Complex ballot propositions, individual voting behavior, and status quo bias

ZOHAL HESSAMI & SVEN RESNJANSKIJ[1]

UNIVERSITY OF KONSTANZ, DEPARTMENT OF ECONOMICS

One concern about direct democracy is that citizens may not be sufficiently competent to decide about complex policy issues. This may result in a status quo bias due to exaggerated conservatism in citizens' voting behavior. Although this concern is often voiced, there is no evidence on how the complexity of ballot propositions affects individual voting behavior. We develop a novel measure of proposition complexity (using information provided in official pre-referendum booklets) which we combine with micro-data from post-referendum surveys in Switzerland. Using Heckman selection estimations to account for participation bias in the voting decision, we find that increasing proposition complexity from the 10th to the 90th percentile would decrease the approval rate by 5.6 ppts. This decline is often decisive: an additional 12% of the propositions in our sample would be rejected. We also find that these effects are twice as large for less educated citizens and that campaign ads in newspapers aggravate the status quo bias.

JEL codes: *D71, D72, D81*

Keywords: *Voting behavior, proposition complexity, direct democracy, status quo bias, Heckman probit model*

[1] Corresponding author: Zohal Hessami, University of Konstanz, Department of Economics, Box 138, 78457 Konstanz, Germany, zohal.hessami@uni-konstanz.de. The authors thank Kira Gährken, Sebastian Kopf, Christoph Siemroth and Tobias Steinhauser for excellent research assistance and gratefully acknowledge funding by the German National Science Foundation (DFG) through Collaborative Research Center 884 ("Political Economy of Reforms") at the University of Mannheim. The authors also thank the Young Scholar Fund at the University of Konstanz for funding the follow-up project "The complexity of direct-democratic propositions, voter preferences, and the feasibility of political reforms". This article has benefitted from comments and suggestions by participants at the 2013 Workshop on Political Economy at CESifo in Dresden, the 2014 EEA meeting in Toulouse, the 2014 IIPF meeting in Lugano, the 2015 Silvaplana Workshop on Political Economy, and various seminars and a political economy workshop in Konstanz.

I. Introduction

Direct legislation provides citizens with the opportunity to directly set policies. This influence comes at a cost, however. Voters have to decide about ballot propositions on a broad variety of issues. Often several propositions are at stake on the same day where citizens are asked to choose between the status quo (rejecting the proposition) and the proposed ballot measure (accepting the proposition).[2] Since "a defining characteristic of many propositions is complexity" (Lupia 1994, p.63), voters may face considerable difficulties in estimating the personal consequences of accepting or rejecting a proposition. Various authors even go so far to assert that ordinary voters may not be sufficiently competent decide about complex policy issues (Magleby 1984; Cronin 1999).[3]

If voters do face such difficulties in deciding about complex ballot propositions, how may this affect their voting behavior? First, it is likely that complexity has a negative influence on the individual decision to participate in a referendum since voters may derive less utility from turning out when they are uncertain about their voting decision (Matsusaka 1995). Second, the theoretical literature suggests that complexity increases the likelihood that voters reject a proposition (i.e. prefer the status quo). Samuelson and Zeckhauser (1988) were among the first to state that the existence of uncertainty – for instance due to the complexity and a resulting poor understanding of an issue - may lead to a status quo bias. Related to this, Eichenberger and Serna (1996) argue that complexity increases the likelihood that voters make random errors in assessing the costs and benefits of a proposed ballot measure. Due to the nature of the voting procedure these random errors have asymmetric effects on the outcome of a referendum and may lead to a higher likelihood of rejecting a proposition. Finally, a behavioral literature emphasizes that heuristics are primarily used when people have to trade off the effort required in decision-making and the accuracy of one's decision (Gigerenzer and Gaissmaier 2010). Hence, voters are particularly likely to rely on the status quo heuristic when they face highly complex propositions.

To be able to empirically investigate whether the complexity of proposition affects whether voters turn out in a referendum and whether they accept or reject a proposition, one has to address two challenges. First, one needs to find a way to measure proposition complexity. Most of the existing empirical literature on complexity in direct legislation resorts to a convenient measure of *ballot complexity* which counts the number of propositions that were at stake on the same day (Selb 2008; Stadelmann and Torgler 2013). Alternatively, some studies use the number of words or number of lines devoted to a proposition on a ballot as a measure of *proposition prolixity* (Bowler, Donovan, and Happ 1992; Nicholson 2003; Kriesi 2008).[4,5] The ballot complexity measure only takes into account how many

[2] Crowded ballots have been subject to criticism for many decades (see for instance Lapalombara and Hagan (1951)).

[3] In one of the most influential previous studies on proposition complexity, Magleby (1984) examines California ballots during the 1970s and argues that more than 17 years of formal education would be required to understand an average proposition as stated on a ballot.

[4] In a precursory study, Hessami (2016) analyzes the effect of proposition complexity on aggregate referendum outcomes using the number of subjects per proposition as a measure of complexity.

[5] Ballot complexity has been found to be positively associated with (i) a lower awareness of the propositions at stake (Nicholson 2003; Kriesi 2008), (ii) an interference with the pre-referendum deliberative process (Frey 1994), difficulties to translate political preferences into policy choices (Selb 2008); (iii) lower turnout or higher roll-off5 (Bowler, Donovan, and Happ 1992; Bowler and Donovan 1998; Reilly and Richey 2011), (iv) a stronger inclination to reject propositions (Bowler and Donovan 1998), and (v) a stronger reliance on parliamentary recommendations (Stadelmann and Torgler 2013).

propositions voters have to deal with on the same day and thus does not capture the complexity of an individual proposition.[6] The proposition prolixity measure refers to a point in time where a voter has already taken a first decision – when a voter has already turned out and is sitting in front of the ballot and thus one would ignore citizens who abstain in the first place. Second, one has to think carefully about how to most accurately capture the effect of complexity on individual voting behavior. While at first sight it may seem that the participation decision (whether to turn out) is irrelevant to the effect of proposition complexity on the referendum outcome and a potential status quo bias, there are good reasons to believe that this is not the case. In particular, theoretical contributions such as the swing voter's curse theory by Feddersen and Pesendorfer (1996) suggest that it is important to take into account potential interrelations between the participation and voting decision. Therefore, one needs a dataset that includes information on individual decisions to turn out and to reject or accept a proposition as well as a suitable econometric specification that takes into account that the two decisions are interrelated.

In this paper, we empirically study how the complexity of propositions affects individual voting behavior. Our dataset covers all 276 federal referenda that were held in Switzerland in the 1981-2010 period. The data on individual participation and voting decisions is taken from the VOX post-referendum survey series. We combine this rich micro data-set with a novel measure of proposition complexity which we construct based on information provided in official pre-referendum booklets that are sent to all Swiss households prior to a referendum. Our regression model takes into account that the participation and voting decisions are interrelated.

We make three substantive contributions to the literature. Our first contribution is to develop a new measure for proposition complexity. Conceptually, we take a different approach than the existing literature by constructing a measure of the *underlying* complexity of ballot propositions.[7] The complexity of a proposition that voters are exposed to *prior* to a referendum is the variable of interest which we capture with our novel complexity measure based on an extensive data collection effort for 276 federal referenda in Switzerland over the time period from 1981 to 2010. We use information provided in official pre-referendum booklets which the Swiss government is legally obliged to disseminate before each referendum since 1978 (Schweizer Bundesrat 1978). In constructing this measure, we follow the literature which regards complex propositions as those that are "lengthy (…) and technical" (Lupia 1994, p.65). While the level of technicality is subjective and difficult to measure,

[6] Reilly and Richey (2011) study the effect of the readability of propositions in terms of language and find that the number of propositions per ballot (as a measure of ballot complexity) has no independent effect on aggregate drop-off rates once readability is included in the regression model. This provides evidence that the individual complexity of propositions has an impact on voting decisions that goes beyond ballot-specific complexity measures.

[7] Existing measures (ballot complexity and proposition prolixity) conceptually target complexity at a more superficial level, i.e. the number of propositions on a ballot and the proposition content as stated on the ballot. Our measure of proposition complexity does not require citizens to read the information booklets to be exposed to the complexity of a proposition. Instead, the length of the information text in the official information booklet concerning a proposition serves as a proxy of the underlying complexity of a proposition.

the length of the description of propositions can be measured.[8] Hence, we record the number of words in the information and debate section in the official booklets for each proposition.

Our second contribution is that to our knowledge we are the first to study the effect of complexity on individual voting behavior – or on voting behavior in direct legislation more generally – with an econometric specification which addresses sample selection. For any proposition voters face two decisions: (i) whether to participate in the referendum and (ii) conditional on participation whether to choose the status quo (reject the proposition) or the ballot measure (accept the proposition).[9] Since the second decision – the vote decision – is only observable for citizens that participate in the referendum, the classic *sample selection problem* arises (Heckman 1978, 1979). In our case, the selection bias is a participation bias.[10] Previous studies ignore sample selection and typically use aggregated data on the turnout rate or the share of yes- or no-votes as the dependent variable. We apply a structural approach based on two estimation equations (Heckman selection model). We solve the endogeneity problem by means of an exclusion restriction, i.e. we include a variable that influences the individual participation decision but which is arguably orthogonal to the voting decision. This allows us to portray more accurately how complexity affects individual voting behavior and to disentangle a direct effect of complexity on the voting decision and possible indirect effects via the participation decision.

Our third contribution is to validate underlying theoretical mechanisms proposed by the literature.[11] We examine how the effect of complexity on individual voting behavior varies with voters' education level as well as the amount of information on a proposition that is provided in newspaper ads. Various contributions in the literature suggest that voters' decisions to participate in an election or to accept or reject a proposition depends on how informed or educated they are and on how much information on a policy issue is readily available. For instance, Matsusaka (1995) points out that citizens may be more likely to turn out when their issue-related knowledge and education level is high. Gerber and Lupia (1999) argue that voter characteristics that mitigate the uncertainty due to a proposition's content should reduce the status quo bias. Hence, better educated voters have to invest less effort to understand the content of a complex proposition and are less likely to rely on the status quo heuristic. Gerber and Lupia (1999) also describe conditions under which information on campaign spending increases voter competence and thereby may reduce the reliance on the status quo heuristic.

We provide evidence that voters that face more complex propositions are more likely to reject them. Increasing proposition complexity from the 10th to the 90th percentile would decrease the approval rate by 5.6 ppts. This decline is often decisive: an additional 12% of the propositions in our

[8] This approach is partially inspired by Lupia (1994). He analyzes data on exit polls for five insurance reform propositions in California in 1988 to study the role of information shortcuts (partisan cues, etc.) for individual voting behavior. Information is a particularly salient issue for these five propositions due to their unusual complexity and the fact that official information pamphlets were handed out to voters.

[9] In the context of our setting, the pre-reform situation represents the status quo. Voters are asked to approve a set of reforms in the ballot measure. Hence, a rejection implies that the pre-reform situation will prevail and voters are more likely to prefer the pre-reform situation. In addition, note that the default in referendums is always the pre-reform situation. Therefore, a no-vote, which is conceptually a vote in favor of the status quo, is in effect a vote for the pre-reform situation.

[10] For propositions on immigration issues, Krishnakumar and Müller (2012) find a participation bias of 17 ppts which is quite substantial.

[11] More generally, we are able to investigate the transmission channel for status quo bias due to proposition complexity. This has not been investigated in previous studies. Samuelson and Zeckhauser (1988) list a wide variety of explanations and channels for status quo bias but the empirical literature has so far rarely attempted to provide evidence on underlying channels.

sample would be rejected. Two competing mechanisms determine the impact of complexity on the status quo bias in the vote outcome. A *direct effect* – confronted with higher complexity, voters tend to vote in favor of the status quo, and an *indirect participation effect* – increasing complexity reduces the percentage of biased citizens (through vote abstentions) among the voters and therefore mitigates the status quo bias. We contribute to the literature by estimating the magnitude of both effects. We find that neglecting the indirect participation effect would overestimate the effect of complexity on the referendum outcome by almost 50%. The sheer magnitude of the opposing indirect effect calls attention to the importance of treating voting behavior as an outcome of two sequential choices.

Voters with a lower ability to understand complex issues (i.e. with a lower education level) are on average more than 10% more likely than highly educated voters to abstain from voting and to reject propositions. This effect is twice as large for propositions at the 90th percentile compared to propositions at the 10th percentile of the complexity distribution in our sample. Finally, when the campaign intensity in newspapers is higher, voters are on average more inclined to participate in a referendum but are also more likely to reject a proposition. Thus, higher campaign intensity leads to an even larger status quo bias.[12] Our findings hence suggest that less educated citizens, i.e. citizens with weaker cognitive abilities, are disinclined to turn out when propositions are highly complex. Lijphart (1997, p.1) claims that such "unequal participation spells unequal influence" and thereby calls the legitimacy of referenda on complex issues into question. However, in line with the concept of the swing voter's curse (Feddersen and Pesendorfer 1996), we find that biased citizens are more likely to abstain from voting which renders the vote outcome informationally superior since it mitigates the status quo bias in the vote outcome. By estimating the participation bias, we contribute a quantitative dimension to the discussion of the underlying trade-off between a *representative vote outcome* (Lijphart 1997) and an *informationally superior vote outcome* (Feddersen and Pesendorfer 1996).

We conduct five robustness tests that address potential concerns regarding our complexity measure and the possibility that other mechanisms explain our main estimation results. First, our objective complexity measure has the advantage that it is unrelated to individual characteristics and can thus be used to study the interaction between voters' characteristics and proposition complexity. Nevertheless, as a robustness test we use an alternative measure of proposition complexity based on voters' perception of how difficult it was for them to form an opinion about the consequences of rejecting or accepting a propositions. The estimation results are similar. Second, we use data on a survey question that indicates whether a respondent has or has not used the official information booklet. We find that our estimation results for the effect of complexity on individual voting behavior do not differ between booklet users and nonusers. This indicates that our complexity measure indeed captures proposition complexity at a deeper level and that its validity does not require citizens to read the information booklets. Third, additional estimations show that our complexity measure and our main estimation

[12] The increase in campaign intensity represents a decrease in voting costs as it becomes cheaper for voters to inform themselves about an issue. Related to this, Hodler, Luechinger, and Stutzer (2015) also find that lowering voting costs due to the introduction of postal voting leads to an increase in turnout. The authors additionally find that the lowering of voting costs has especially led individuals with less political knowledge and education to turn out.

results are not confounded by the fact that more complex propositions may be perceived by voters as more important. Fourth, we provide evidence that the government does not strategically manipulate proposition complexity to induce voters to vote in line with the government's recommendation. Finally, we provide evidence that our main estimates are not contaminated by those propositions where survey bias - as identified by Funk (2015) – may be an issue.

The findings in this paper are interesting not only because they provide a comprehensive and accurate analysis of how complexity in direct legislation affects individual voting behavior, but also because direct democracy is becoming a more important tool for decision-making in various contexts (see for instance the British referendum on EU membership or the referendum in Greece on the bailout packages). From a policy perspective, it is important to know how the complexity of policy issues influences individual voting behavior as well as aggregate referendum outcomes. An obvious implication of the finding that campaign intensity as measured by ad space in newspapers reduces the participation bias, but does not mitigate the status quo bias is that media attention does not necessarily increase voters' competence. Instead our results suggest that it is preferable that governments invest more resources in general education to form politically mature citizens able to make informed decisions even if propositions are unusually complex.

The rest of the paper proceeds as follows. Section II discusses theoretical considerations for the effect of proposition complexity on individual voting behavior. Section III discusses the empirical strategy. Section IV describes the data. Section V presents the main estimation results. Section VI presents evidence on two extensions. Section VII discusses five robustness tests. Section VIII concludes.

II. Theoretical Considerations

A. *Proposition complexity and the participation decision*

Downs (1957) and Riker and Ordeshook (1968) introduce a theory of voting based on a rational trade-off between the costs of voting and the expected benefits. Matsusaka (1995) extends the traditional rational voter model by highlighting the role of limited information. Voters are portrayed as utility-maximizing consumers who receive higher payoffs from casting their vote when they are more confident of their vote choice. We develop this argument further by taking into account that the amount of information that a voter needs to reach a certain level of confidence depends on the complexity of the issue at stake. In addition, Matsusaka (1995) states that information is meant to comprise specific information about an issue (which may be provided by the media) as well as general knowledge (the education level). This implies that when voters are more educated and when the media provide more information on a proposition voters are less inclined to abstain from voting. We propose that it is the interaction between the complexity of a policy issue and voters' levels of education as well as the campaign intensity in newspapers that influences whether voters will participate in an election or not.

