liberal bias. If nonvoters were more liberal than voters and said they had voted in the survey, then a higher share of nonvoters in the survey should lead to a larger liberal bias. However, the interaction term between the (measurable and ballot-specific) share of overrepresentation of voters in the survey times the dummy “Recommendation Left” is statistically insignificant.

6 The respondent sample is constant in this case because the survey asks the respondent how he or she voted on all votes that were held on a given day (see section II).
Italian-, and French-speaking areas), religions (nearly an equal share of Protestants and Catholics), and economic opportunities. It turns out that culture (measured by a canton’s language) and population size matter the most for cross-sectional differences in survey accuracy: cantons of German language display significantly higher survey biases than the other cantons (French or Italian speaking), and larger cantons have smaller biases than smaller cantons. Religion and income do not matter once language is accounted for.

The paper relates to various strands of the literature. First, it is relevant for a growing economic literature based on survey data. Although economists have traditionally been skeptical with regard to surveys on attitudes and preferences (Bertrand & Mullainathan, 2001), there has been a recent surge of influential papers explaining certain types of attitudes and preferences (Fong, 2001; Guiso, Sapienza, & Zingales, 2003; Alesina & Fuchs-Schündeln, 2007). This paper shows that skepticism toward surveyed preferences is justified in certain policy areas (e.g. racial attitudes, attitudes on gender equality) but less so in others (preferences for direct democracy, health, or federal finances). Furthermore, the data allow one to assess how innocent it is to compare survey responses across cultures and religions. Second, the paper relates to a growing literature investigating the consequences of social pressure and image concerns (Benabou & Tirole, 2006; Tadelis, 2011). So far, various studies have established that image concerns matter for voter participation (Gerber, Green, & Larimer, 2008; Funk, 2010), contributions to charity (DellaVigna, List, & Malmendier, 2012), or worker effort (Mas & Moretti, 2009). This paper documents that opinion polls are particularly biased on topics with a predominant politically correct view, which is consistent with citizens caring about their image. Third, the paper complements a strand of papers (mostly in political science) that analyze the accuracy of polls in elections (Baretto et al., 2006; Stromberg, 2008; Hopkins, 2009). The key addition to these papers is to add knowledge on the accuracy of polls on issues. The setting of a direct democracy gives rise to a measure of true preferences (as revealed at the ballot box) with which survey responses can be compared. Fourth, a related literature in political science explores the role of item nonresponse in the survey quality of a given respondent sample (Berinsky, 1999, 2004). The data at hand allow me to go one step further by contrasting survey responses to the true underlying preferences of the voting population. In line with Berinsky (1999, 2004), I find poor survey quality on issues involving race. Finally, the paper relates to a strand of laboratory experiments that investigate the nature of lying (Gneezy, 2005; Lundquist, Ellingsen, & Johannesson, 2009; Erat & Gneezy, 2011). This paper suggests that for some policy areas, citizens prefer to hide their true opinion, even if it is merely a survey conducted by telephone.

The rest of this paper is structured as follows. Section II describes the data. Section III characterizes the major survey biases, and section IV sheds light on possible channels. Section V analyzes cantonal differences in survey accuracy, and section VI concludes.

II. Data

A. Official Voting Results

For Swiss citizens, having a voice in politics is almost daily business. Switzerland has a long tradition of direct democracy; at the federal level alone, citizens have voted more than 300 ballots in the past fifty years. Citizens can propose an initiative for a partial or total revision of the federal constitution. In addition, they can request a referendum on all laws issued by the federal government if 50,000 signatures are collected. Moreover, a voter referendum is mandatory for any changes to the constitution and all international treaties. As a consequence, citizens vote on federal ballots several times each year. In Switzerland, every citizen over 18 is allowed to vote (before March 1991, the minimum age was 21). Eligible voters receive balloting documents by mail at home. These documents include all relevant information on the ballots (there are usually several ballots bundled for a given voting day), such as the precise questions, the arguments for and against the propositions, a printed version of the parliamentary debates (if any), and often outside opinions from interest groups. Hence, Swiss citizens have easy access to information about the ballots through the distributed documentation and discussions in the media. Returning the ballot is also relatively convenient. In contrast to the United States, no registration to vote is necessary at all. Since 1995, voters have been granted the option of voting by mail, in addition to at the voting booth. Voter turnout in the past twenty years is 42% on average, with some variation depending on the topic. The web page of the federal authorities (http://www.admin.ch) lists all federal votes ever held. Information on the votes includes the title, the date, the number of eligible voters, the number of effective voters, the number of valid votes, the number of blank votes, the number of yes votes, and the number of no votes. The share yes votes are calculated as the number of yes votes as a proportion of the total number of valid nonblank votes, and the share no votes are calculated as the number of no votes in proportion of the total number of valid (non blank) votes. The share yes and share no add up to 100%. The main variable of interest is the approval percentage for each vote measured as the share yes, which will be compared with the stated approval percentage in the VOX survey.