In a similar vein, the swing voter's curse theory argues that less informed voters may rationally prefer to abstain from voting even when they have a strict preference in favor or against a proposition and voting is costless (Feddersen and Pesendorfer 1996). Assuming that voters have homogeneous preferences but differ in their ability to identify the preferred choice, uninformed voters can only be pivotal if they vote differently than informed voters which is irrational since informed voters vote for the preferred option with certainty. Therefore, some uniformed voters may find it rational to vote against their prior beliefs as long as enough uniformed voters still vote in opposition to informed voters. If all uniformed voters vote against their prior beliefs, they may dominate informed voters and the inferior alternative will be elected. As a consequence, uniformed voters find it optimal to abstain from voting to maximize the probability that informed voters determine to vote outcome (Feddersen and Pesendorfer 1996). It is straightforward to assume that the share of uniformed voters is positively related to the complexity of the policy issue and negatively related to the level of education. As a result, educated citizens are more likely to participate and turnout is expected to be lower for complex topics.

The empirical implication of these two theories is that on average voters should be more reluctant to participate in a referendum with increasing proposition complexity, and this effect should be stronger for individuals with a low level of education and when campaign intensity is high.

B. *Proposition complexity and the voting decision*

Samuelson and Zeckhauser (1988) were among the first to draw wide attention to the issue of status quo bias in decision-making. They conducted a number of experiments showing that individuals disproportionately tend to stick with the status quo. The authors draw on a broad range of insights from economics, psychology, and decision theory to provide theoretical explanations for status quo bias. One of these explanations is the existence of uncertainty. The authors state that an early choice may have substantial advantage over an alternative. This could be for instance the status quo in a certain policy area before a referendum takes place. From a consumer choice perspective, Samuelson and Zeckhauser (1988) argue that consumers remain loyal to a chosen brand as long as their utility from consuming this product is above a certain threshold. In the context of referendums this implies that as long as voters enjoy a minimum level of utility or do not significantly suffer from the status quo they may reject a proposition and may not even bother to find out whether they might benefit from a policy change. The authors also argue that "the choice to undertake a decision analysis is itself a decision" (p. 35). If these costs are high, voters may only conduct this analysis once, take a decision, and then defer to this choice in the future. If we apply these ideas to the context of direct democracy, one could argue that both the complexity of propositions as well as voters' education levels and the amount of information provided in the media determine how large the cost of analysis is and whether a proposition is rejected.

Second, a key insight of Eichenberger and Serna (1996) is that individual errors in the assessment of expected benefits, even if random, have asymmetric effects on the referendum outcome due to the nature of the voting procedure. A proposition may benefit an average voter; i.e. the associated policy changes generally increase net-utility (which seems plausible based on the relatively high quality of governance in Switzerland). For an individual voter, however, these policy changes may or may not be beneficial. The complexity of a proposition increases the variance of the expected benefits of a proposition for a given voter. When complexity is higher, a larger number of voters will believe that this proposition is very beneficial or very harmful for them, i.e. individual errors become larger. Technically speaking, the tails of the distribution of expected benefits become fatter. The larger number of voters who (wrongly) believe that the proposition is very beneficial for them is inconsequential for the referendum outcome. These voters would vote in favor of the proposition even if it were less complex. The larger number of voters who (wrongly) believe that the proposition is very harmful, however, has substantive consequences for the referendum outcome. Some of the voters who actually benefit from the proposition now underestimate its benefits and reject the proposition. This mechanism is likely to vary with the education level of voters. Random errors may be less relevant for more educated voters who are better equipped to process complex information or more specifically to gauge the personal costs and benefits of complex propositions. Related to this, Eichenberger and Serna (1996) state: "it is difficult to measure the complexity of an issue independently from the individuals' human capital" (p.140).

Third, a behavioral literature emphasizes that heuristics (such as the status quo heuristic) are primarily used in situations where people have to trade off the effort required in decision-making and the accuracy of one's decision (Gigerenzer and Gaissmaier 2010). Therefore, better informed or better educated voters may have to invest less effort and are less likely to rely on the status quo heuristic in their decision-making process. Highly educated voters will therefore be less likely than less educated voters to exhibit a status quo bias, even though they may also reject propositions because they are highly complex. Gerber and Lupia (1999) identify a consensus among scholars that average citizens in direct democracies lack competence to make informed choices. From a rational choice perspective, voter characteristics that mitigate the uncertainty due to a proposition's content should reduce the status quo bias. Hence, better informed or better educated voters may have to invest less effort to understand the content of a proposition and are less likely to rely on the status quo heuristic. This makes highly educated voters less likely to reject propositions due to their complexity.

The empirical implication of these theories is that on average voters should be more likely to reject propositions with increasing proposition complexity; this effect should be stronger for individuals with a low level of education and when campaign intensity is low.

III. Empirical Strategy

A. *Participation bias and endogeneity*

For any proposition that is at stake, voters face two decisions: (i) participation versus abstention and (ii) conditional on participation the status quo (reject the proposition) versus the ballot measure (accept the proposition). Since the vote decision is only observable for the subset of citizens that participate in the referendum, the classic *sample selection problem* may arise (Heckman 1978, 1979).

We hypothesize that the participation and the vote decision are both influenced by the complexity of a proposition. The two decisions can be represented by the following binary choice models:

(1) $$Participate^{*} = \alpha\ Complexity + \boldsymbol{\beta}'\mathbf{x} + u, \quad Participate = 1 \text{ if } Participate^{*} > 0,$$

$$Participate = 0 \text{ otherwise.}$$

(2) $$Yes\text{-}vote^* = \alpha\, Complexity + \boldsymbol{\beta}'\mathbf{x} + \epsilon,\ Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^* > 0,$$
$$Yes\text{-}vote = 0 \text{ otherwise.}$$

where vector $\mathbf{x}$ includes a set of control variables.[13] In equation (1), the complexity of the entire ballot is relevant for the participation decision (i.e. the sum of complexity across all propositions on a ballot), whereas in equation (2) the complexity of individual propositions is relevant for the voting decision.

According to Heckman (1979), the effect of self-selection of citizens into the voting sample can be interpreted as omitted variable problem in the voting equation (2). Several solutions to this problem may come to mind.

A first idea would be to restrict the sample to the voting population, i.e. to ignore the first equation and only rely on the data of the subsample of citizens who actually voted. This approach ignores that complexity or any other variable may alter the decision to vote in favor of a proposition and may also change the composition of voters participating in the election. This would cause either an upward or downward bias in α depending on how voters and non-voters differ in their response to complexity. If an increase in complexity causes citizens, which otherwise would have voted against the proposition, to abstain from voting, then α (a measure of the magnitude of the status quo bias) would be downward biased in regression equation (2), overstating the extent of the status quo bias.[14]

A second potential solution is to control for participation in the vote decision equation and to use an instrument for the potentially endogenous participation decision. However, an IV approach is not feasible since the voting decision is only observable for citizens who participate in a referendum.

Another solution for the omitted variable problem in equation (2) is to control for all characteristics of the participating decision by adding additional variables to equation (2). Even after controlling for all observable characteristics, the selection process might still be driven by unobservable factors. Building on the literature on status quo bias (Samuelson and Zeckhauser 1988), unobserved characteristics of respondents such as the ability to cope with complex propositions or preference parameters like risk aversion may influence whether they are willing to bear the participation costs and whether they are biased towards supporting the status quo.

Our solution to the selection problem is to use a Heckman selection model which is identified with the help of an exclusion restriction, i.e. we include a variable in the first equation that influences the participation decision but which is arguably orthogonal to the vote decision. In doing so, we can estimate the magnitude of the selection bias – which is synonymous to the participation bias – and correct our estimates regarding the effect of complexity on the individual voting decision for this bias.

[13] For notational convenience, we use the same Greek letters indicating the coefficients to be estimated in equation (1) and (2). However, they can represent different estimates in each equation. We stick to this convention throughout the paper.
[14] In equation (2), a status quo bias corresponds to a negative estimate for α.

In this paper, the analysis of self-selection of citizens into the sample of voters should not be understood as a primarily statistical problem, but rather as an attempt to bring the statistical analysis closer to the structure of the underlying political-economic theory related to the act of voting.

B. *Exclusion restriction and Heckman selection approach*

Without an exclusion restriction in equation (1), identification would solely rely on the bivariate normality assumption for the functional form of the error terms. Wooldridge (2010) shows that identification based on this assumption alone can be misleading and produce spurious results. In our setting, a valid exclusion restriction requires a variable that influences participation but that has no direct effect on the vote decision.

Based on a sizable literature on the relationship between the closeness of elections and turnout, we include the ex-post approval share of a proposition, i.e. the share of yes-votes among all valid votes that are cast, as a valid exclusion restriction. A positive correlation between the closeness of an election and the individual likelihood to cast one's vote is firmly grounded on various theoretical arguments. The seminal literature asserts that the benefit of voting increases with the probability of casting the decisive vote (Downs 1957; Riker and Ordeshook 1968). This probability is higher when an election is contested. This has been denoted in the literature as the *Downsian Closeness Hypothesis* (Matsusaka and Palda 1993). A second prominent explanation for this positive relationship relies on more mobilization efforts (lowering participation costs) by stakeholders in contested elections which affects the individual probability to participate in the referendum (Denver and Hands 1974; Key and Heard 1984; Cox and Munger 1989). On the other hand, there is no reason why a close election would make it more likely that voters support or reject propositions, i.e. there is no correlation between the vote decision and the closeness of elections.[15]

Our Heckman selection approach can be described by the following two binary choice equations:

(3) $$Participate^* = \alpha\ Complexity + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\delta}'\mathbf{z} + u,\ Participate = 1 \text{ if } Participate^* > 0,\quad Participate = 0 \text{ otherwise},$$

(4) $$Yes\text{-}vote^* = \alpha\ Complexity + \boldsymbol{\beta}'\mathbf{x} + \epsilon,\ Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^* > 0,\quad Yes\text{-}vote = 0 \text{ otherwise},$$

$$\text{with } \begin{pmatrix} u \\ \epsilon \end{pmatrix} | \mathbf{x}, \mathbf{z} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right],$$

where $\mathbf{z}$ captures the exclusion restriction.[16] The error terms u and ϵ are assumed to be distributed bivariate normal with $(u, \epsilon) \sim\ bivariate\ normal[0, 0, 1, 1, \rho]$, where ρ denotes the correlation between the error terms. By estimating ρ within the Heckman selection model, we are able to control for

[15] In principle, the share of yes-votes depends on the sum of individual voting decisions and may appear to be an invalid exclusion restriction at first sight. However, the absolute number of valid votes for each proposition in the time interval covered in our sample (1981-2011) was on average 2 million. Therefore, the individual voting decision has a negligible influence on the share of yes-votes. Another thought experiment illustrating the validity of our exclusion restriction. Imagine a new variable for the share of yes-votes is constructed where we exclude the vote of one individual. The value of the adjusted share of yes-votes would change only to a very small degree, which would not affect our estimation results.

[16] To be more precise, we will include the approval share as well as its square since we expect an inversely U-shaped relationship.

unobserved factors influencing both the turnout and voting decision. An estimate of ρ different from zero would point towards a presence of a selection (participation) bias justifying the Heckman selection model as preferred estimation strategy over the simple probit estimates based on equations (1) and (2). As in section III.A, in the participation equation complexity is aggregated at the ballot level, whereas in the voting equation complexity is included at the proposition level. Vector $\mathbf{x}$ includes a number of control variables for standard voter characteristics: female dummy, age, education level, knowledge about the proposition, married dummy, Protestant dummy, employed dummy.[17] We additionally include dummies for the canton in which the respondent is living, the year in which a referendum is held and the policy area in which a proposition falls.[18] The inclusion of additional controls is a straightforward way to control for a selection bias in equation (4), if selection is only driven by observables. However, it is unlikely that voters' preferences as well as their capacity to understand complex proposition can be fully accounted for by the inclusion of standard socio-economic variables. The empirical results for the selection model presented in section V.B support this view.

We use the log of the number of words for the information text of a proposition as the complexity measure (more details will be provided in section IV.B) since we expect that the same absolute increase in the number of words of the information text has a stronger effect on voting behavior for instance for an increase from 100 to 200 words than for an increase from 1000 to 1100 words, i.e. relative rather than absolute differences are relevant for voters. All hypothesis tests are based on standard errors that are clustered at the ballot level. As suggested by Freedman and Sekhon (2010), we solve the two-equation model in equations (3) and (4) using full information maximum likelihood estimation[19] instead of the two-step procedure originally introduced by (Heckman 1978, 1979). The first precursory study that estimates two probit equations in a Heckman selection framework is Van de Ven and van Praag (1981).

IV. Data Description

A. *Post-referendum survey data*

We use data from standardized and representative polls conducted after each national referendum in Switzerland since 1981. The GfS Research Institute in Berne conducts these surveys on behalf of the Institutes of Political Science at the Universities of Berne, Geneva, and Zurich (FORS -

[17] Brunner, Ross, and Washington (2011) show that economic conditions shape preferences on direct-democratic legislation. In our estimations, this would at least be partially captured by the employed dummy. Funk and Gathmann (2015) provide evidence that female voters make different choices on direct-democratic propositions in Switzerland than male voters.

[18] Status quo bias and the willingness to participate in referenda may differ across cantons for cultural reasons. Year dummies allow us to capture common shocks in specific time periods as well as trends in participation and vote decisions. Policy area fixed effects reduce the variation in proposition complexity by the amount that is exclusively due to the policy area.

[19] The log-likelihood which has to be estimated is given by: $\ln L\left(\alpha_\pi, \alpha_v, \boldsymbol{\beta}_{\boldsymbol{\pi}}, \boldsymbol{\beta}_{\mathbf{v}}, \boldsymbol{\delta}; v, \pi, c_\pi, c_v, \mathbf{x}, \mathbf{z}\right) = \sum_{\pi,v=1} \ln \Phi_2(\alpha_\pi c_\pi + \boldsymbol{\beta}'_{\boldsymbol{\pi}}\mathbf{x} + \boldsymbol{\delta}'\mathbf{z},\ \alpha_v c_v + \boldsymbol{\beta}'_{\mathbf{v}}\mathbf{x},\ \rho) + \sum_{\pi=1,v=0} \ln \Phi_2(\alpha_\pi c_\pi + \boldsymbol{\beta}'_{\boldsymbol{\pi}}\mathbf{x} + \boldsymbol{\delta}'\mathbf{z},\ -\alpha_v c_v - \boldsymbol{\beta}'_{\mathbf{v}}\mathbf{x},\ -\rho) + \sum_{\pi=0} \Phi(-\alpha_\pi c_\pi - \boldsymbol{\beta}'_{\boldsymbol{\pi}}\mathbf{x} - \boldsymbol{\delta}'\mathbf{z})$. The variables *Participate* and *Yes-vote* are denoted by π and v. Parameters corresponding to the participation equation (3) are denoted by a subscript π, whereas parameters denoted by v correspond to the voting equation (4). Φ_2 represents the bivariate normal cdf. Φ denotes the standard normal cdf.

Swiss foundation for research in social sciences 2012). A random sample of 700 to 1000 eligible voters is selected from the Swiss telephone book and surveyed within two weeks after the elections.[20]

The VOX survey asks citizens about their participation in each referendum as well as their individual voting decision. These are our dependent variables in equations (3) and (4). The respondent is also asked about his knowledge about the proposition, the kind of media consulted prior to the referendum, the perceived importance of the vote, and various personal characteristics (age, gender, education, marital status, profession, etc.). Summary statistics for all variables at the respondent level are to be found in table A.1 in the appendix.

B. *Official information booklets and complexity measure*

Since 1978, the Swiss government is obliged by law (Schweizer Bundesrat 1978) to mail a written information booklet before each national referendum to all eligible voters. The office in charge of writing the information booklet ("Bundeskanzlei") has to follow strict legal rules regarding the content of the information material. The booklets are required to be short, objective, transparent, and in line with the principle of proportionality (Bundesgericht 2008). It is explicitly forbidden by law to influence the decision-making process of voters towards accepting the proposition.[21]

Each proposition has a separate chapter in the booklet. The booklets on average have a total size of around 50 pages. Each chapter usually consists of four sections, a short summary, a detailed information section, a debate section comparing arguments against and in favor of the proposition, and a legal section in which parts of the wording of the law are published that would change if the referendum is successful.