B. The Postelection Surveys (VOX Surveys)

Since 1981 VOX surveys have been conducted after each federal vote. These surveys are conducted with samples of roughly 1,000 eligible voters (700 voters until 1987) and take place during the two or three weeks following the vote.
As described in the technical documentation on the VOX surveys, the basis for selecting households is the Swiss telephone book. A random sample stratified by language area (German speaking, French speaking, and Italian speaking) is applied, and households are contacted until roughly 1,000 respondents have been gained. Response rates fluctuate between 25% and 48% for the surveys conducted between 1998 and 2007.8

The main objective of these postelection surveys is to understand the motives underlying the individual voting decision and possible connections with individuals’ knowledge of the issues. Most relevant for this study, the VOX survey asks about participation in the most recent federal vote and participants’ voting decisions. The exact procedure of the survey is the following. The interviewer starts by introducing himself or herself and asks whether there is an eligible voter in the household. The exact wording is, “Hello, I am an employee of the GFS Research Institute, Bern. We are conducting a survey on the last federal votes. The survey is anonymous. May I ask whether there is at least one household member older than 18? In the case of several persons, which one has his or her birthday the earliest in the year? May I conduct an interview with that person?” If a survey respondent is recruited, the survey asks about voter participation and voter decisions in the following way: For participation, the question is, “Was it possible for you to participate in the federal vote on the date [DD.MM.YYYY]?” Responses are yes, no, don’t know, no answer. As for voter decisions, the precise question is: “How did you vote on the federal ballot [title X]?” Possible answers are yes, no, blank, don’t know, no answer. All votes that were bundled on a given voting day are included. Hence, a respondent gives answers for all these votes. Apart from these questions directly related to the vote, the survey also asks about various aspects relating to the voting decision, such as knowledgeability on the topic, types of media consulted prior to the decision, or the perceived importance of the vote. An extensive set of questions aimed at gathering individual characteristics (e.g., age, education, marital status, profession) completes the questionnaire.

C. Definition Survey Bias

To compare approval in the survey for a certain vote with that revealed at the ballot box, I first define the share yes in the survey in an equivalent manner to the share yes of the voting result. This means taking the number of yes votes (from citizens who indicate they voted) divided by the sum of yes and no votes. The key variable of interest is the difference between the share yes in the survey and the share yes from the official voting outcome, which I define as survey bias. A positive survey bias indicates that the approval stated in the survey is larger than the official one, and a negative survey bias indicates the opposite. I start by comparing official voting outcomes with stated voting outcomes for all votes (initiatives and referenda) held in 1987 or later, where the VOX survey had a sample size of roughly 1,000 citizens. The latest available data were VOX surveys conducted in 2007, which gives a sample of 187 votes in total, spanning all relevant policy areas over the twenty years. Since three votes (numbers 462, 463, 464) have an identical reported share yes, I drop these votes due to high likelihood of error. That leaves a sample of 184 valid votes.

III. The Survey Biases

A. Descriptive Evidence

To get a sense of the magnitude of the survey biases for the 184 votes, figure 1 displays the kernel density. As can be seen, the reported share yes is slightly bigger than the actual share yes, with a wide variation across different votes (the maximum difference between reported and real share yes is nearly 20 percentage points).

What are the possible sources for these biases? First, the sample of survey respondents may differ from the population sample (in terms of observables and unobservables). Second, for a given respondent sample, a certain share of respondents may not be willing to reveal their vote (in the survey, such an individual would answer “blank,” “don’t know” or “no answer” even though he voted yes or no). Third, survey respondents who answer yes or no as their voting outcome falsify either their policy preferences or their voting decision (i.e., nonvoters declare they have voted).

The first possible explanation for survey bias is having a selected respondent sample. To investigate this possibility, I compare the respondents’ characteristics in terms of age, gender, religion, language, and education with a representative sample of the Swiss population. Note that here, the right comparison is between all survey respondents (voters and nonvoters) and the Swiss population. Information on the latter can be gained by using existing data on a random

---

8Technical reports are not officially available for the earlier votes.
sample of 5% of the Swiss citizenry, the Public Use Sample (PUS), compiled by the Swiss Federal Office of Statistics for various years. To ensure anonymity, the PUS uses age classes of the respondents.9 Subsequently, I focus on individuals of 20 years of age or more in both the VOX surveys and the PUS data.

As shown in table 1, VOX respondents are quite similar to the random PUS sample of Swiss citizens. In 1990, the share of Protestants and the share of highly educated in the survey are slightly higher than in the population counterpart, and in the year 2000, the share of elderly people is additionally overrepresented at the cost of the younger people. Overall, however, the differences in average characteristics between survey respondents and the population sample are small. To assess the role of sample selection in generating the observed survey biases, I reweight all the survey samples to match the population precisely on religion, age above 60, and higher education (where the highest deviations were found).10 Note that reweighting on a few observables is the standard procedure to correct for sample selection in opinion polls and surveys. As a rule of thumb, the recommendation is to reweight if the difference between a specific survey characteristic (e.g., share of elderly) and population characteristic is more than 5 percentage points (DeBell & Krosnick, 2009). As can be seen from figure 2, the survey bias becomes reduced somewhat if reweighted samples are used, but there is still considerable variance.

Taking as a starting point the survey bias left after reweighting the surveys, I can investigate whether there is a significant relationship between survey bias and survey response rate, as could be expected if there was selection on unobservables. The first column of table 2, however, shows no such relationship. One could imagine that people with certain views (e.g., politically incorrect views) respond less frequently to surveys. Or the opposite may be the case for people with a politically correct view and offset the effect on response rates. Next, for a given respondent sample, a bias may be created by respondents hiding their vote with a "no answer," "blank," or "don’t know." If so, the higher the share of voters who reveal their voting decisions, the lower the bias should be. The second column of table 2 shows that this is indeed the case, although the variation in vote reveal rates cannot explain a large part of the variation in survey bias (the R² including the explanatory power of the fixed effects is 0.22, and the R² “within” is a mere 0.058). Finally, misrepresentation of voter participation may create survey bias. Nonvoters may claim to have voted and create a survey bias if their preferences are different from those of genuine voters.11 If voters are more likely to respond to surveys (but there is no lying on turnout and policy preferences),

10 I have to reweight the complete survey (including voters and nonvoters) since I have information on the characteristics of the whole population but not the voters alone. Lack of this information is no major drawback, however, since reweighting the whole survey sample is sufficient to eliminate selection on observables. Assume, as an example, that the population consists of 50% young and 50% old, but since elderly people are more likely to respond to surveys, the share of elderly in the surveys is 80% (one can easily allow this share to vary by survey). Reweighting the whole survey ensures that the survey sample corresponds to the population in age: 50% are old and 50% are young. If on one ballot, mostly young people voted (e.g., the share of elderly voters is 20%), this share will be accurately depicted when analyzing approval of declared voters in the reweighted survey sample.
11 León (2013), however, provides evidence that voters and nonvoters have comparable preferences on average.
there should be no relationship between overrepresentation of voters in the surveys and the magnitude of the survey bias (since the survey bias is calculated for survey respondents who declare they voted). Since overrepresentation of voters in a given survey is directly measurable (the share of survey respondents who declare they voted minus real turnout, per ballot), I can easily investigate whether there is a significant effect of overrepresentation of voters on survey bias. The answer is no (see column 4). This preliminary evidence suggests that neither variation in the survey composition of observable individual characteristics nor differences in overrepresentation of voters, item nonresponse or, overall response rates to the surveys can explain a significant part of the variation in survey biases. I suspect that the topic of the vote matters and investigate this in the next section.

B. Survey Bias and Topic of the Vote

This section tests the main hypothesis of the paper, according to which survey biases should be larger in votes with a predominant politically correct view (e.g., race, gender).

To illustrate the votes with particularly large survey biases, appendix table 2 (online) reports the 184 valid votes, sorted by amount of survey bias. The table also indicates the year of the vote, the VOX number, the title of the proposition, and the number of survey respondents who declared they voted and reveal their voting decision (# Obs.). Next to displaying the amount of survey bias, I analyze which part could have been generated by sampling error. For this, I calculate (based on the survey sample share yes of voters) the predicted mean share yes and its confidence intervals for the population of voters, taking into account the relevant sampling design (random sampling stratified by language area). Based on the 95% confidence intervals for the predicted mean share yes, I calculate confidence intervals for the survey bias as true share yes (at the ballot box) minus the lower or upper bound of the confidence intervals. If the confidence intervals include 0, the survey bias could be generated by sampling error. Likewise, I report p-values indicating whether, based on the (survey) sample share yes of self-declared voters, the null hypothesis (mean share yes for the population of voters equals the known true share yes as given by the voting result) can be rejected. As can be seen from appendix table 2 (online), the null is rejected for roughly one-third of the votes at standard levels of significance. To give the reader an intuitive grasp of whether the votes with the highest survey biases have particularly low survey response rates, low vote reveal rates, or a high overrepresentation of voters, I report the numbers for these last three variables as well. In line with table 2, there does not seem to be a strong relationship with any of those. What are the topics of the votes with the largest gaps? From visual inspection, it looks as if there are quite a few votes in the areas of environment, immigration, and redistribution where the differences in stated and real ballot outcomes are high. To make a more systematic comparison, I define twelve broad policy areas that are relevant beyond the Swiss context: international integration, immigration, military, protection of environment, nuclear energy, federal budget, direct democracy, health, redistribution, retirement age, gender equality and liberal attitudes. I selected all votes in a policy area that had either the same or the opposite goal (e.g., either to facilitate immigration or make it more difficult; either increase or decrease the size of the Swiss army) Since the framing of a ballot may matter (see Buettler & Marechal, 2007), I display the votes with opposing goals separately. Appendix table 3 (online) shows the selection of the votes per policy area. Overall, 90 votes could be assigned to these twelve policy areas, which leaves another 94 unclassified votes. Every selection process is subjective to a certain degree. To make this process as transparent as possible, I describe the goal of all 184 the votes (including the unclassified ones) in the online appendix.