We construct a novel proposition complexity measure based on the information booklets. We use standard office software that transfers the booklets into a machine-readable format that allows us to count the number of words in the information and debate section for each proposition. Highly complex propositions are associated with a more detailed and therefore longer description in the booklet. The strict legal framework requires a short and balanced booklet text and prohibits that the government agency in charge of writing the information booklets influences voters by exaggerating the view of the government. We therefore argue that the length of the information text is determined solely by the necessity to provide longer descriptions of more complex propositions to ensure that the content of the information booklets complies with legal requirements.[22]

Several mechanisms may cause the complexity of a proposition to be positively associated with the number words used in its information text. First, the content of the proposition itself is difficult to understand. In that case, one would expect a more extensive description to make the topic of the

[20] The interviews are conducted as follows. The interviewer calls, introduces himself and asks whether there is an eligible voter in the household. If there are several eligible voters in a household, the one who has his birthday on the earliest day in the year is interviewed.

[21] In a robustness test reported in section VII.D, we find no empirical evidence for a correlation between the government recommendation and our text based complexity measure.

[22] An alternative measure for complexity could be constructed based on the legal text that would change if a referendum is successful. Each proposition gives rise to a change to the constitution or to existing laws. Huber and Shipan (2002), however, argue that the detailed language and the resulting length of the legal text might be driven by politicians' incentives to delegate policy making to other policymaking authorities such as bureaucrats.

proposition accessible to the average citizen. Second, a proposition that implies various individual policy measures in a certain policy area requires a description of each policy measure and therefore requires a longer description in the official booklet.

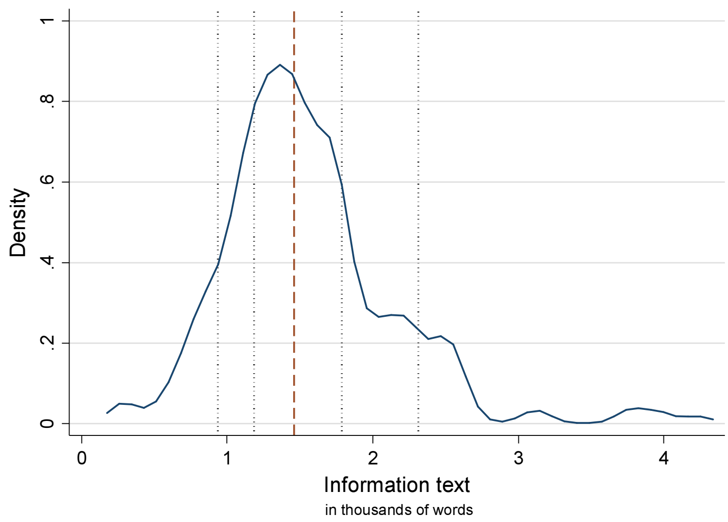

FIGURE 1. DISTRIBUTION OF THE PROPOSITION COMPLEXITY MEASURE

Notes: This figure depicts the variation in our proposition complexity measure using a Gaussian kernel density plot with a kernel bandwidth of 100 words. The dashed red line represents the median of the complexity measure. The dotted grey lines correspond to the 10th, 25th, 75th, and 90th percentiles, respectively. For better readability, the information text axis is restricted to values below the 99th percentile.

Figure 1 illustrates the large variation in proposition complexity in our sample of 276 federal referendums in Switzerland between 1981 and 2010. The median length of the information text is around 1,500 words; roughly 80 percent of the observations lie between 900 and 2,300 words.

V. Estimation Results

A. *Probit estimations for the participation and vote decision*

In this section, we report estimation results that we obtain when ignoring participation bias. We estimate two separate estimation equations (see equations (1) and (2)) using the probit estimator. The benefit of this exercise is to obtain a benchmark based on a naïve empirical strategy which allows us to assess how the results change when we account for participation bias (see section V.B).

Table 1 collects the regression results for the participation decision (models (1) to (4)) and the voting decision (models (5) to (8)).[23] We report average marginal effects instead of probit coefficients. Note that the sample size for the estimation of the voting decision is only about half as large ($N = 107.420$) as the sample used for the estimation of the participation decision ($N = 204.818$) since we only observe the voting decision for citizens who participated in the referendum.

[23] Summary statistics for the variables used in the regressions are provided in table A.1 in the appendix.

TABLE 1—PROBIT ESTIMATION RESULTS: COMPLEXITY AND THE INDIVIDUAL PARTICIPATION AND VOTE DECISION

Avg. Marginal Effects reported	Dep. Var.: Participation				Dep. Var.: Yes-Vote			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Complexity (ballot)	-0.098***	-0.104***	-0.115***	-0.122***				
	(0.036)	(0.036)	(0.043)	(0.036)				
Complexity (proposition)					-0.085**	-0.081**	-0.093***	-0.105***
					(0.040)	(0.035)	(0.034)	(0.030)
Rural	-0.003	-0.010**	-0.008*	-0.008	-0.040***	-0.035***	-0.036***	-0.037***
	(0.005)	(0.004)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Female	-0.003	-0.004	-0.004	-0.002	0.020***	0.023***	0.024***	0.024***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.006)
Age	0.005***	0.005***	0.005***	0.005***	-0.001**	-0.000*	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.044***	0.044***	0.036***	0.035***	0.016***	0.017***	0.018***	0.018***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.004)	(0.003)	(0.003)
Proposition Knowledge	0.056***	0.056***	0.065***	0.068***	0.027**	0.031***	0.028***	0.028***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.012)	(0.008)	(0.007)	(0.006)
Married	0.071***	0.069***	0.073***	0.072***	-0.012*	-0.012**	-0.013***	-0.013***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Protestant	0.017***	0.011*	0.021***	0.020***	-0.006	-0.009*	-0.007	-0.007
	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)	(0.005)	(0.004)	(0.005)
Employed	-0.002	-0.006	-0.005	-0.005	0.003	0.001	-0.001	-0.001
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Canton dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Referenda type dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes	No	No	Yes	Yes
Policy area dummies	No	No	No	Yes	No	No	No	Yes
Pseudo R^2	0.140	0.149	0.164	0.172	0.009	0.058	0.070	0.087
Observations	204818	204818	204818	204818	107420	107420	107420	107420

Notes: The table establishes the negative and significant effect of complexity on eligible voter's probability to participate in a referendum and to vote in favor of a proposition. Average marginal effects based on probit regressions are reported in all specifications. The marginal effect of age is based on age and its squared term. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses. Summary statistics are provided in table A.1 in the online appendix.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

The results for models (1) to (4) suggest that the (average marginal effect) coefficients of all complexity measures are negatively and significantly related with the probability of participating in the referendum.[24] The magnitude of the negative effect of complexity on participation is considerable. Based on our estimates in model (4), when complexity increases by one standard deviation (i.e. the information text is 3992 words longer at the ballot level), citizens on average have a 6.4 ppts lower probability to participate in the referendum.[25] These results are robust to the inclusion of fixed effects for cantons, referenda type, years, and policy areas. The results also show that more educated, more politically interested and more knowledgeable voters are significantly more likely to participate in a referendum. A reasonable explanation is that these voters have to invest fewer resources to estimate the consequences

[24] The high significance levels in our statistical tests are not driven by the large sample size used in the regression analysis. Our objective complexity measure varies only at the proposition, respectively ballot level. Therefore, we correct our standard errors by clustering at the ballot level, allowing observations within a ballot to be correlated. The power of our statistical tests is therefore determined by the number of independent observations (ballots) in our estimation sample. With respect to the relatively low number of 74 ballots, high significance levels cannot be attributed to the size of the entire sample, but rather to a large quantitative effect (as we will illustrate later in this section) and maybe to relatively low noise in our estimates regarding the effect of complexity on voting behavior.

[25] This value is calculated based on a centralized change of one standard deviation in the complexity measure based on the log of the word count.

of their voting decision. The results for models (5) to (8) show that proposition complexity has a significantly negative effect on the probability of voting in favor of a proposition. Based on our estimates in model (8), when complexity increases by one standard deviation (i.e. the information text is 895 words longer at the proposition level), citizens on average are 4.6 percentage points less likely to vote in favor of a proposition. The inclusion of fixed effects slightly increases the precision (through lower standard errors) and the size of the estimated marginal effect.[26]

In Figure 2, we plot the results for the most complete models (models (4) and (8)) to discuss in more detail the size of our estimates.[27] The shaded areas indicate the 95 percent confidence intervals. The vertical dotted lines illustrate the distribution of proposition complexity in our sample by indicating the 10th, 25th, 50th, 75th, and 90th percentiles.

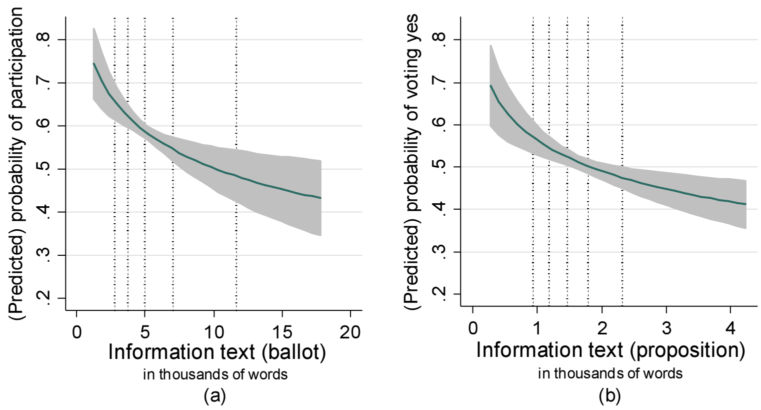

FIGURE 2. EFFECT OF COMPLEXITY ON THE INDIVIDUAL PARTICIPATION AND VOTING DECISION

Notes: This figure depicts the statistically significant negative effect of complexity on the probability that a voter participates in a referendum and votes in favor of a proposition. We plot average predicted probabilities against complexity. The estimates in panel (a), and (b) are calculated based on the estimation results for models (4) and (8) in table 1. The shaded area in panels (a) and (b) represents the 95 percent confidence interval band of the predicted probabilities. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentiles of complexity, respectively. For better readability, the information text axis is restricted to values below the 99th percentile is plotted.

According to Figure 2, the predicted participation rate for a low-complexity ballot (complexity at 90th percentile) is around 65% and falls below 50% for a high-complexity ballot (10th percentile). The predicted probability of voting in favor of a proposition drops by slightly less than 10 percentage points when comparing a low-complexity proposition (90th percentile) with a high-complexity proposition (10th percentile). Our estimation results based on a naïve empirical strategy that ignores participation bias point toward substantial status quo bias in voting behavior due to complexity. In the next section, we

[26] These results derived from our objective, information text-based complexity measure are quantitatively in the same ballpark as the results derived with a subjective, survey-based measure. For comparison, we report the results based on our subjective complexity measure in table A.8 in the appendix. We find that citizens who reported difficulties to form an opinion have a 12.3 percentage point lower probability to participate in the election and a more than 5 percentage point lower probability to vote in favor of a proposition than citizens reporting no difficulties. The results of both, objective and subjective complexity measures provide strong empirical evidence for the existence of a status quo bias in the behavior of participating voters. However, the subjective survey based complexity measure might be confounded with observable and unobservable individual characteristics. We discuss this point in detail in section VII.A

[27] We follow McCloskey and Ziliak (1996) and Ziliak and McCloskey (2008), suggesting that the size of the estimate is at least as important as the statistical significance. The following graphical representation of the nonlinear relationship relies on comments in Wooldridge (2004) and Greene (2010).

will apply a more sophisticated empirical strategy to account for participation bias and thereby obtain unbiased estimation results.

B. *Heckman estimations to identify and correct for participation bias*

As described in section III.B, we include the variables *approval share* and *approval share squared* in our estimations to implement the exclusion restriction.[28] To test the validity of this approach, we regress the participation dummy on our measure for the voter's expectation of a narrow voting decision using a probit estimator. Table 2 provides the regression results. We find a statistically significant hump-shaped relationship which we also illustrate in Figure 3. The highest participation rate is indeed associated with a close election outcome (share of yes-votes ~ 50%).[29]

TABLE 2—CLOSENESS OF THE REFERENDUM OUTCOME AND PARTICIPATION

Probit coefficients reported	Dep. Var.: Participation			
	(1)	(2)	(3)	(4)
Complexity (ballot)	-0.322***	-0.326***	-0.404***	-0.412***
	(0.110)	(0.111)	(0.124)	(0.109)
Exclusion Restriction				
Approval Share	1.721***	1.677***	1.785***	1.709***
	(0.593)	(0.636)	(0.658)	(0.656)
Approval Share squared	-2.076***	-1.980***	-2.223***	-2.058***
	(0.579)	(0.617)	(0.678)	(0.668)
Canton dummies	No	Yes	Yes	Yes
Referenda type dummies	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes
Policy area dummies	No	No	No	Yes
Pseudo R-squared	0.140	0.149	0.164	0.172
Observations	204818	204818	204818	204818
p-value for joint significance of linear and quadratic terms in:				
Approval Share	0.000	0.003	0.001	0.003

Notes: The table illustrates the significant hump-shaped effect of the approval share on the probability to participate in a referendum. Regression coefficients based on probit regressions are reported in all specifications. The variable approval share measures the share of yes-votes obtained from the official Swiss election data (University of Bern, Institute of Political Science 2013). The table also reports the p-value for the joint significance of the variable approval share and its squared term. All equations are estimated including individual controls as in table 1. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

[28]Instead of relying on the quadratic specification, we could use a measure for the distance of the share of yes-votes from the 50% threshold. Our approach, however, is preferable. It allows the effect to be nonlinear, thereby we do not impose a hump-shaped relationship with a peak around 50%, but it is a result of the estimation. We therefore not only test whether the share of yes votes turns out to be a statistical significant regressor, but also whether the implied quadratic functional form is (i) indeed hump-shaped (negative coefficient of the *approval share squared variable*) and (ii) has its peak around 50% (see Figure 3) as theory would predict.

[29]A second condition for the validity of our exclusion restriction is that the expectation of a narrow voting decision should have no direct effect on the decision to vote in favor of a proposition. This does not exclude the possibility of an indirect effect via the decision to participate, which does not violate the assumptions regarding to a valid exclusion restriction in the Heckman selection model.

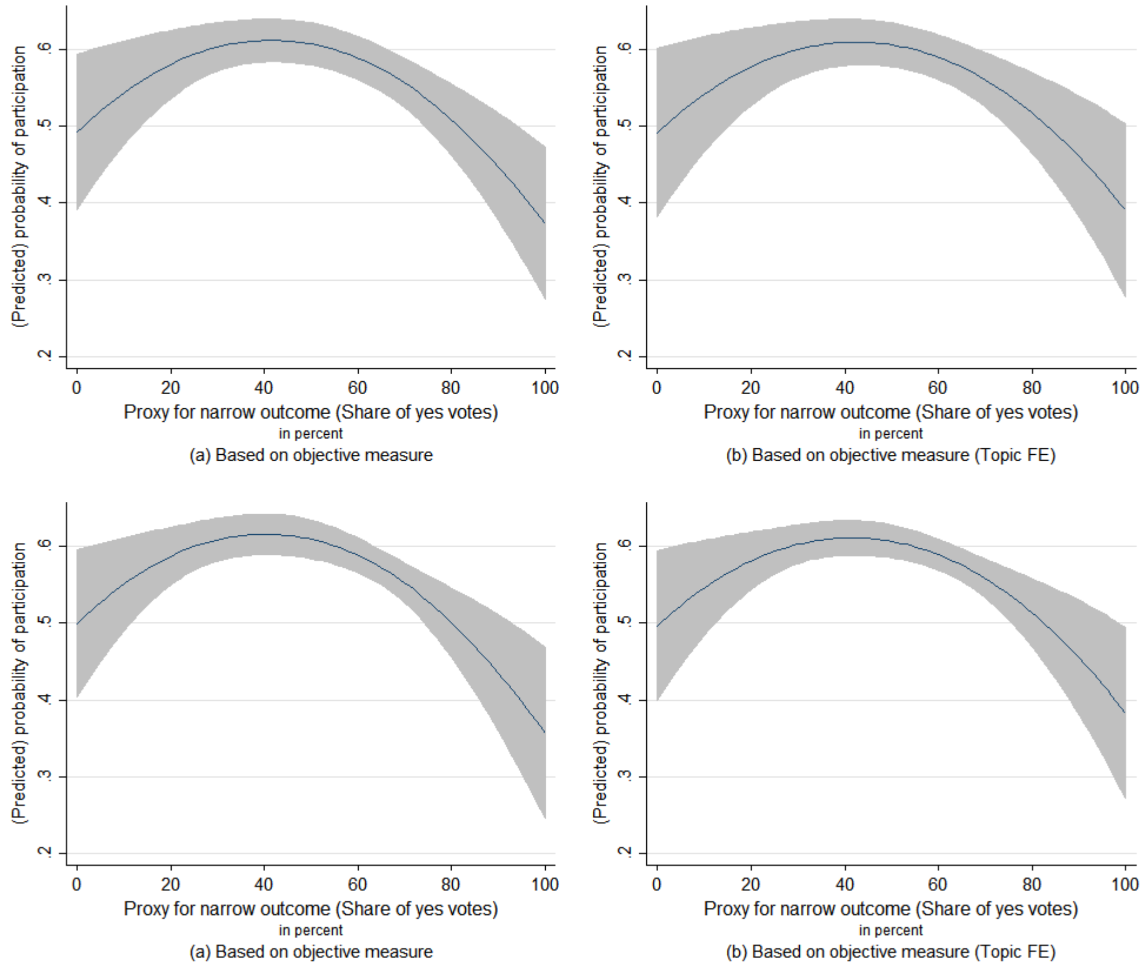

FIGURE 3. (EXPECTED) NARROW OUTCOME ON YES VOTE

Notes: This figure illustrates the hump-shaped relationship between an (expected) narrow referendum outcome and the probability that a citizen participates in a referendum. We plot the average predicted probability against the share of yes-votes. The estimates in panel (a), (b), (c) and (d) are calculated based on the results for probit regressions, in particular models (1) to (4) in Table 2. The shaded area represents the 95 percent confidence interval band of the predicted probability of participation.