12 From research on elections, it is known that there are usually more voters in the survey samples than in the share of voters in the population (Karp & Brockington, 2005; Holbrook & Kronick, 2010). The same pattern is found for current surveys, where the share of voters in the survey is on average 15 percentage points higher than official voter turnout. However, as can be seen from the summary statistics in appendix table 1 (online), there is substantial variation in the difference between reported and real turnout across surveys.

13 At the 1% level of significance, I can reject the null for 51 votes, and at the 5% level for 60 votes.
Subsequently, I test whether the survey bias differs by policy area. The model I estimate is

\[ Y_{ijt} = \beta_j \times D_j + \epsilon_{ijt}. \] (1)

The dependent variable is the Survey-Bias (the difference between the stated share yes of voters in the survey and the effective share yes of the voting result) per vote \( i \) that falls into policy area \( j \) and was voted on in year \( t. \) \( D_j \) is a dummy for each of the twelve policy areas. Standard errors are clustered at the voting-day level to account for possible correlation of errors within a given survey sample.

The first column of table 3 shows the differences in survey biases by policy area. As can be seen, the survey bias is positive and relatively high for votes aiming at facilitating immigration (12 percentage points), fostering international integration (5.6 percentage points), against nuclear energy (5.2 percentage points), for the protection of the environment (3.6 percentage points), for gender equality (6.7 percentage points), or two votes involving a liberal attitude (8.5 percentage points). These last two votes with very high survey biases were giving more rights to homosexual couples and liberalizing the sexual rights of teenagers. High negative survey biases are found for votes that attempt to restrict immigration (−5.1 percentage points). Here, the share of respondents who admit having voted for tighter immigration laws is lower in the survey than the ballot box. Finally, note also that there are some policy areas (health, retirement age, direct democracy, federal finances) with no significant gaps between survey approval and real approval. Unsurprisingly, I can reject the null hypothesis that the gaps are the same across policy areas at the 1% significance level.

What are the reasons for these survey biases? The main hypothesis is that people with certain preferences (e.g., against immigration) do not respond to the survey or that they do respond but lie. However, it could also be the case that overrepresentation of individuals with certain observable characteristics explains these gaps. As shown in table 1, people may deliberately lie or unconsciously misrepresent their preferences. See Bazerman and Gino (2012) for various examples of and explanations for unintentional dishonesty.

The dependent variable in columns 1 to 4 is the difference between a ballot’s approval in the survey (“stated share yes”) and true approval as measured at the ballot box (“real share yes”). The number in column 5 is the share of votes per policy area, where the left-wing party recommended a yes. In column 2, the surveys are reweighted to match the population on age, religion, and education. In column 3, surveys are reweighted for all effective characteristics (age, religion, (left-wing voters)) voters)

\[ \text{Table 3.—Survey Bias, by Topic} \]

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of Votes Supported by Left</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Controls</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Voting result</td>
</tr>
<tr>
<td>Number of votes</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Use of weights to calculate survey bias characteristics</td>
<td>No</td>
<td>Yes (age, religion, education)</td>
<td>Yes (left-wing voters)</td>
<td>No</td>
</tr>
<tr>
<td>Controls</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Number of votes</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Use of weights to calculate survey bias characteristics</td>
<td>No</td>
<td>Yes (age, religion, education)</td>
<td>Yes (left-wing voters)</td>
<td>No</td>
</tr>
<tr>
<td>R²</td>
<td>0.563</td>
<td>0.450</td>
<td>0.508</td>
<td>0.534</td>
</tr>
<tr>
<td>H0: Equality of estimated coefficients (P-Value):</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