Table 3 presents the estimated coefficients for the Heckman selection model.[30] The coefficients for the complexity measure have the expected negative sign and are significant at the 1 percent level. The estimated coefficient ρ measures the correlation between the error terms of the participation and the voting equation and can be interpreted as a measure of unobserved factors affecting both the participation and outcome decision. The estimate for ρ is positive and statistically significant in all specifications. This implies that unobserved factors affect the probability to participate and the probability to vote in favour of the proposition in the same direction.[31] This means that the indirect effect of complexity on the voting decision through the participation decision is positive. Higher complexity

[30] Note that the number of observations is smaller in table 3 than in tables 1 and 2 (191669 instead of 204818). The reason is that in the two-equation Heckman selection model any missing observation in the voting equation will also lead to a missing observation in the participation equation.

[31] In all specifications we find the size of the correlation decreasing, as we include further controls and fixed effects. This result nicely illustrates the interplay between observables and unobservables in the model. Controlling for observable determinants decreases the role that unobserved factors play in determining the participation and voting decision. However, even in model (4) in Table 3, where we control for canton, referenda, year, and policy area differences (as well as socio-economic factors, which are included in all specifications in Table 3), the correlation coefficient ρ is still quantitatively large and precisely measured.

increases the probability that citizens (who otherwise would have voted against the proposition) are overwhelmed by complexity and therefore abstain from voting.

TABLE 3—RESULTS FOR HECKMAN SELECTION MODELS: COMPLEXITY AND THE INDIVIDUAL PARTICIPATION AND VOTE DECISION

Heckman coefficients reported	Heckman (1)		Heckman (2)		Heckman (3)		Heckman (4)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)	(4-1)	(4-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.266*** (0.054)		-0.282*** (0.082)		-0.331*** (0.082)		-0.341*** (0.079)
Complexity (ballot)	-0.357*** (0.106)		-0.384*** (0.121)		-0.479*** (0.134)		-0.502*** (0.119)	
Exclusion Restriction								
Approval Share	0.750 (0.542)		1.185 (0.774)		1.260* (0.761)		1.428* (0.733)	
Approval Share square	-1.918*** (0.506)		-2.026*** (0.646)		-2.208*** (0.763)		-2.149*** (0.755)	
Unobserved Factors								
rho	0.892		0.555		0.517		0.395	
Wald test (*p-value*)	0.000		0.033		0.005		0.001	
Canton dummies	No		Yes		Yes		Yes	
Referenda type dummies	No		Yes		Yes		Yes	
Year dummies	No		No		Yes		Yes	
Policy area dummies	No		No		No		Yes	
Observations	191669		191669		191669		191669	

Notes: The table provides the estimated coefficients of the Heckman selection model and establishes the negative and significant effect of complexity on voter's probability to participate and vote in favor of a proposition. The table also reports the correlation ρ between the error terms of both equations, as well as the corresponding p-values. All equations are estimated including individual controls as in Table 1. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

The participation bias arises due to observed and unobserved factors which affect both decisions. Table 4 reports estimates for the resulting participation bias based on the regression results in column (4) in Table 3. Whereas on average only 46.8% of those citizens who are against the proposition participated in the referendums, 67.8% of the citizens who are in favour of the proposition participated in the referendums. The resulting participation bias in the average referendum outcome equals roughly 11 ppts.

TABLE 4—PARTICIPATION BIAS OF THE VOTING OUTCOME

Voting Preferences and Behavior		
Voting Yes	$Pr(v=1\|\pi=1)$	0.525
Preferring Yes	$Pr(v=1)$	0.417
Participation Bias	$Pr(v=1\|\pi=1) - Pr(v=1)$	0.108
Participation Behavior		
Participation	$Pr(\pi=1)$	0.555
Participation of Yes-Voters	$Pr(\pi=1\|v=1)$	0.678
Participation of No-Voters	$Pr(\pi=1\|v=0)$	0.468

Notes: The table establishes the resulting participation bias of 10.8 percentage points and illustrates that potential yes-voters are more likely to participate than potential No-voters. The estimates are based on the model estimates in column (4) of Table 3.

Since coefficients in nonlinear models (especially when these coefficients are associated with variables appearing in both the selection and the outcome equation of a Heckman selection model) are difficult to interpret, we report average partial effects for our main variables of interest in Table 5. We also compare them with the respective results from the naïve single equation probit specifications as described in equations (1) and (2). The indirect effect of complexity on the vote outcome, $(v = 1|\pi = 1)$ – that can be identified with the Heckman approach – is positive. An increase of one standard deviation in complexity increases the average probability of voting in favor of a proposition by 2.3 percentage points (5th column in Table 5). Higher complexity reduces the turnout rate of potential no-voters more strongly than for potential yes-voters. However, the indirect effect is quantitatively not large enough to offset the opposing negative direct effect of complexity on the voting decision (-5.4 percentage points).

TABLE 5—PARTIAL EFFECTS OF HECKMAN MODELS

APE of ± 0.5 SD	Probit	Probit	Heckman		
	π=1	v=1|π=1	π=1	v=1|π=1	v=1
Complexity (proposition)		-0.046*** (0.013)		-0.054*** (0.013)	-0.051*** (0.011)
Complexity (ballot)	-0.082*** (0.020)		-0.082*** (0.019)	0.023*** (0.007)	

Notes: The table summarizes the average partial effects of a change of one standard deviation (centered, ± 0.5 SD) in each complexity measure on participation and voting behavior based on the single equation models (column (4) in Table 2, and column (8) in Table 1) and the Heckman selection model (column (4-1) and (4-2) in Table 3). The Heckman model allows for an indirect effect of the variable Complexity (ballot) on the vote outcome via altering the participation decision. This indirect effect increases the probability of voting in favor of a proposition by 2.3 percentage points if complexity changes by one SD (centered). Even though quantitatively important the indirect effect is outweighed by the negative direct effect of complexity on the probability of voting in favor of a proposition of -5.4 percentage points. All results are based on regression estimates using the estimation sample used in the Heckman regressions (n=191669).

Based on the estimates in Table 5, one might be tempted to accept the probit estimate for the average effect of complexity on voting behavior (-4.6 ppts) as a reasonable approximation for the Heckman estimates consisting of both direct (5.4 ppts) and indirect (2.3 ppts) effects of complexity. Yet, the direction of the bias of the probit estimates is systematically related to the complexity of the ballot. Figure 4 illustrates this by comparing the predictions for the vote outcome of the probit and the Heckman approach. Since probit estimates only based on the sample of voters neglect the participation decision of potential voters, they underestimate the probability of voting in favor of a proposition in cases in which complexity of the corresponding ballot is very high. This leads to lower participation and the neglected positive participation effect is strong, whereas they overestimate the probability of voting in favor of a proposition when ballot complexity is low quantitatively very low. The difference in the predictions between the probit and Heckman estimation results is quantitatively substantial and sometimes the probit estimator predicts that the proposition is accepted, while the Heckman estimator predicts that the proposition is rejected in the case of low complexity ballots. The opposite sometimes occurs for highly complex ballots.

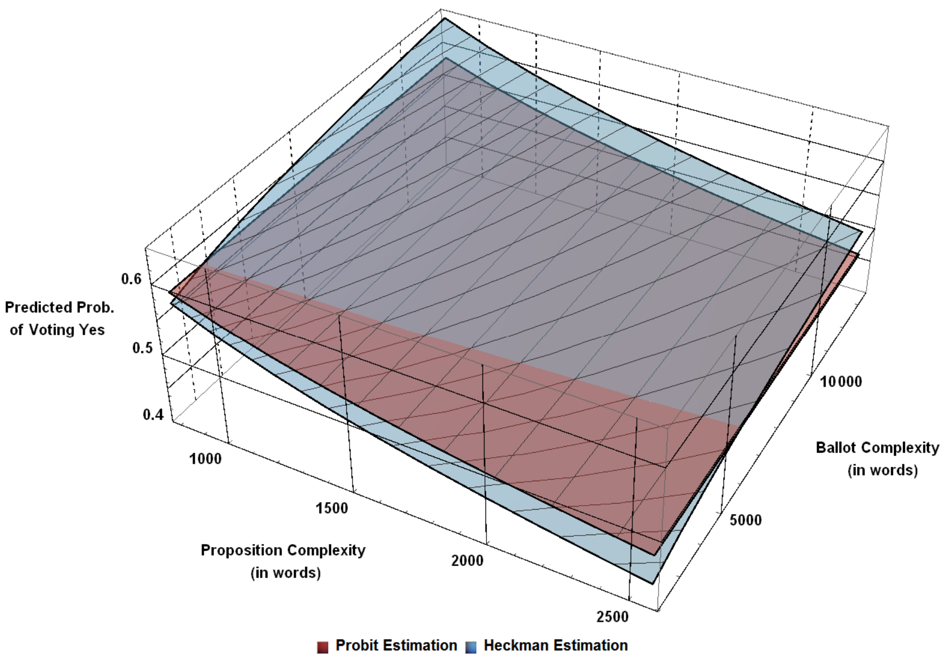

FIGURE 4. PROBIT VS. HECKMAN ESTIMATION RESULTS: PREDICTED PROBABILTIES OF VOTING IN FAVOR OF A PROPOSITION

Notes: This figure compares the results for the naïve probit estimations in section V.A with the Heckman estimation results in section V.B for the effect of complexity on individual voting behavior. We plot the average predicted probability of voting in favor of a proposition against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). The probit estimation (equation (2)) neglects the indirect effect. Therefore, the probit estimations are independent of ballot complexity. The estimates in the above figure are based on the estimates for model (4) in Table 3. The dotted vertical lines correspond to the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the respective complexity measure, respectively. For better readability, the information text axes are restricted to values between the 5^{th} and 95^{th} percentile.

The last column in Table 5 provides the estimate of the effect of proposition complexity on the voting behavior of the entire electorate ($\Pr(v = 1)$), including voters and non-voters. This estimate can be interpreted as the effect of complexity on the preference in favor of an approval of a proposition of the country's population. This effect can only be identified with the Heckman model.

In Table 4, the difference between the average probability of voting in favour of a proposition ($\Pr(v = 1|\pi = 1) = 52.5\%$) and the average probability of preferring an approval of a proposition ($\Pr(v = 1) = 41.7\%$) becomes evident. As denoted in the last column in Table 5, an increase of proposition complexity by one standard deviation reduces the average preference in the population for an approval of the proposition by 5.1 percentage points. This effect can be interpreted as the status quo bias in the entire population of a country, independently of the turnout decision.[32]

C. *Simulations of the policy impact*

We conclude the presentation of the Heckman estimation results with a policy simulation exercise. We investigate to what extent proposition complexity may decisively influence a referendum outcome, i.e. tilt the aggregate outcome from approval to rejection and vice versa. In our policy simulation we fix complexity at the ballot and proposition level at the 10^{th} percentile and calculate the

[32] With a single equation regression model (such as the probit model in equation (2)), which focuses on the behavior of voters, these results cannot be obtained.

individual predictions for each of the 191,669 observations in the sample. Afterwards, we repeat this exercise based on the 90th percentile.[33] We refer to these as *low* and *high* complexity scenarios.

In line with the results in Table 5, we find that the direct effect of proposition complexity (ignoring the participation effect) leads to a reduction in the probability of voting in favor of a proposition by 11.8 ppts. However, the participation effect partially offsets the decline in approval. The predicted participation rate decreases by 21.6 ppts. Since the participation rate of citizens who would vote against the proposition declines disproportionately, the participation effect of ballot complexity (via the participation effect) on the approval rate leads to an increase of 6.1 ppts. In total, an increase in ballot and proposition complexity from the 10th percentile to the 90th percentile causes the approval rate to decline by 5.6 ppts. Relying on a simple probit estimation (for the subsample of voters, ignoring the participation effect) would result in a predicted decline in the approval rate by 9.4 ppts. This implies an overestimation of the complexity effect by more than 67%.

In the final step of our simulation exercise, we average the individual predictions across propositions to obtain a collapsed data set containing the predicted approval rate for 223 propositions. Is complexity likely to alter the referendum outcome? We investigate in how many cases complexity causes the approval rate to decline below 50%. Figure 5 plots the cumulative distribution of the 223 propositions with respect to their (predicted) approval rate in the case of the low (red line) and high (blue line) complexity scenario. The left panel illustrates the joint complexity effect based on the Heckman approach.

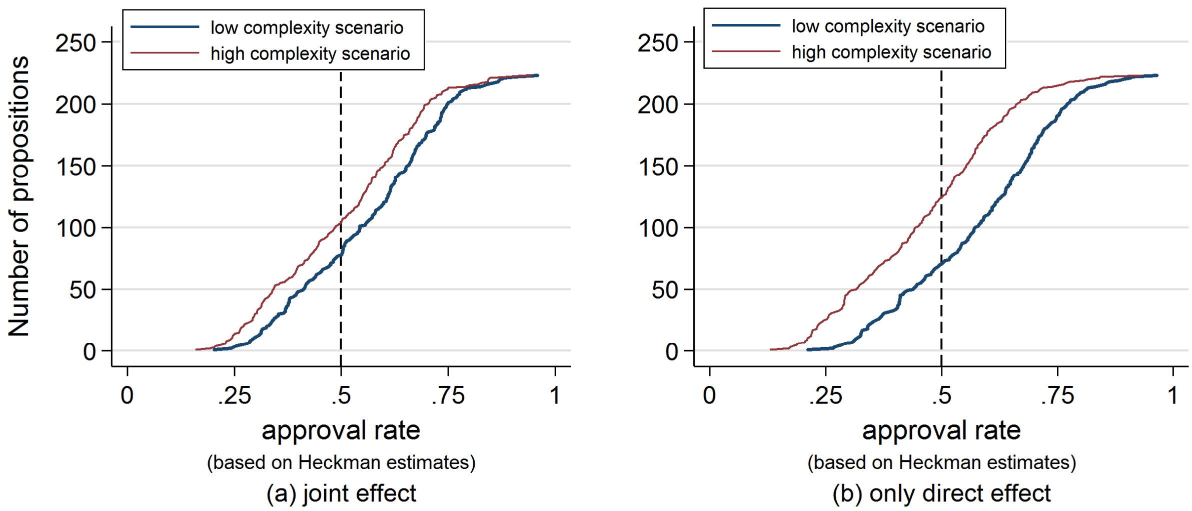

FIGURE 5. SIMULATION ON THE EFFECT OF COMPLEXITY ON APPROVAL

Notes: N=223. This figure illustrates the effect of an increase in the complexity from the 10th percentile (low complexity scenario) towards the 90th percentile (high complexity scenario). The figure is based on estimates of the approval rate for 223 propositions from 74 propositions.