14 People may deliberately lie or unconsciously misrepresent their preferences. See Bazerman and Gino (2012) for various examples of and explanations for unintentional dishonesty.
the shares of Protestants, the elderly, and the highly educated are slightly higher in the survey compared to the shares in the census. To assess the relevance of sample selection for the observed differences across policy areas, I reweight the data, as is standard in polling research. I correct for overrepresentation of the elderly, the share of Protestants, and the share of the highly educated (column 2). As can be seen from table 3, column 2, the biases from the reweighted samples are often somewhat smaller, but do not disappear. Note that this type of reweighting procedure corresponds to the classic strategy used by most of the major U.S. news media, including the New York Times, Gallup, CNN, and USA Today (Blumenthal, 2004). There, opinion polls are weighted to match the U.S. Census for gender, race, education, and usually some geographic classification. Self-reported party affiliation is typically not used for reweighting because it is subject to error itself. Since I know the true voter preferences for the 184 surveys, it is nevertheless interesting to see whether the biases disappear if data are reweighted according to self-reported party identification. It turns out that the average share of self-declared left-wing voters in the survey (29%) is higher than the average share of left-wing vote shares in parliamentary elections (21%). Either left-wing voters are more willing to respond to surveys or some (plausibly populist) right-wing voters do not indicate their favored party. Should the second explanation have some truth, then the share of self-declared left-wing voters in the survey is higher than the real share of left-wing voters in the survey. Table 3, column 3, presents the results when overrepresentation of (self-declared) left-wing voters in the surveys is corrected for. Many biases still persist, and the one on budget balance becomes even larger. Finally, I investigate the sensitivity of the results with respect to controlling for the voting decision (column 4). The magnitude of the estimated coefficients is sometimes affected, but the results remain qualitatively similar.

Summing up, the accuracy of the postballot survey differs by policy area. As shown in table 3, selection on observables does not seem to have generated the differences in survey biases across topics. Also, the policy areas with particularly high biases (international integration, immigration, nuclear energy, liberal attitudes) are not characterized by systematically low survey response or vote reveal rates (see appendix table 4 [online]). As such, it seems likely that either voters with certain (unobservable) policy preferences select into survey response or survey respondents misrepresent their voting decisions on certain topics. In either case, there is not much the researcher can do to eliminate survey bias.

More important, the analysis suggests that surveys are inaccurate for topics on international integration, immigration, gender equality, and votes involving a liberal attitude. These are the policy areas where a politically correct view is most obvious. Preferences on direct democracy, health, or the retirement age, however, appear to be less contaminated in surveys. From visual inspection of table 3, it also looks as if left-wing policies tend to have a positive gap. As shown in column 5, the policy areas with the biggest positive gaps (international integration, immigration, liberal attitudes) were largely supported by the left-wing party. That raises the question about the existence of a liberal bias.

C. Liberal Bias

To investigate systematically whether support of the left-wing party relates to the survey bias, I first define a dummy variable, Recommendation Left-Wing Party, that takes a value of 1 if the vote was supported by the left-wing party and 0 otherwise. Votes where the left-wing party made no recommendation are coded as a missing value (10 votes in total).

The first column of table 4 shows that votes that were supported by the left-wing party had a 4.9 percentage point higher survey bias than votes where the left-wing party recommended a no. As can be seen from the constant, votes where the left-wing party recommended a no had a negative survey bias (stated – true approval) of −1.7. Votes where the left-wing party recommended a yes had a positive average survey bias of 3.2 (4.9 − 1.7), which is statistically different from 0. The result that votes supported by the left-wing party have a higher approval in the survey (I call this liberal bias) is robust to adjusting for differences in the composition of survey respondents, with regard to age, religion, and education (column 2) and self-declared party-affiliation (column 3). The liberal bias does not disappear when controlling for the ballot result (column 4) or when identifying the effect from within-survey sample variation (column 5).

Clearly, this section shows that surveys are less accurate on sensitive topics and are distorted toward a left-wing view. The next section addresses the question why this is the case.

15 To be precise, I reweight the complete survey (including voters and nonvoters) to match the population counterpart on certain characteristics for each vote separately. Then I newly calculate the share yes for the citizens who indicate they voted. The correction for oversampling of one specific characteristic in Strata can be done by specifying poststrata within svy-estimation and indicating poststrata weights (see Levy & Lemeshow, 1999, for a concrete example); or if one aims at correcting oversampling of various individual characteristics, one can use a raking procedure. I perform the raking procedure with the stata command survwgt. Apart from what is shown in table 3, I also corrected for individual characteristics separately and found that correcting for education affects the estimates the most.

16 The share of self-declared left-wing voters is calculated as a percentage of all survey respondents who indicate a favored party.

17 To take advantage of the larger sample of all (classified and nonclassified) votes, I first regress the survey bias on the share yes of the ballot and then use the residuals as the dependent variable in column 4.

18 Note also that the estimated interaction terms recommendation Left-Wing Party times turnout gap, and recommendation Left-Wing Party times vote reveal rate are insignificant. Therefore, the liberal bias is unlikely to be caused by item nonresponse or nonvoters being more liberal and pretending to have voted.
IV. Mechanism

A natural interpretation for the detected liberal bias is that left-wing parties support more politically correct views (e.g., Liberal attitude, pro-environment, pro-redistribution, pro-immigration), and this causes people with politically incorrect views either not to respond to the survey or to respond and misrepresent their preferences.

What speaks against selection as being the only channel is the fact that the liberal bias persists when reweighting for self-reported party affiliation. To test more explicitly the hypothesis that some of the survey respondents falsify their preferences, I state conditions under which respondent lying is more likely to occur. Of first-order priority seems to be the importance of the vote. Votes with major consequences for the country are more salient and also more frequently discussed in the media. Admitting a politically incorrect view seems more costly in this situation. Since the VOX survey asks for the perceived importance of the vote for Switzerland as a whole, I can classify the votes according to whether they were ranked above or below the mean importance of the vote.

The left-hand side of table 5 investigates whether the liberal bias is larger in important votes. I define a dummy variable, Importance Vote, which takes a value of 1 if the vote was above the mean and 0 otherwise. As can be seen in table 5, column 1, the bias is significantly larger for important votes. Column 2 relies on variation across votes that were held on the same day. Here, the set of survey respondents is the same. Again the liberal bias is present only in salient votes, even though the estimated interaction term is insignificant. Note also that results are qualitatively similar when relying on other measures for the importance or salience of a vote, such as intensity of advertisement in newspapers prior to the vote or voter turnout.

The fact that the survey bias exists only for important votes suggests that some types of voters pretend having voted liberal when in fact they did not. Who are these voters? Intuitively, I expect a self-reported left-wing voter to be more embarrassed to admit a politically incorrect right-wing view than, for instance, a self-reported populist right-wing voter. In fact, if a survey respondent admits to vote for the populist right-wing party, he or she does not seem to bother too much about image concerns and therefore is unlikely to misrepresent preferences on referendum votes.

While it is hard to fully answer who misreports preferences as too liberal, some insights can be gained by looking at respondent behavior broken down by the importance of the vote. To rule out selection, I compare survey responses for a fixed sample of respondents, who respond on how they voted on a mix of important and unimportant votes held on the same day. Columns 3 to 9 in table 5 document that for the unimportant votes (see row 1), a self-reported left-wing voter’s probability of voting yes goes up by 25 percentage points if the left-wing party recommends a yes, while for voters with different party affiliations, this probability goes down. As left- and right-wing positions are often negatively correlated, this is a natural pattern to suspect (at least for the right-wing voters). Interestingly, though, the negative effect remains (to a smaller extent) even when own-party recommendation is controlled for (columns 5, 7, and 9). As such, having the left-wing party suggest a yes vote makes non-leftist voters less likely to vote yes in relatively unimportant votes.

What effect is to be expected for important votes? If non-left-wing voters dislike proposals supported by the left-wing party, this effect is plausibly larger when the votes are important and the stakes are higher. Finding the opposite may therefore be indicative of respondent lying. As can be seen from the second row of table 5, we observe a negative interaction term only for subjects who indicate the populist right-wing party as their favorite party (and unlikely care about image concerns). For voters of the centrist/right-wing party, the estimated interaction term is positive. As such, the observed liberal bias detected in important votes is unlikely caused by self-reported populist right-wing voters but, rather, voters of other affiliations overstating their support for left-wing policies.

---

19 I cannot do the same analysis with the individual topics, as the variance within topics is sometimes small (e.g., immigration).

20 Experimental Evidence points to individuals being less altruistic in large-stakes compared to small-stakes games (Andersen et al., 2011).
All of this evidence suggests that some of the respondents misrepresent their preferences as too left wing in surveys. Apart from political correctness concerns, what other behavioral motives could lead survey respondents to misrepresent their political views? A desire to conform to underlying norms and views (Bernheim, 1994; Kuran, 1995) may create a conflictive situation for minority-view holders. Since the voting result is known at the time of the survey, it is unusually clear what the majority view is. As such, I suspect some of the survey respondents who voted against the majority were reluctant to reveal their true voting result.

To test this conformity motive with individual-level data, I select from the surveyed votes of each respondent the ones where the respondent’s favorite party recommended a no. An individual deviates from the party’s recommendation if he or she states voting yes. The empirical strategy will be to test whether an individual is more likely to report a deviation from his or her favorite party’s recommendation (which is no) if the vote was accepted. Columns 1 to 5 in table 6 regress the indicator variable Deviation Party on the dummy vote accepted and individual fixed effects for voters of the largest four parties separately (columns 1 to 4), and together (column 5). Individual fixed effects capture an individual’s innate propensity to deviate from the policy recommendations of their favorite party. Clearly, a deviation is up to 35 percentage points more likely if the vote passed. Column 6 restricts the set of votes to those where announced approval by the parties (weighted by their vote shares in elections) is similar to approval by the voters (i.e., the difference is smaller than 10 percentage points). The conformity bias persists and alleviates concerns that a vote could have been accepted precisely because voters deviate from party recommendations. Finally, columns 7 and 8 restrict the sample to accepted votes with a share yes in that range, and Margin 45–55 are votes with approval rates between 45% and 55% (column 8). Robust standard errors in parentheses. Significant at ***1%, **5%, *10%.