The vertical distance between the red and the blue line at the 50% approval rate threshold (dashed vertical line) is 27. This means that 27 out of 223 propositions would have been rejected in the

[33] For the proposition complexity measure, the increase in complexity from the 10th to the 90th percentile represent an increase from 937 and 2346 words. Whereas an increase in the ballot complexity level from the 10th to the 90th percentile is associated with an increase from 2820 and 11635 words. The estimates in this section are based on the Heckman model (4) in Table 4.

high complexity scenario but not in the low complexity scenario. The right panel presents the difference in the approval rate if only the direct effect of complexity is at work, i.e. ignoring the offsetting participation effect. Without the participation effect, 54 propositions would have fallen below the 50% approval rate threshold. Again, the naïve probit model strongly exaggerates the effect of complexity on the number of tilted referendum outcomes: it predicting 44 additional rejected proposals. The difference between the predicted joint effect (Heckman model) and the probit model predication is again large (27 vs. 44 additional rejected propositions) and also statistically highly significant ($n = 223, t = 4.14, p < 0.001$).[34]

VI. Extensions

In this section, we investigate whether the complexity-induced status quo bias in the voting decision is indeed driven by the mechanisms proposed in the theoretical literature. If this were the case, we should observe that citizens with a lower education level – who are less informed and less capable of understanding the consequences of a proposition – are more likely to reject a proposition. In addition, we investigate whether a higher campaign intensity in the newspapers is able to mitigate the effect of complexity on status quo bias by lowering information costs and the likelihood of random errors.

A. *Complexity and education*

The availability of an objective complexity measure is particularly valuable when interacting proposition complexity with individual voter characteristics such as education.[35] Let $c_s(c_o, \mathbf{i_{obs}}, \mathbf{i_{uno}})$ denote the subjective complexity of a proposition which depends on the objective complexity of the proposition c_o, a vector of observable individual characteristics $\mathbf{i_{obs}}$ and a vector of unobservable individual characteristics (such as intelligence or cognitive skills) denoted by $\mathbf{i_{uno}}$. Highly educated people may be better able to deal with complex issues leading to differences in voters' reaction to complex propositions across education levels.[36] Voter's education is, however, also likely to be correlated with unobservable characteristics $\mathbf{i_{obs}}$. The corresponding interaction term between subjective complexity and education is $c_s(c_o, edu, \mathbf{i_{obs}}, \mathbf{i_{uno}}) \times edu$. It is, however, not clear how to interpret the estimate for this interaction term because as education changes, the perceived subjective complexity c_s changes as well.[37] We circumvent the problems related to a subjective complexity measure and use our objective text-based measure of complexity which is uncorrelated with individual voter characteristics. Thus, we estimate $c_o \times edu$ which has a clear interpretation.

---

[34] The test is based on a paired t-test. To correct the test procedure for clustering at the ballot level, we used a block bootstrap t-test relying on 999 replications.

[35] Related to this, Eichenberger and Serna (1996) state: "it is difficult to measure the complexity of an issue independently from the individuals' human capital" (p.140).

[36] We indeed find that subjective complexity is positively related to objective complexity and negatively related to education. Table A.8 in the appendix provides the empirical results.

[37] The only way how the level of subjective complexity can stay constant w.r.t. to a change of complexity is if we assume a simultaneous change in unobservable characteristics offsetting the effect of education on the level of subjective complexity.

Since educated people are more likely to understand the consequences of complex propositions, we formulate the following hypothesis: with increasing complexity, less educated voters are more likely to use the status-quo heuristic. To test this hypothesis, we re-estimate the Heckman selection model described in section III.B including an interaction term between proposition complexity and a university degree dummy which equals 1 for citizens with a university degree and 0 otherwise.[38] We estimate the following Heckman selection model:

$$(5) \quad Participate^* = \alpha\, Complexity + \lambda\, University + \gamma\, Complexity \times University + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\delta}'\mathbf{z} + u,$$
$$Participate = 1 \text{ if } Participate^* > 0, Participate = 0 \text{ otherwise},$$

$$(6) \quad Yes\text{-}vote^* = \alpha\, Complexity + \lambda\, University + \gamma\, Complexity \times University + \boldsymbol{\beta}'\mathbf{x} + \epsilon,$$
$$Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^* > 0, Yes\text{-}vote = 0 \text{ otherwise}.$$

Table 6 presents the estimation results. The upper part of the table reports regression coefficients; the lower part of the table reports the marginal effect for the variables of main interest.

TABLE 6—STATUS QUO BIAS AND THE INTERACTION BETWEEN COMPLEXITY AND EDUCATION

	Heckman (1)		Heckman (2)		Heckman (3)		Heckman (4)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)	(4-1)	(4-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.344*** (0.083)		-0.334*** (0.086)		-0.343*** (0.082)		-0.353*** (0.082)
Complexity (ballot)	-0.527*** (0.121)		-0.493*** (0.136)		-0.517*** (0.120)		-0.516*** (0.121)	
University degree			0.349*** (0.026)	0.241*** (0.040)	0.341*** (0.025)	0.217*** (0.035)	0.390 (0.399)	-0.521 (0.537)
Log Info Text (proposition) x Uni								0.100 (0.073)
Log Info Text (ballot) x Uni							-0.006 (0.047)	
Marginal Effects	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)
Complexity (proposition)		-0.128*** (0.033)		-0.131*** (0.036)		-0.128*** (0.032)		-0.128*** (0.032)
Complexity (ballot)	-0.168*** (0.038)	0.039*** (0.013)	-0.158*** (0.042)	0.055*** (0.016)	-0.164*** (0.037)	0.041*** (0.013)	-0.164*** (0.037)	0.041*** (0.013)
University degree			0.110*** (0.009)	0.057*** (0.011)	0.106*** (0.008)	0.055*** (0.011)	0.106*** (0.008)	0.053*** (0.011)
Topic dummies	Yes		No		Yes		Yes	
Observations	191669		191669		191669		191669	

Notes: The table reports the estimates of the interaction effect between the objective complexity measure and education (university degree vs. no university degree) and therefore indicates heterogeneity in the response to complexity for voters with different education levels. Probit coefficients are reported in the upper half of the table. Average marginal effects for the variables of interest are reported in the lower half of the table. The average marginal effect associated with the interaction term is illustrated Figure 6. All regressions are estimated with fixed effects for year, canton and referenda type and controls for individual characteristics. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

[38] Distinguishing citizens' education levels only w.r.t to university and non-university degree keeps the analysis tractable and ensures that we can interpret our education unambiguously w.r.t. high and low education.

Model (1) in Table 6 provides the estimation results that we obtain when we do not control for citizen's education.[39] A university degree dummy is introduced in model (2), (3) and (4). Citizens with a university degree have an ~11 ppts higher probability to participate in a referendum and a ~6 ppts higher probability to vote in favor of a proposition. The coefficient for the interaction term between university education and complexity is positive but insignificant. However, neither the size nor the sign nor the statistical significance can be interpreted for interaction effects in nonlinear models (Ai and Norton 2003; Berry, DeMeritt, and Esarey 2010; Greene 2010).[40] To quantify the size and statistical significance of the interaction effect we follow the suggestion by Greene (2010) and analyze the predicted probabilities of participation and yes-voting for citizens with and without a university degree. Figure 6 and Figure 7 present the corresponding graphs.

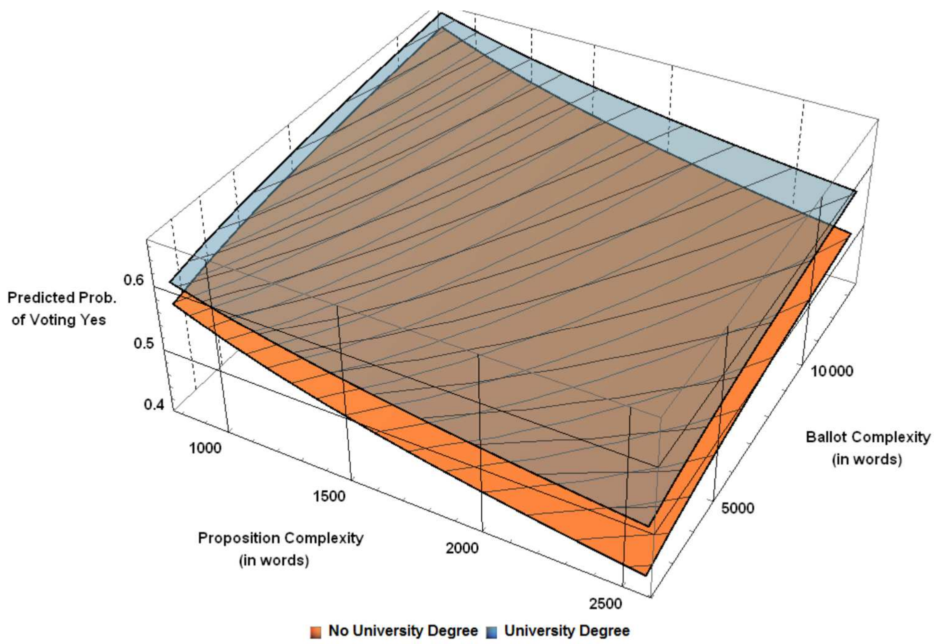

FIGURE 6. PREDICTED PROBABILTIES OF VOTING IN FAVOR OF A PROPOSITION - VOTERS WITH AND WITHOUT UNIVERSITY DEGREE

Notes: This figure illustrates the heterogeneous effect of complexity on voters' probability to vote in favor or against a proposition for voters with respect to different education levels. Voters without a university degree change their voting behavior stronger towards the status-quo as complexity rises. The figure plots the average predicted probability against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). The estimates in Figure 6 are based on the estimates of model (4) in Table 6. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

The slope of the surfaces in Figure 6 illustrates the marginal effect of complexity on the likelihood of accepting a proposition conditional on voters' education level. Panel (a) illustrates status quo bias for both education groups. Voters with a lower education level vote against a proposition more often than highly educated voters holding a university degree, even when propositions have a low level of complexity. However, the effect of complexity on the likelihood of rejecting a proposition is weaker for highly educated voters. The difference in the expected probability of voting in favor of a proposition

[39] This means that Model (1) in Table 6 only differs from Model (4) in Table 3 in the sense that the education is not controlled for.
[40] Hence, using a t-test to assess statistical significance of the coefficient of the interaction term is also invalid.

with a low level of proposition complexity is below 3%. This gap increases to about 8% for more complex propositions.

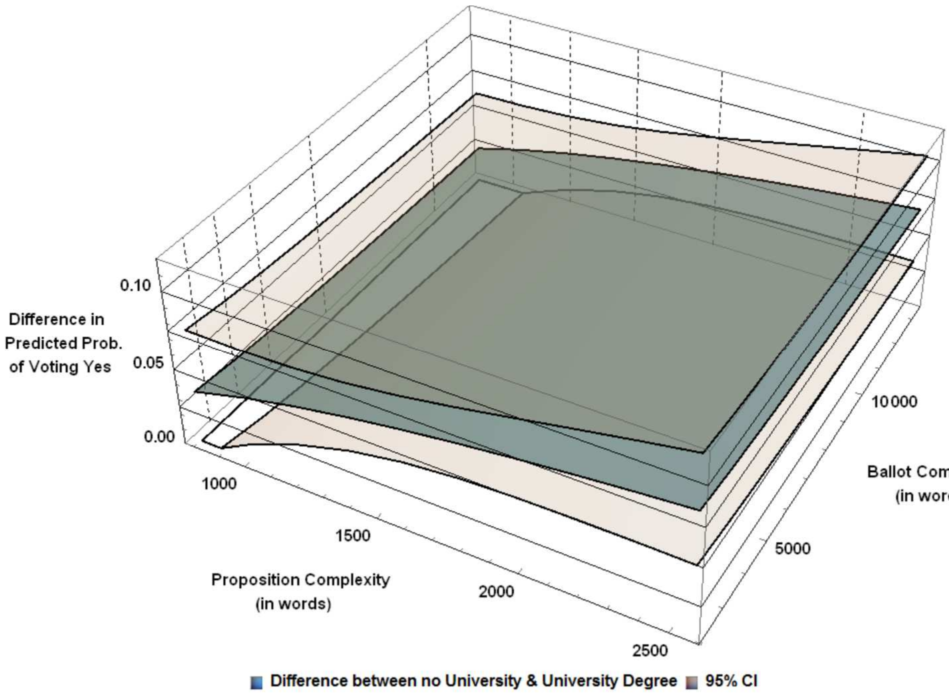

FIGURE 7. DIFFERENCE IN THE PREDICTED PROBABILTY OF „YES-VOTE" BETWEEN VOTERS WITH & WITHOUT UNIVERSITY DEGREE

Notes: This figure illustrates the "difference in difference" w.r.t. the response to complexity of citizens with and without a university degree. The estimates in Figure 7 are based on the estimates of model (4) in Table 6. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

Figure 7 plots the difference in the predicted probability of voting in favor of a proposition for university and non-university educated citizens. The predicted difference is not statistically different from zero for relatively easy propositions and increases to a statistically significant difference of more than 8 percent for more complex propositions. The gap between university and non-university citizens more than doubles if complexity increases. Based on the empirical results, we find lower educated citizens to be more affected by an increase in complexity than higher educated ones.

B. *Complexity and campaign intensity*

A second implication of the channels proposed in the theoretical literature is that a higher campaign intensity lowers information costs and therefore mitigates the status quo bias. In this subsection, we test whether the status quo bias is less pronounced if costs for the acquisition of information on the proposition are lower. Lowering costs to access information – for instance due to more information campaigns – is expected to mitigate perceived complexity. Our proxy for information costs is related to the campaign intensity measured by the number of ads related to a given proposition in the six major Swiss newspapers before the voting day.[41] We estimate the following Heckman model[42]:

[41] We thank Hans-Peter Kriesi for kindly providing access to this data.

[42] The assumed structure of the error term in all following Heckman estimations is similar to the one described in section III.B.

(7) $$Participate^{*} = \alpha\ Complexity + \lambda\ CampaignIntensity$$
$$+\gamma\ Complexity \times CampaignIntensity + \boldsymbol{\beta}'\mathbf{x}\ + \boldsymbol{\delta}'\mathbf{z} + u,$$
$$Participate = 1 \text{ if } Participate^{*} > 0, Participate = 0 \text{ otherwise,}$$

(8) $$Yes\text{-}vote^{*} = \alpha\ Complexity + \lambda\ CampaignIntensity$$
$$+\gamma\ Complexity \times CampaignIntensity + \boldsymbol{\beta}'\mathbf{x} + \epsilon,$$
$$Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^{*} > 0, Yes\text{-}vote = 0 \text{ otherwise.}$$

In model (1) in Table 7, we reproduce our main results for the subsample for which data on information costs is available.

TABLE 7—STATUS QUO BIAS AND THE INTERACTION BETWEEN COMPLEXITY AND CAMPAIGN INTENSITY

	Heckman (1)		Heckman (2)		Heckman (3)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.325*** (0.082)		-0.341*** (0.101)		-0.341*** (0.125)
Complexity (ballot)	-0.493*** (0.120)		-0.571*** (0.097)		-0.660*** (0.126)	
Campaign Intensity (Number of Ads x 100)			0.167*** (0.032)	0.035 (0.023)	-0.485 (0.319)	0.010 (0.170)
Log Info Text (proposition) x Campaign						0.003 (0.020)
Log Info Text (ballot) x Campaign					0.072** (0.035)	
Marginal Effects	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)
Complexity (proposition)		-0.121*** (0.032)		-0.126*** (0.038)		-0.125*** (0.042)
Complexity (ballot)	-0.155*** (0.037)	0.041*** (0.013)	-0.178*** (0.030)	0.041*** (0.013)	-0.184*** (0.034)	0.042*** (0.013)
Campaign Intenstity (Number of Ads)			0.052*** (0.010)	0.001 (0.008)	0.040*** (0.009)	0.002 (0.011)
Observations	180337		180337		180337	

This table illustrates that the intensity of the coverage of the campaign regarding to a proposition in the media does not change the general tendency of the effect of complexity on voter's probability to participate and vote in favor of a proposition. The variable *campaign intensity* is a measure of the number of ads related to a given proposition in the 6 major Swiss newspapers. All regressions are estimated with fixed effects for year, canton and referenda type and controls for individual characteristics. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

In models (2) and (3), we introduce the campaign intensity measure as well as an interaction term. A higher campaign intensity measured by an absolute increase in ads by 100 is associated with a statistically significant increase of 4 to 5.2 ppts in the participation rate. However, the effect of complexity on the approval rate is quantitatively low and insignificant. The results suggest that higher campaign intensity mitigates the negative effect of complexity on participation. This result is supported by Figure 8. The effect of complexity on participation rates for a representative low (25th percentile in campaign intensity distribution) proposition is compared to proposition characterized by high proposition complexity (75^{th} percentile in campaign intensity distribution).