Table 5.—Mechanism Liberal Bias

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation Left-Wing Party</td>
<td>0.997</td>
<td>−0.592</td>
<td>0.247**</td>
<td>−0.101***</td>
<td>−0.0640**</td>
<td>−0.179***</td>
<td>−0.121***</td>
<td>−0.0944***</td>
<td>−0.0578**</td>
</tr>
<tr>
<td>(1.705)</td>
<td>(1.761)</td>
<td>(0.0195)</td>
<td>(0.0290)</td>
<td>(0.0291)</td>
<td>(0.0221)</td>
<td>(0.0225)</td>
<td>(0.0239)</td>
<td>(0.0242)</td>
<td></td>
</tr>
<tr>
<td>Rec. Left-Wing × Importance Vote</td>
<td>5.440**</td>
<td>4.269</td>
<td>0.172**</td>
<td>0.0887</td>
<td>0.0303</td>
<td>0.0795***</td>
<td>0.0560*</td>
<td>−0.200***</td>
<td>−0.183**</td>
</tr>
<tr>
<td>(2.120)</td>
<td>(2.768)</td>
<td>(0.0254)</td>
<td>(0.0381)</td>
<td>(0.0389)</td>
<td>(0.0295)</td>
<td>(0.0294)</td>
<td>(0.0311)</td>
<td>(0.0315)</td>
<td></td>
</tr>
<tr>
<td>Importance Vote</td>
<td>−3.959**</td>
<td>−3.474</td>
<td>−0.136**</td>
<td>−0.0744**</td>
<td>−0.0253</td>
<td>−0.0529**</td>
<td>−0.0277</td>
<td>0.110***</td>
<td>0.101***</td>
</tr>
<tr>
<td>(1.868)</td>
<td>(2.612)</td>
<td>(0.0231)</td>
<td>(0.0342)</td>
<td>(0.0343)</td>
<td>(0.0266)</td>
<td>(0.0265)</td>
<td>(0.0281)</td>
<td>(0.0282)</td>
<td></td>
</tr>
</tbody>
</table>

Data

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Vote Level</th>
<th>Individual Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control voting result</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Control for own party recommendation</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Party affiliation</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of votes</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>R²</td>
<td>0.415</td>
<td>0.673</td>
</tr>
</tbody>
</table>

Table 6.—Deviations from Party Recommendations

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vote Accepted</td>
<td>0.283***</td>
<td>0.174**</td>
<td>0.347***</td>
<td>0.125**</td>
<td>0.235***</td>
<td>0.331***</td>
<td>0.129***</td>
<td>0.137**</td>
</tr>
<tr>
<td>(0.0330)</td>
<td>(0.0270)</td>
<td>(0.0414)</td>
<td>(0.0591)</td>
<td>(0.0178)</td>
<td>(0.0607)</td>
<td>(0.0388)</td>
<td>(0.0628)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.179**</td>
<td>0.166**</td>
<td>0.185**</td>
<td>0.241**</td>
<td>0.191***</td>
<td>0.166**</td>
<td>0.245***</td>
<td>0.243**</td>
</tr>
<tr>
<td>(0.0153)</td>
<td>(0.00902)</td>
<td>(0.00642)</td>
<td>(0.00800)</td>
<td>(0.00489)</td>
<td>(0.0172)</td>
<td>(0.0164)</td>
<td>(0.0237)</td>
<td></td>
</tr>
<tr>
<td>Individual fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Votes</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Congruent</td>
<td>Margin 40–60</td>
<td>Margin 45–55</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4,500</td>
<td>5,075</td>
<td>5,302</td>
<td>3,421</td>
<td>18,298</td>
<td>7,871</td>
<td>6,145</td>
<td>3,231</td>
</tr>
<tr>
<td>R²</td>
<td>0.821</td>
<td>0.643</td>
<td>0.751</td>
<td>0.718</td>
<td>0.738</td>
<td>0.860</td>
<td>0.878</td>
<td>0.926</td>
</tr>
</tbody>
</table>

The dependent variable is whether an individual deviates from the recommendation of her preferred party (which is to vote no). The variable takes a value of 1 if the individual states voted yes and a value of 0 if the individual states voted no. Vote Accepted is a dummy variable taking a value of 1 if the ballot has been approved by the Swiss electorate and 0 otherwise. Congruent Votes (column 6) are votes where the parties and voters had similar approval (difference <10%). Margin 40–60 (column 7) are votes with a share yes in that range, and Margin 45–55 are votes with approval rates between 45% and 55% (column 8). Robust standard errors reported. Significant at ***1%, **5%, *10%.

Note also that the conformity bias shows up in the aggregate data. Accepted votes have a more positive survey bias than rejected votes. This result holds even with voting-day fixed effects and can therefore not be explained by selection.
differences in survey responses between groups (i.e., gender, religion) and the degree of misrepresentation is the same between groups. The next section sheds light on this issue.

V. Canton Characteristics and Differences between Stated and Official Approval

Switzerland offers a rare opportunity to shed some light on whether survey biases vary across subgroups of people. This is especially relevant for researchers using the Eurobarometer or the World Value Survey to compare political attitudes or values across countries. If survey biases are comparable across countries, this type of study makes perfect sense. If, however, due to different social norms, different types of people respond or lie, such an analysis may be less convincing. There is no way to assess the accuracy of the measured attitudes across countries (since the “true attitudes” for the population are unobserved). The Swiss data allow a comparison of at least whether survey accuracy differs for different cultures and religions within a country. Voting results are the basis for this analysis at the canton level (available for every federal vote), which I can compare with the stated share of yes votes of residents of a given canton.22 The Swiss cantons are highly diverse in terms of culture, religion, and economic richness. While the majority of Swiss cantons are German speaking, there are also a couple of French-speaking cantons and an Italian-speaking canton (plus one canton, Graubuenden, with Romansch as an official language). It is well known that cultural differences between the German- and non-German-speaking cantons are large. For instance, an official term, Roesti-graben, refers to the consistently different voting outcomes between the German- and non-German-speaking cantons. The Swiss cantons are also heterogeneous with regard to religion. While certain cantons are predominantly Catholic, others are predominantly Protestant or mixed. Cantons can also vary with regard to other characteristics such as income, age structure, education, and population size.

These data allow me to uncover interesting correlations but not necessarily causal relationships. However, there are usually no data available to shed light on whether survey biases on reported attitudes (in this case, for different policies) vary across cultures and religions. Again, the reason is that there is information as given in the surveys but no comparison group revealing “true preferences.” Keeping this caveat in mind, I analyze correlations between the absolute value of survey bias and various canton characteristics, taking canton-level voting results (from federal votes) between 1987 and 2000 as the basis of the analysis. Since some smaller cantons have very few survey respondents, I run weighted regressions, with the number of surveyed voters per canton and ballot assigned as weights.

As can be seen from table 7, cantons with a higher share of Protestants have lower biases on average. However, the

---

22 The VOX survey asks for the respondent’s canton of residence.
surveys should be handled with great care.

cates that cross-country comparisons of stated attitudes in
and which part of between-subject variance is explained by
lying on socially sensitive issues differs between subjects
noteworthy of further investigation. For instance, using list-
the finding that survey accuracy depends on culture seems
precisely the amount of survey bias into selection and
avenue for future research would be to try to decompose
likely to be contaminated.

As for the behavioral mechanisms, the data point toward 
respondent lying as being part of the story. An interesting 
avenue for future research would be to try to decompose 
precisely the amount of survey bias into selection and 
respondent lying and deepen our knowledge through which 
survey methods these two biases can be reduced. Also, 
the finding that survey accuracy depends on culture seems 
worthy of further investigation. For instance, using list-
randomization techniques, one could investigate whether 
lying on socially sensitive issues differs between subjects 
and which part of between-subject variance is explained by 
the subjects’ nationalities. If the latter is important, it indi-
cates that cross-country comparisons of stated attitudes in 
surveys should be handled with great care.

VI. Conclusion

This paper analyzes how accurately political preferences are represented in surveys. Using unique data on all Swiss 
votes between 1987 and 2007, I find that the average dif-
ference between stated and real approval is 4.7 percentage 
points. More important, I show that the amount of survey bias 
differs considerably with regard to the policy areas involved. 
For instance, citizens inaccurately reveal their preferences 
on issues related to integration, immigration, the environ-
ment, and certain types of regulation, but not on federal 
finances, health, and institutions. Therefore, the paper sheds 
light on which types of survey questions are more or less 
likely to be contaminated.

As for the behavioral mechanisms, the data point toward respondent lying as being part of the story. An interesting avenue for future research would be to try to decompose precisely the amount of survey bias into selection and respondent lying and deepen our knowledge through which survey methods these two biases can be reduced. Also, the finding that survey accuracy depends on culture seems noteworthy of further investigation. For instance, using list-randomization techniques, one could investigate whether lying on socially sensitive issues differs between subjects and which part of between-subject variance is explained by the subjects’ nationalities. If the latter is important, it indicates that cross-country comparisons of stated attitudes in surveys should be handled with great care.

REFERENCES