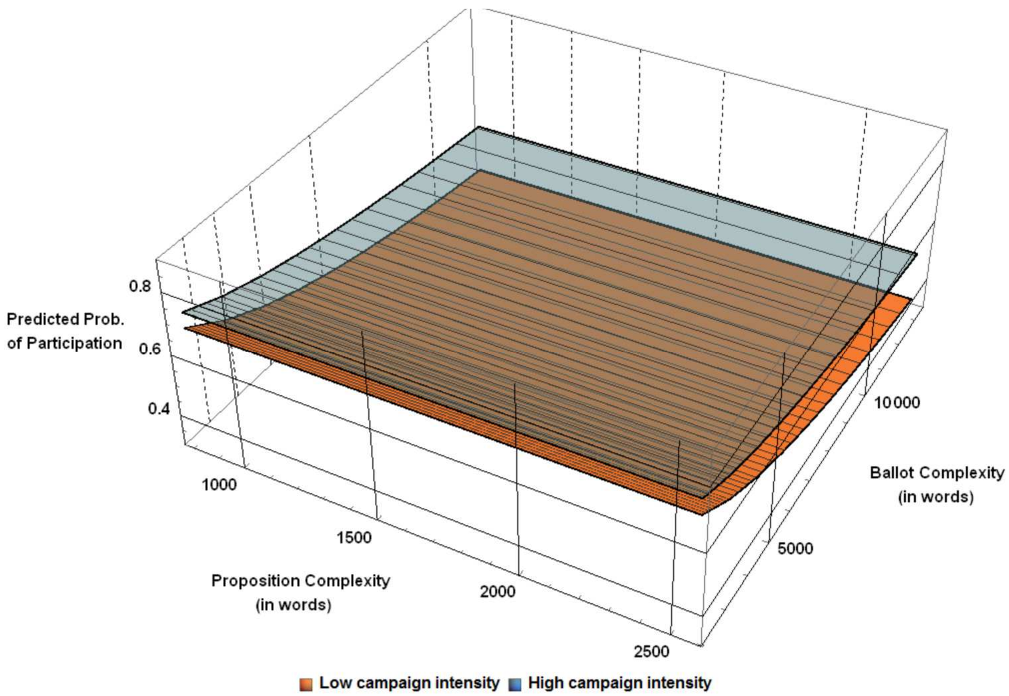

FIGURE 8. PREDICTED PROBABILTY OF PARTICIPATION W.R.T. CAMPAIGN INTENSITY

Notes: This figure illustrates the heterogeneous effect of complexity on voters' probability to participate in an election with respect to different levels of campaign intensity. Low (high) campaign intensity is represented by the value at the 25th (75th) percentile in the campaign intensity measure. As ballot complexity increases, citizens' participation rate decreases less, if campaign intensity is high. The figure plots the average predicted probability against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). The estimates in Figure 8 are based on the estimates of model (3) in Table 7. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

When ballot complexity increases, participation rates decrease more in case of relatively low campaign intensity. This suggests that a high campaign intensity may weaken the negative effect of complexity on participation.

Figure 9 plots the predicted probability of voting in favor of a proposition.[43] The estimates in the figure illustrate that the status quo bias is arising in roughly similar size in proposition with low and high campaign intensity, when proposition complexity increases. The intersection of both surfaces in Figure 9 along the ballot complexity axis illustrates the difference in the indirect effect transmitted through the heterogeneous response with respect to participation illustrated in figure 8.

[43] The participation decision is only affected by the ballot complexity channel. Hence, the two surfaces in Figure 8 are parallel with respect to proposition complexity. We still use the 3-dimensional illustration, to make comparison among the Figure 8 and Figure 9 easier.

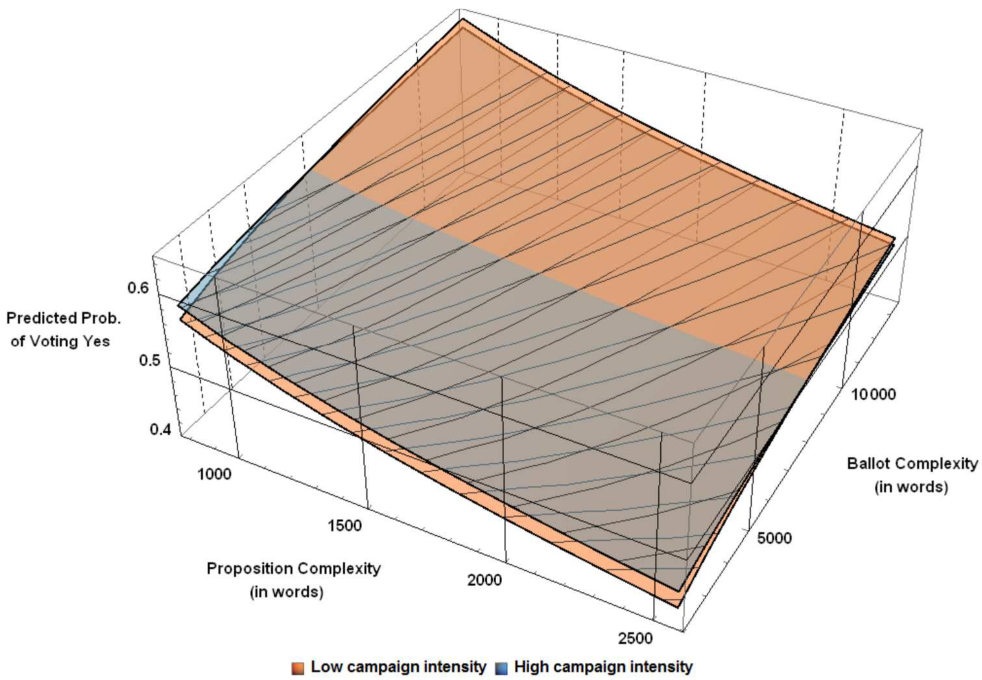

FIGURE 9. PREDICTED PROBABILTY OF „YES-VOTE“ W.R.T. CAMPAIGN INTENSITY

Notes: This figure illustrates the heterogeneous effect of complexity on voters' probability to vote in favor of a proposition in an election with respect to different levels of campaign intensity. Low (high) campaign intensity is represented by the value at the 25th (75th) percentile in the campaign intensity measure. The figure plots the average predicted probability against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). The estimates in Figure 9 are based on the estimates of model (4) in Table 7. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

The (indirect) participation effect of complexity mitigates the status quo bias in the vote outcome, because biased voters are more likely to stay at home. This effect is lower if campaign intensity is high, which can be seen by a lower slope of the blue surface with respect to ballot complexity in Figure 9. Even the quantitative effect is fairly small, still our results suggest that lower information cost not necessarily decrease the status quo bias because a higher share of relatively uninformed or uneducated voters is participating in the elections. This effect may be somewhat surprising, however it is within the range of possible expected results, if one includes the effect of the participation decision on the vote outcome. When exposed to a high campaign intensity, citizens may overestimate their competence, which would explain the increase in turnout, whereas a high campaign intensity does not appear to mitigate the increase in the status quo bias for complex propositions.

Why high campaign intensity does not mitigate the status quo bias.—We conclude the discussion on campaign intensity with an example that illustrates the magnitude of our results. Similar to the estimates presented in Table 5, we investigate how the complexity-induced change in voting behavior is mediated by different levels of campaign intensity. Figure 10 is based on the estimates of the Heckman model (3) reported in Table 7.

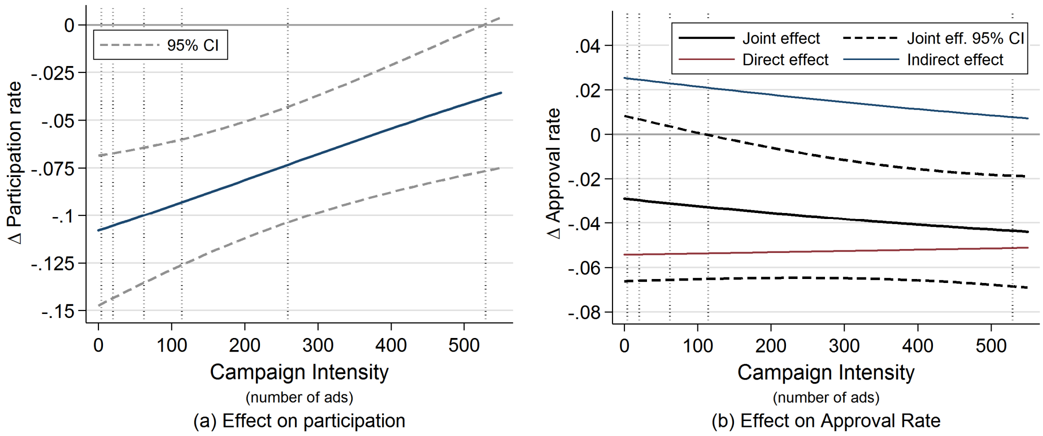

FIGURE 10. THE INTERACTION OF A CHANGE IN COMPLEXITY AND CAMPAIGN INTENSITY

Notes: This figure illustrates the effect of a change by one standard deviation (centered, ± 0.5 SD) in ballot and proposition on the participation and voting decision for different levels of campaign intensity. The vertical dotted lines correspond to the 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentile in the distribution of the campaign intensity measure. The dashed lines indicate the 95% confidence intervals. The figures are based on the estimates reported in Table 7.

Panel (a) in Figure 10 shows how the participation rate changes when ballot complexity increases by one standard deviation for different levels of campaign intensity. The complexity-induced decrease in the participation rate ranges from about 10 ppts (for low campaign intensities) to less than 5 ppts (for high campaign intensities). The total effect of complexity (i.e. the sum of the direct and the indirect effect) on the approval rate is illustrated in panel (b). Since the participation rate decreases with campaign intensity, the positive participation effect of campaign intensity on the approval rate – represented by the blue line in panel (b) – declines, while the negative direct effect is almost constant as campaign intensity increases. The total effect of complexity on the approval rate is illustrated by the black line; the effect slightly increases with increasing campaign intensity indicating that higher campaign intensity does not mitigate the arising status quo bias as complexity increases. This supports our findings above.

VII. Robustness tests

In this section, we discuss the results for five robustness tests. We only discuss the main findings here. Tables and figures are available in the appendix.

A. *Subjective proposition complexity*

First, we re-run our main estimations using an alternative complexity measure which is based on survey-based subjective perceptions of the complexity of individual propositions as stated by Swiss citizens.[44] In the post-referendum VOX surveys, citizens were asked whether it was difficult for them to

[44] We do not use the subjective complexity measure in the main analysis of this paper because it has several shortcomings compared to our objective complexity measure. First, it is potentially endogenous. For instance, non-participants may ex post justify their absenteeism with the excuse that it was difficult to decide. A second shortcoming arises due to the correlation of subjective complexity with citizens' characteristics (see results in Table A.8 in the appendix). While the difficulty to form an opinion about a proposition is clearly associated with the objective complexity of a proposition (c_o), the subjective measure is most likely confounded with observable ($\mathbf{i}_{\mathbf{obs}}$) and unobservable ($\mathbf{i}_{\mathbf{uno}}$) individual characteristics like education and income or intelligence, which determine the individual ability to understand the content of complex

form an opinion about the proposition (survey question: "*Did you find it rather easy or rather difficult given the provided information to imagine the impact of a yes- or no-vote on yourself with regard to this proposition?*"). The binary variable *difficulty to form an opinion* is a straightforward indicator for the subjective complexity of the proposition.[45]

All estimation results are qualitatively in line with our previous results reported in section V.B (see tables A.2 and A.3 in the online appendix for comparison). Voters who find it rather difficult to form an opinion are 11.5 ppts less likely to turn out and 4.7 ppts more likely to reject a proposition. This provides additional support that our booklet-based objective complexity measure is indeed a valid measure of the underlying complexity of a proposition and that complexity has a considerable effect on individual voting behavior in referenda.

B. *Does it matter whether voters read the information booklet?*

We argue that the length of the text describing a proposition in the official information booklet serves as proxy for the complexity of the proposition, independently of whether the voter has actually seen or bothered to read the booklet.[46] An alternative mechanism which may explain the link between the length of the information text and the complexity of the proposition works as follows: If voters read a complex description of a proposition in the information booklet, they might be overwhelmed by the length of the information text itself. This would open up the possibility that a voter is overwhelmed not because the underlying proposition is complex, but rather because the description of the proposition is complex.

If this alternative mechanism drives our main estimation results, a first empirical implication is that including a dummy that indicates whether a citizen has used the relevant information booklet should significantly affect our estimates for the effect of complexity on individual voting behavior. A second empirical implication of this potential mechanism is that the length of the information text should alter individual voting behavior only for those voters who actually use the information booklet. We test whether this mechanism is empirically relevant by estimating the following model:

$$\text{(9)} \qquad Participate^* = \alpha\, Complexity + \lambda\, Booklet + \gamma\, Complexity \times Booklet + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\delta}'\mathbf{z} + u,$$
$$Participate = 1 \text{ if } Participate^* > 0, Participate = 0 \text{ otherwise},$$

$$\text{(10)} \qquad Yes\text{-}vote^* = \alpha\, Complexity + \lambda\, Booklet + \gamma\, Complexity \times Booklet + \boldsymbol{\beta}'\mathbf{x} + \epsilon,$$
$$Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^* > 0, Yes\text{-}vote = 0 \text{ otherwise},$$

propositions. Therefore, the survey measure gives rise only to a subjective measure of complexity, which we denote by $c_s(c_o, \mathbf{i_{obs}}, \mathbf{i_{uno}})$. The confoundedness with variables such as education will make it difficult to use the subjective measure to identify heterogeneity in voters' response to complexity with respect to education, since interacting education with the subjective measure is problematic as we discussed in section VI.A. A problem we do not face when using our objective booklet-based complexity measure, which is independent of individual characteristics.

[45] In the participation equation, we use the more general survey question: *"In general, did you find it rather easy or rather difficult given the provided information to imagine the impact of a yes- or no-vote on yourself?"*

[46] The booklets provide a reputable and widely used information source for a majority of voters (Rohner 2012).

where *Booklet* is a dummy variable that indicates whether a voter has used the information booklet. If the alternative channel is relevant, the effect of our complexity measure on the status quo bias should diminish or at least decrease substantially.

Table A.4 in the online appendix presents the estimation results. The lower part of the table reports average marginal effects. The inclusion of the *booklet* dummy does not affect the joint effect of proposition and ballot complexity on the likelihood to vote against a proposition (model (2)). Model (3) includes the interaction term.[47] These effects are independent of proposition complexity and are therefore not directly linked to the effect of proposition complexity on the vote outcome.[48] As mentioned before, an insignificant point estimate for the interaction term does not indicate a nonlinear interaction effect between proposition complexity and booklet use (Greene 2010). Therefore, we plot the predicted probabilities of participation (figure A.2 in the online appendix) and supporting a proposition (figure A.3 in the online appendix) for booklet readers and nonreaders. The results show that regardless of whether voters have read or not read the booklet, proposition complexity has a similar effect on voting behavior. We conclude that the alternative channel is not empirically relevant and that the underlying complexity of a proposition is relevant for individual voting behavior.

C. *Proposition complexity and proposition importance*

More important propositions may be associated with longer booklet texts. Therefore, our complexity measure may be confounded. If so, the question arises whether importance and not complexity is the mechanism that drives the effect of our complexity measure on voting behavior. Voters may be more likely to participate in a referendum if they perceive the proposition to be important. In addition, the importance of a proposition likely mitigates the status quo bias because voters should be more likely to invest resources in understanding complex propositions when they perceive them as important. If our text-based complexity measure is indeed confounded with the importance of the proposition, our estimates should be downward biased, i.e. the true effect of complexity on the probability of abstaining from voting and rejecting a proposition is larger. To check for this possibility, we estimate the following models:

(11)
$$Participate^* = \alpha\, Complexity + \lambda\, Importance + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\delta}'\mathbf{z} + u,$$
$$Participate = 1 \text{ if } Participate^* > 0, Participate = 0 \text{ otherwise},$$

(12)
$$Yes\text{-}vote^* = \alpha\, Complexity + \lambda\, Importance + \boldsymbol{\beta}'\mathbf{x} + \epsilon,$$
$$Yes\text{-}vote = 1 \text{ if } Yes\text{-}vote^* > 0, Yes\text{-}vote = 0 \text{ otherwise},$$

[47] These results also provide interesting evidence of an effect that is not important for the mechanism that is tested here but which may be of interest to readers. Voters using the information booklet are 2 ppts more likely to vote in favor of a proposition and almost 18 ppts more likely to participate in a referendum (see models (2) and (3)) in Table A.4. This shows that when voters make an effort of collecting more information they are more likely to turn out and less likely to reject a proposition. This is in line with previous theoretical considerations on the role of information for voting behavior (Matsusaka 1995; Feddersen and Pesendorfer 1996), even though of course the decision to read the booklet is endogenous.

[48] The role of the media – official sources and private media sources like television and newspapers – in the decision process in direct democracies and the relation between proposition complexity and media use by citizens is left for further research.

where *Importance* is a categorical dummy variable (scaled from 0 (unimportant) to 10 (highly important) measuring two types of importance that a survey respondent attaches to a proposition. In particular, the VOX survey asks the following questions: "*How important are the consequences of the proposition for you personally?*" and "*How important are the consequences of the proposition for our country?*"

Since the survey questions about the perceived proposition importance were only asked in a subset of the referendums in our sample, we re-estimate our baseline model in column (1) in table A.5 in the online appendix to obtain a benchmark with this smaller sample. The estimates are not substantially affected by this change in sample size. When we control for proposition importance, the negative effect of (objective) complexity on voting behavior changes only slightly (see columns (2) and (3): from -0.14 to -0.157 and from -0.137 to -0.161). This suggests that proposition complexity, and not proposition importance, is the driving force behind the negative effect. The average marginal effect of proposition importance on the probability of voting in favor of a proposition is positive and highly significant as we expected. This suggests that the importance of a proposition indeed influences voting behavior. We conclude that while proposition importance makes it more likely that voters turn out and support a proposition, our complexity measure does not seem to be confounded by proposition importance and our baseline results remain unaffected.

D. *Endogeneity of proposition complexity: Strategic manipulation by the government?*

Another concern is that the government may manipulate the information text of a proposition to influence citizens' voting behavior. Note however that the scope for manipulation by changing the booklet text is limited due to legal restrictions on the drafting of the booklets (see section IV.B) as well as the absence of systematic differences in the effect of complexity on voting behavior between booklet users and non-users (see section VII.B above). Nevertheless, in this section we investigate whether the complexity of propositions (as measured by the length of the information text) is systematically smaller (larger) when the Swiss government supports (is against) a proposition. We estimate the following model:

(13) $$Complexity = \alpha\ GovernmentAgainst + \boldsymbol{\beta}'\mathbf{x} + u\ ,$$

where *GovernmentAgainst* is a dummy variable that is 1 when the National Council advises voters to reject a proposition and 0 when the National Council advises voters to support a proposition. The data for this variable is taken from official election data provide by the University of Bern, Institute of Political Science (2013). If indeed the government attempts to manipulate voters, we expect that the estimate for α is positive and significant.

The estimation results are reported in table A.6 in the appendix. We find that there is no significantly positive correlation between the government's voting recommendation and our complexity measure. We conclude that our complexity measure is not confounded by attempts of the government to influence referendum outcomes in its favor.[49]

E. *Potential survey bias in post-referendum surveys*

A general concern with post-election surveys is so-called survey bias, i.e. the possibility that voters' responses are not truthful. Comparing the aggregate results of the VOX survey (FORS - Swiss foundation for research in social sciences, 2012) and official election data (University of Bern, Institute of Political Science, 2013), Funk (2015) provides evidence for a significant difference in the share of yes-votes in the VOX survey compared to the official data in about half of the referendums. She also clearly indicates in her paper which propositions are affected by survey bias.[50] This allows us to test whether our estimation results on the effect of complexity on individual voting behavior may be contaminated by survey bias. Due to a sense of civic duty, voters may feel pressured to state that they participated in a referendum when they actually did not. It is, however, ex ante not clear how this may relate to the effect of complexity on turnout or especially on the actual voting decision. The main purpose of this exercise is to obtain somewhat "cleaner" estimates in our baseline models by excluding those propositions where Funk (2015) provides evidence for survey bias.

We re-estimate our baseline Heckman model for different subsamples, in which we systematically exclude propositions with the highest survey bias as identified in Funk (2015). Table A.7 in the online appendix summarizes the estimation results. In columns (1) to (4), we report estimates for different subsamples excluding 5%, 10%, 25%, and 50% of the propositions with the highest survey bias as reported in appendix table 2 in Funk (2015). Even if we exclude 50% of the propositions with the highest survey bias our results remain fairly unaffected. We conclude that there is no systematic influence of survey bias on the effect of complexity on individual voting behavior.

VIII. Conclusion

One concern that is often voiced by scholars about direct democracy is that citizens may not be sufficiently competent to decide about complex policy issues. This paper is the first to study how the complexity of propositions affects individual voting behavior in a direct democracy using a Heckman selection approach. Our dataset combines a novel complexity measure based on information provided in official information booklets with individual post-referendum survey data for 276 referenda at the federal level in Switzerland over the 1981-2010 period.

We find that the more complex a proposition is, the less likely are citizens to participate and, if they participate, to support a proposition. These findings are consistent with the idea of a status quo bias

[49] Note that our results do not prove that the government has no power at all to influence voting decision in a referendum. However, our results show that such manipulation does not appear to occur via the complexity of propositions. See Selb (2008) for details on the limits of of Swiss national government to influence the composition and content of ballots.

[50] Note that our objective is to explain voting behavior at the individual level rather than predicting exact aggregate referendum outcomes. Whenever we rely on aggregate data in our estimations we use official election data. In the estimations in Table 2, we rely on the share of approval votes to identify the Heckman selection model, where we took the data from the official election data from the University of Bern, Institute of Political Science (2013). Funk (2015) also shows that the survey bias varies across specific proposition topics. We include topic fixed effects meaning we rely on differences in the probability of voting yes within each topic category. Our point estimates usually increase and become more precise with topic fixed effects, possibly due to control for biases mentioned in Funk (2015).

when issues are too complex. More educated voters respond less to increasing complexity and exhibit a lower status quo bias than less educated voters. A higher campaign intensity increases turnout, but has no mitigating effect on the status quo bias.

Two competing mechanisms determine the impact of complexity on the status quo bias in the vote outcome. A *direct effect* – confronted with higher complexity, voters tend to vote in favor of the status quo, and an *indirect participation effect* – increasing complexity reduces the percentage of biased citizens (through vote abstentions) among the voters and therefore mitigates the status quo bias. We contribute to the literature by estimating the magnitude of both effects. We find that neglecting the indirect participation effect would overestimate the effect of complexity on the referendum outcome by almost 50%. The sheer magnitude of the opposing indirect effect calls attention to the importance of treating voting behavior as an outcome of two sequential choices.

The normative assessment of the decline in turnout triggered by proposition complexity depends on the trade-off between a *representative vote outcome* (Lijphart 1997) and an *informationally superior vote outcome* (Feddersen and Pesendorfer 1996). We provide evidence for the latter mechanism suggesting that policy measures like the introduction of mandatory voting may be counterproductive. Finally, our results suggest that improving the general level of education seems to be an appropriate measure to reduce the bias in the vote outcome.

References

Ai, Chunrong and Edward C. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters*, 80(1): 123–29.

Augenblick, Ned and Scott Nicholson. 2016. "Ballot Position, Choice Fatigue, and Voter Behaviour." *The Review of Economic Studies*, 83(2): 460–80.

Berry, William D., Jacqueline H. R. DeMeritt, and Justin Esarey. 2010. "Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?" *American Journal of Political Science*, 54(1): 248–66.

Bowler, Shaun and Todd Donovan. 1998. *Demanding choices. Opinion, voting, and direct democracy*. Ann Arbor: University of Michigan Press.

Bowler, Shaun, Todd Donovan, and Trudi Happ. 1992. "Ballot Propositions and Information Costs: Direct Democracy and the Fatigued Voter." *The Western Political Quarterly*, 45(2): 559–68.

Brunner, Eric, Stephen L. Ross, and Ebonya Washington. 2011. "Economics and Policy Preferences: Causal Evidence of the Impact of Economic Conditions on Support for Redistribution and Other Ballot Proposals." *Review of Economics and Statistics*, 93(3): 888–906.

2008, *Bundesgerichtsurteil 1C_412/2007 E.6.1*.

1978. Vol. 130, *Bundesratsbeschluss über das Ergebnis der Volksabstimmung vom 4. Dezember 1977 (Politische Rechte; Ausgleich des Bundeshaushalts)*.

Cox, Gary W. and Michael C. Munger. 1989. "Closeness, Expenditures, and Turnout in the 1982 U.S. House Elections." *American Political Science Review*, 83(01): 217–31.

Cronin, Thomas E. 1999. *Direct democracy. The politics of initiative, referendum, and recall*. Cambridge, Mass: Harvard University Press.

2013. "Datensatz der eidgenössischen Volksabstimmungen ab 1848." Année Politique Suisse. http://www.swissvotes.ch/uploads/codebuch.pdf.

Denver, D. T. and H. T. G. Hands. 1974. "Marginality and Turnout in British General Elections." *British Journal of Political Science*, 4(01): 17–35.

Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper and Row.

Eichenberger, Reiner and Angel Serna. 1996. "Random errors, dirty information, and politics." *Public Choice*, 86(1): 137–56.

Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *The American Economic Review*, 86(3): 408–24.

Freedman, David A. and Jasjeet S. Sekhon. 2010. "Endogeneity in Probit Response Models." *Political Analysis*, 18(2): 138–50.

Frey, Bruno S. 1994. "Direct Democracy: Politico-Economic Lessons from Swiss Experience." *American Economic Review*, 84(2): 338–42.

Funk, Patricia. 2015. "How Accurate are Surveyed Preferences for Public Policies? Evidence from a Unique Institutional Setup." *forthc. Review of Economics and Statistics*.

Funk, Patricia and Christina Gathmann. 2015. “Gender Gaps in Policy Preferences: Evidence from Direct Democracy in Switzerland.” *Economic Policy*, 30(81): 141–81.

Gerber, Elisabeth R. and Arthur Lupia. 1999. “Voter Competence in Direct Legislation Elections.” In *Citizen competence and democratic institutions*, ed. Stephen L. Elkin and Karol E. Sołtan. University Park, Pa.: Pennsylvania State University Press.

Gigerenzer, Gerd and Wolfgang Gaissmaier. 2010. “Heuristic Decision Making.” *Annual Review of Psychology*, 62(1): 451–82.

Greene, William. 2010. “Testing hypotheses about interaction terms in nonlinear models.” *Economics Letters*, 107(2): 291–96.

Heckman, James J. 1978. “Dummy Endogenous Variables in a Simultaneous Equation System.” *Econometrica*, 46(4): 931–59.

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*, 47(1): 153–61.

Hessami, Zohal. 2016. “How do voters react to complex choices in a direct democracy? Evidence from Switzerland.” *forthc. Kyklos*.

Hodler, Roland, Simon Luechinger, and Alois Stutzer. 2015. “The Effects of Voting Costs on the Democratic Process and Public Finances.” *American Economic Journal: Economic Policy*, 7(1): 141–71.

Huber, John D. and Charles R. Shipan. 2002. *Cambridge studies in comparative politics*, *Deliberate discretion? The institutional foundations of bureaucratic autonomy / John D. Huber, Charles R. Shipan*. Cambridge: Cambridge University Press.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias.” *The Journal of Economic Perspectives*, 5(1): 193–206.

Key, V. O. and Alexander Heard. 1984. *Southern politics in state and nation*. Knoxville: University of Tennessee Press.

Kriesi, Hanspeter. 2008. *Direct democratic choice. The Swiss experience*. 1st ed. Lanham, Md: Lexington, Books.

Krishnakumar, Jaya and Tobias Müller. 2012. “The political economy of immigration in a direct democracy: The case of Switzerland.” *European Economic Review*, 56(2): 174–89.

Lapalombara, Joseph G. and Charles B. Hagan. 1951. “Direct Legislation: An Appraisal and a Suggestion.” *American Political Science Review*, 45(02): 400–21.

Lijphart, Arend. 1997. “Unequal Participation: Democracy's Unresolved Dilemma Presidential Address, American Political Science Association, 1996.” *American Political Science Review*, 91(01): 1–14.

Lupia, Arthur. 1994. “Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections.” *American Political Science Review*, 88(1): 63–76.

Magleby, David B. 1984. *Direct legislation. Voting on ballot propositions in the United States*. Baltimore: Johns Hopkins University Press.

Matsusaka, John G. 1992. "Economics of Direct Legislation." *Quarterly Journal of Economics*, 107(2): 541–71.

Matsusaka, John G. 1995. "Explaining voter turnout patterns: An information theory." *Public Choice*, 84(1): 91–117.

Matsusaka, John G. 2015. "Ballot Order Effects in Referendum Elections." *University of Southern California Law School, Law and Economics Working Paper*(No. 176).

Matsusaka, John G. and Filip Palda. 1993. "The Downsian voter meets the ecological fallacy." *Public Choice*, 77(4): 855–78.

McCloskey, Deirdre N. and Stephen T. Ziliak. 1996. "The Standard Error of Regressions." *Journal of Economic Literature*, 34(1): 97–114.

Nicholson, Stephen P. 2003. "The Political Environment and Ballot Proposition Awareness." *American Journal of Political Science*, 47(3): 403.

Quattrone, George A. and Amos Tversky. 1988. "Contrasting Rational and Psychological Analyses of Political Choice." *American Political Science Review*, 82(03): 719–36.

Reilly, Shauna and Sean Richey. 2011. "Ballot Question Readability and Roll-Off: The Impact of Language Complexity." *Political Research Quarterly*, 64(1): 59–67.

Riker, William H. and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *The American Political Science Review*, 62(1): 25–42.

Rohner, Gabriela. 2012. *Schriften zur Demokratieforschung*. Vol. 4, *Die Wirksamkeit von Volksinitiativen im Bund, 1848-2010*. Zürich [etc.]: Schulthess.

Samuelson, William and Richard Zeckhauser. 1988. "Status Quo Bias in Decision Making." *Journal of Risk & Uncertainty*, 1(1): 7–59.

Selb, Peter. 2008. "Supersized Votes: Ballot Length, Uncertainty, and Choice in Direct Legislation Elections." *Public Choice*, 135(3-4): 319–36.

Stadelmann, David and Benno Torgler. 2013. "Bounded Rationality and Voting Decisions over 160 Years: Voter Behavior and Increasing Complexity in Decision-Making." *PLoS ONE*, 8(12): e84078.

Van de Ven, Wynand P. and Bernard M. van Praag. 1981. "The demand for deductibles in private health insurance: A probit model with sample selection." *Journal of Econometrics*, 17(2): 229–52.

2012. "Voxit data set." University of Lausanne.

Wooldridge, Jeffrey M. 2004. "Statistical significance is okay, too: comment on "Size Matters"" *The Journal of Socio-Economics*, 33(5): 577–79.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. 2nd ed. Cambridge, Mass: MIT Press.

Ziliak, Stephen T. and Deirdre N. McCloskey. 2008. *Economics, cognition, and society*, *The cult of statistical significance. How the standard error costs us jobs, justice, and lives*. Ann Arbor: University of Michigan Press.

Appendix

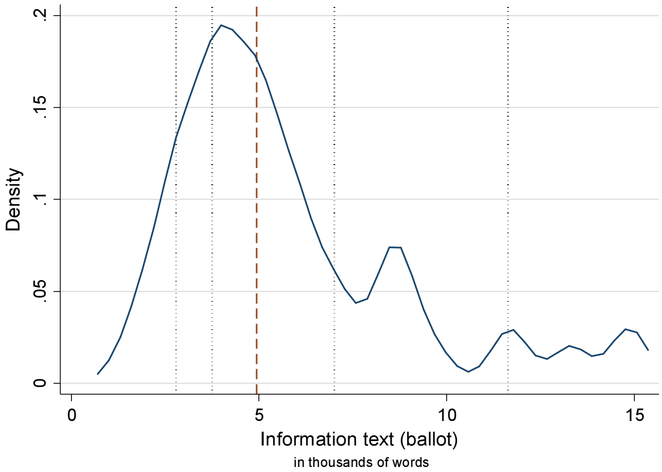

FIGURE A.1. VARIATION IN THE BOOKLET-BASED COMPLEXITY MEASURE

Notes: This figure depicts the variation in the booklet-based objective complexity measure aggregated on the ballot level which is used to identify the complexity of the ballot. It is based on a Gaussian kernel density plot with a half-width of 500 words. The dashed red line denotes the median of the complexity measure. The dotted grey lines correspond to the 10th , 25th , 75th , and 90th percentile. For better readability, the information text axis is restricted to values below the 99th percentile.

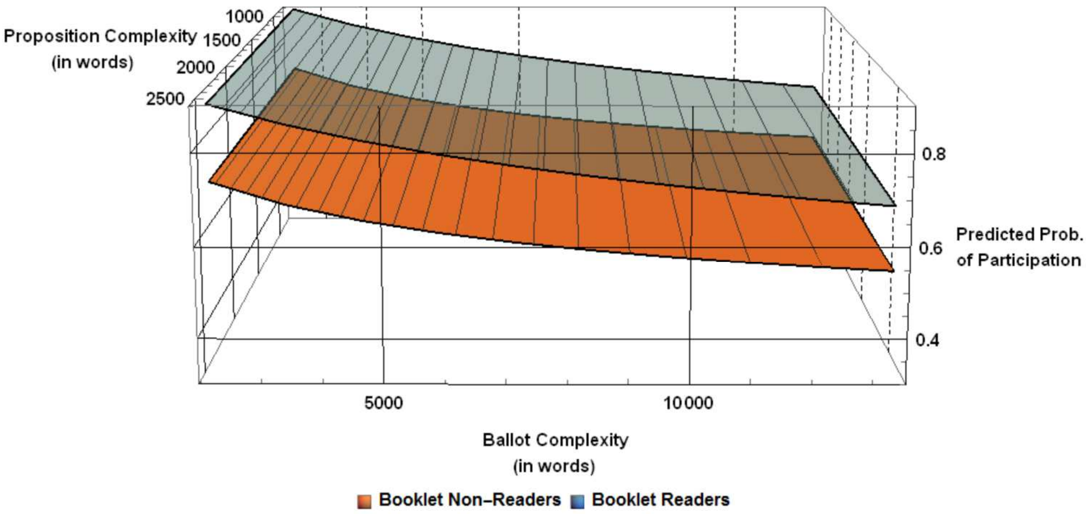

FIGURE A.2. PREDICTED PROBABILTY OF PARTICIPATION AND THE BOOKLET READING CHANNEL

Notes: This figure illustrates the heterogeneous effect of complexity on voters' probability to vote in favor of a proposition in an election with respect to different levels of campaign intensity. Low (high) campaign intensity is represented by the value at the 25th (75th) percentile in the campaign intensity measure. As ballot complexity increases, citizens' participation rate decreases less, if campaign intensity is high. The figure plots the average predicted probability against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). The estimates in Figure A.2 are based on the estimates of model (3) in Table A.4. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

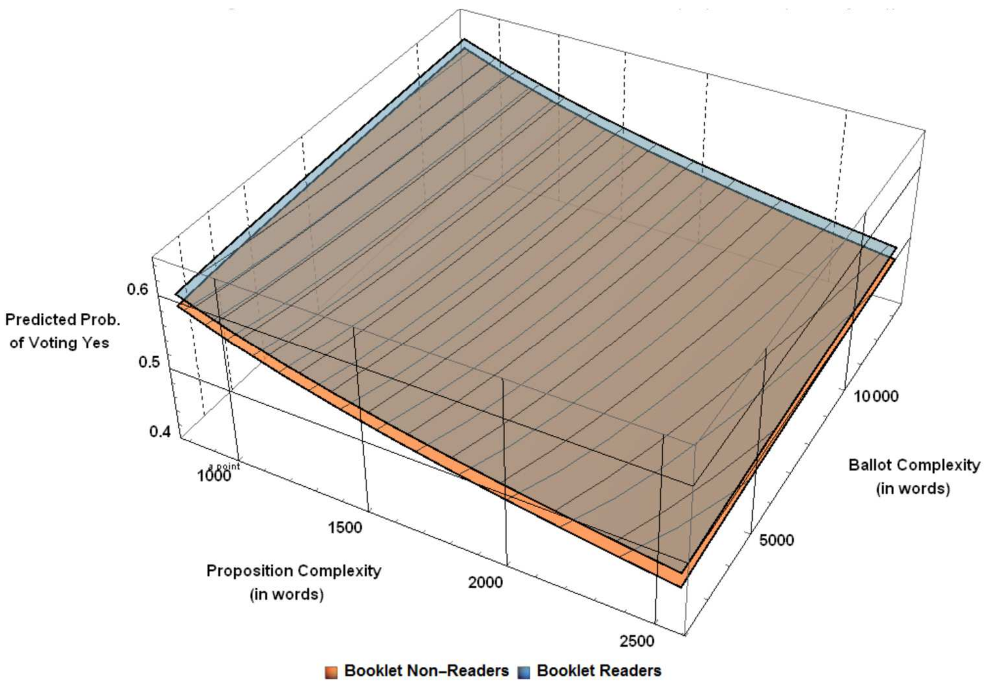

FIGURE A.3. BOOKLET READING, COMPLEXITY AND VOTING BEHAVIOR

Notes: This figure illustrates that booklet readers and non-readers voting behavior is fairly similar w.r.t. their reaction to complexity. The figure plots the average predicted probability against the complexity of the proposition (direct effect) and the complexity of the ballot (indirect effect). Figure A.3 is based on the estimates of model (3) in Table A.4. The dotted vertical lines correspond to the 10th, 25th, 50th, 75th, and 90th percentile of the respective complexity measure. For better readability, the information text axes are restricted to values between the 5th and 95th percentile.

TABLE A.1—DESCRIPTIVE STATISTICS

Voters and non-voters (N=191669)				Only voters (N=106817)			
	Mean	SD	Median		Mean	SD	Median
Log Complexity (ballot)	8.58	0.52	8.5	Log Complexity (proposition)	7.33	0.44	7.3
Complexity (ballot)	6.10	3.38	5.1	Complexity (proposition)	1.73	1.55	1.5
Rural	0.35	0.48	1	Rural	0.35	0.48	1
Female	0.50	0.50	1	Female	0.47	0.50	0
Age	47.00	17.39	44	Age	49.85	16.79	49
Education	2.71	1.54	2	Education	2.93	1.61	2
Proposition Knowledge (ballot)	4.80	2.98	4	Proposition Knowledge	1.62	0.60	2
Married	0.59	0.49	1	Married	0.65	0.48	1
Protestant	0.43	0.49	0	Protestant	0.45	0.50	0
Employed	0.61	0.49	1	Employed	0.60	0.49	1

Notes: The table reports the descriptive statistics for the sample used in regressions in Table 3. The log of the complexity measures are used in the estimations. We also report the descriptive statistics for the untransformed complexity measures (text length in thousands of words).

TABLE A.2—(EXPECTED) NARROW ELECTION DECISION AND PARTICIPATION (SUBJECTIVE COMPLEXITY MEASURE)

Probit coefficients reported	Dep. Var.: Participation			
	(1)	(2)	(3)	(4)
Difficult to Form an Opinion	-0.362***	-0.345***	-0.334***	-0.325***
	(0.018)	(0.017)	(0.016)	(0.016)
Exclusion Restriction				
Approval Share	1.757***	1.697**	1.554**	1.975***
	(0.638)	(0.678)	(0.625)	(0.585)
Approval Share square	-1.944***	-1.819***	-1.718***	-2.001***
	(0.629)	(0.659)	(0.628)	(0.582)
Canton dummies	No	Yes	Yes	Yes
Referenda type dummies	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes
Policy area dummies	No	No	No	Yes
Pseudo R-squared	0.138	0.143	0.161	0.168
Observations	181747	181747	181747	181747
p-value for joint significance of linear and quadratic terms in:				
Approval Share	0.005	0.018	0.020	0.003

Notes: The table illustrates the significant hump-shaped effect of the approval share and the probability to turn out in the elections. Regression coefficients based on probit regression are reported in all specifications. The variable approval share measures the share of yes-votes obtained from the official Swiss election data (University of Bern, Institute of Political Science 2013). The table also reports the p-value for the joint significance of the variable approval share and its squared term. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.3—HECKMAN SELECTION MODELS (SUBJECTIVE COMPLEXITY MEASURE)

	Heckman (1)		Heckman (2)		Heckman (3)		Heckman (4)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)	(4-1)	(4-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Difficult to Form an Opinion		-0.188***		-0.211***		-0.194***		-0.176***
(proposition)		(0.020)		(0.030)		(0.024)		(0.023)
Difficult to Form an Opinion	-0.327***		-0.376***		-0.365***		-0.355***	
(ballot)	(0.033)		(0.021)		(0.019)		(0.018)	
Exclusion Restriction								
Approval Share	1.091*		1.460*		1.236		1.962***	
	(0.632)		(0.833)		(0.786)		(0.748)	
Approval Share squared	-2.153***		-1.961***		-1.683**		-2.151***	
	(0.603)		(0.747)		(0.809)		(0.749)	
Unobserved Factors								
rho	0.879		0.388		0.299		0.233	
Wald test (*p-value*)	0.000		0.048		0.005		0.011	
Canton dummies	No		Yes		Yes		Yes	
Referenda type dummies	No		Yes		Yes		Yes	
Year dummies	No		No		Yes		Yes	
Policy area dummies	No		No		No		Yes	
Observations	166787		166787		166787		166787	

Notes: The table provides the estimated coefficients of the Heckman selection model and establishes the negative and significant effect of complexity on voter's probability to participate and vote in favor of a proposition. The table also reports the correlation ρ between the error terms of both equations, as well as the corresponding p-values. All equations are estimated with individual controls as reported in table A3.2, fixed effects for the ballot year, referenda type and the canton in which the eligible voter lives. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.4—STATUS QUO BIAS AND THE USE OF INFORMATION BOOKLETS

	Heckman (1)		Heckman (2)		Heckman (3)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.317***		-0.311***		-0.317***
		(0.088)		(0.087)		(0.081)
Complexity (ballot)	-0.697***		-0.614***		-0.551***	
	(0.148)		(0.141)		(0.154)	
Booklet-reader			0.864***	0.088***	2.060***	0.022
			(0.049)	(0.028)	(0.694)	(0.308)
Complexity (proposition) x Booklet-reader						0.009
						(0.040)
Complexity (ballot) x Booklet-reader					-0.139*	
					(0.082)	
Avg. Marginal Effects	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)
Complexity (proposition)		-0.114***		-0.111***		-0.112***
		(0.032)		(0.031)		(0.031)
Complexity (ballot)	-0.142***	0.017**	-0.114***	0.011*	-0.115***	0.011*
	(0.029)	(0.008)	(0.026)	(0.006)	(0.025)	(0.006)
Booklet-reader			0.177***	0.016**	0.177***	0.016**
			(0.010)	(0.007)	(0.010)	(0.007)
Observations	128101		128101		128101	

This table establishes the robustness of the negative and significant effect of complexity on voter's probability vote in favor of a proposition when controlling for the actual use of the booklet by the voter. The variable *using information booklet* is binary and equals one if the voter reports the use of the information booklet. All regressions are estimated with fixed effects for year, canton and referenda type and controls for individual characteristics. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.5—PERCEIVED IMPORTANCE OF PROPOSITIONS AND COMPLEXITY

	Personal Importance				Country Importance			
Aver. Marginal Effects reported	Heckman (1)		Heckman (2)		Heckman (3)		Heckman (4)	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)	(4-1)	(4-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.140***		-0.157***		-0.137***		-0.161***
		(0.048)		(0.047)		(0.048)		(0.047)
Complexity (ballot)	-0.158***	0.026**	-0.136***	0.027**	-0.155***	0.028**	-0.155***	0.027**
	(0.035)	(0.013)	(0.028)	(0.013)	(0.035)	(0.014)	(0.034)	(0.013)
Personal Importance			0.033***	0.018***				
			(0.002)	(0.003)				
Country Importance							0.013***	0.022***
							(0.001)	(0.003)
Observations	132022		132022		127439		127439	

Notes: This table establishes the robustness of the negative and significant effect of complexity on voter's probability vote in favor of a proposition when controlling for the perceived importance of the proposition. Both importance measures are measured on a scale from 0 to 10, where 0 = unimportant, 5 = medium importance, 10 = high importance. All regressions are estimated with fixed effects for year, canton and referenda type and controls for individual characteristics. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.6—GOVERNMENT RECOMMENDATION AND COMPLEXITY

	Dep. Var.: Complexity				
	(1)	(2)	(3)	(4)	(5)
Government Recommendation (Federal Council Advice = No)	-0.197 (0.151)	-0.163 (0.129)	-0.370 (0.729)	-0.455 (0.572)	-0.140 (0.713)
Individual Controls	No	Yes	Yes	Yes	Yes
Canton dummies	No	Yes	Yes	No	Yes
Referenda type dummies	No	No	Yes	Yes	Yes
Year Dummies	No	No	Yes	Yes	Yes
Topic dummies	No	No	No	Yes	Yes
R-squared	0.006	0.102	0.109	0.217	0.329
Observations	205175	205175	205175	205175	205175

Notes: This table shows that there is no statistically significant relationship between the government's voting recommendation and our complexity measure. The dependent variable is the number of words used in the information text per proposition. The results suggest that the government does not strategically manipulate the information text of a proposition. OLS coefficients are reported in columns (1) to (5). The variable *government recommendation* is binary and equals one if the government recommends voting against a proposition. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.7—EXCLUSION OF PROPOSITIONS LIKELY TO SUFFER FROM SURVEY BIAS

	Heckman (1) exclude 5%		Heckman (2) exclude 10%		Heckman (3) exclude 25%		Heckman (4) exclude 50%	
	(1-1)	(1-2)	(2-1)	(2-2)	(3-1)	(3-2)	(4-1)	(4-2)
Dep. Variable:	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote	Participation	Yes-Vote
Complexity (proposition)		-0.315*** (0.079)		-0.312*** (0.078)		-0.298*** (0.070)		-0.280*** (0.074)
Complexity (ballot)	-0.536*** (0.115)		-0.526*** (0.117)		-0.523*** (0.117)		-0.597*** (0.113)	
Exclusion Restriction								
Approval Share	1.605** (0.669)		1.537** (0.696)		2.293*** (0.821)		2.871*** (0.895)	
Approval Share squared	-2.494*** (0.698)		-2.450*** (0.736)		-3.145*** (0.850)		-3.808*** (0.901)	
Avg. Marginal Effects	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)	Pr(π=1)	Pr(v=1|π=1)
Complexity (proposition)		-0.119*** (0.032)		-0.118*** (0.031)		-0.111*** (0.027)		-0.104*** (0.029)
Complexity (ballot)	-0.167*** (0.035)	0.049*** (0.014)	-0.164*** (0.036)	0.045*** (0.013)	-0.163*** (0.036)	0.043*** (0.011)	-0.186*** (0.035)	0.051*** (0.016)
Unobserved Factors								
rho	0.420		0.398		0.381		0.394	
Wald test (*p-value*)	0.000		0.001		0.000		0.002	
Observations	183755		176675		154548		112335	

Notes: The table establishes the robustness of the main results in Table 3 with respect to a potential survey bias as described in Funk (2015). In columns (1) to (4), estimates for different subsamples excluding the 5%, 10%, 25%, and 50% of the propositions with the highest survey bias as reported in appendix table 2. Heckman coefficients are reported in the upper half of the table. Average marginal effects for the variables of interest are reported in the lower half of the table. Complexity represents the log of the number of words used in the information text in the official booklets per proposition or aggregated at the ballot level. All regressions are estimated with fixed effects for year, canton and referenda type and controls for individual characteristics. Estimations that use the participation dummy as the dependent variable include complexity at the ballot level; estimations that use the yes-vote dummy as the dependent variable include complexity at the proposition level. Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A.8—SUBJECTIVE COMPLEXITY AS FUNCTION OF OBJECTIVE COMPLEXITY AND EDUCATION

Dep. Var.:	Difficulty to Form Opinion (Ballot)				Difficulty to Form Opinion (Proposition)			
Avg. Marginal Effects reported	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Complexity (ballot)	0.019**	0.092***	0.092***	0.087***				
	(0.009)	(0.023)	(0.022)	(0.018)				
Complexity (proposition)					0.049***	0.075***	0.084***	0.086***
					(0.018)	(0.014)	(0.016)	(0.013)
Education		-0.042***	-0.042***	-0.042***		-0.030***	-0.029***	-0.029***
		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Individual Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Canton dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Referenda type dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes	No	No	Yes	Yes
Policy area dummies	No	No	No	Yes	No	No	No	Yes
Observations	182639	182639	182639	182639	170094	170094	170094	170094

Notes: This table establishes the positive effect of objective complexity on subjectively perceived complexity on the ballot as well as on the proposition level. In all specifications, Education is negatively correlated with the subjective complexity dummies. The results suggest subjective complexity can be described as function of objective complexity and individual characteristics like education. Average marginal effects based on probit regressions are reported in all columns. The variables *complexity (ballot)* and *complexity (proposition)* refer to the objective complexity measure based on the log of the word count. Alternative specifications using the untransformed word count produce similar results. The variable *education* ranges from 1 (mandatory school) to 6 (university). Heteroskedasticity-robust standard errors clustered at the ballot level are reported in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.