To what extent are citizens able to distinguish between fulfilled and unfulfilled election pledges? What explains variation in the extent to which citizens are able to do so accurately? The answers to these questions are central to the idea of promissory representation (Mansbridge, 2003, p. 515), which is found in the responsible party model and the mandate theory of democracy (Downs, 1957; Klingemann, Hofferbert and Budge, 1994; Grossback, Peterson and Stimson, 2005; McDonald and Budge, 2005). Promissory representation holds that parties make commitments during election campaigns and attempt to follow through on those commitments if they enter government after elections. Citizens' capacity to respond accurately to policy performance is as vital as parties' behaviour to ensuring a strong democratic chain of command and control. Without such a capacity on the part of citizens, the responsible party model would fail. For VO Key (1966), the responsible electorate, which rewards and punishes parties according to those parties' performance in government, is the counterpart to responsible parties. If voters are to respond in this way, they must be able to distinguish accurately between promises that were kept and those that were broken.

Existing research on election pledges mainly describes and explains variation in actual fulfilment, rather than citizens' evaluations of fulfilment. The findings suggest a puzzle that we address. Generally speaking, well above 50 percent of pledges are fulfilled at least partially, and for some single-party governments the figure is above 80 percent (Pomper and Lederman, 1980; Rallings, 1987; Royed 1996; Thomson *et al.*, 2017). However, in most countries the majority of citizens believe that parties break their promises. In the single-party governments that are typical of the United Kingdom, researchers typically find that above 70 percent of election pledges are fulfilled (Rallings, 1987; Royed, 1996). By contrast, only a

minority of UK respondents agree with the statement "People we elect as MPs try to keep the promises they have made during the election" (ISSP, 2008).¹

Few studies have addressed this puzzle to date. Through in-depth interviews with Swedish citizens, Naurin (2011) found that citizens' generally hold a broad and amorphous understanding of pledges. By contrast, published research on pledge fulfilment is based on clear definitions of election pledges that distinguish these statements from campaign rhetoric. Thomson's (2011) study of citizens' evaluations of pledge fulfilment in Ireland and Naurin and Oscarsson's (2017) study of Sweden incorporated this insight. Rather than ask citizens general questions about the fulfilment of pledges, which allow respondents to define pledges as they will, they asked citizens to assess the fulfilment of specific pledges that had been made in a previous election campaign. The headline findings were that citizens are able to distinguish between pledges that were in fact fulfilled or unfulfilled, and that their evaluations are also shaped by a range of individual-level characteristics. Given the scarcity of research on this important topic for representative democracy, the new data we present on British citizens' evaluations of pledge fulfilment are welcome.

We examine whether citizens' trust in governing parties affects their evaluations of the extent to which parties keep their promises. Following Lenard's (2008) distinction between the concepts of mistrust and distrust, we argue that distinct

Other approaches have been used to examine the link between campaign statements and subsequent government policies. These include comparisons of the general direction of campaign commitments and subsequent policies (e.g. Stokes, 2001), and of thematic emphases in election programs and subsequent public spending priorities (e.g. Klingemann, Hofferbert and Budge, 1994). The overall conclusion is that elected representatives generally follow through on their commitments.

aspects of trust refer firstly to a knowledge-based and vigilant responsiveness to actual government performance and, secondly, to a heuristic based on a general expectation about performance, which is at most weakly related to actual performance. Lenard (2008) defines mistrust as a healthy vigilance, which implies that mistrustful citizens are well informed and able to identify fulfilled pledges as fulfilled and unfulfilled pledges as unfulfilled. Following this line of reasoning, we examine and find evidence of an interaction between actual pledge fulfilment, citizens' knowledge about politics, and citizens' evaluations of pledge fulfilment. Actual performance, in terms of whether parties actually fulfilled their pledges, has the strongest effect on the evaluations of the most knowledgeable citizens, although it also impacts on the evaluations of the less knowledgeable.

Distrust, by contrast, equates with cynicism and the expectation of betrayal and disappointment. Distrust is a negative heuristic that people use to inform their evaluations, which accords with recent research that treats trust as a heuristic (Hetherington, 2004; Rudolph and Evans, 2005). We use a survey experiment in the 2014-17 British Election Study, which alters the salience of heuristic thinking based on citizens' pre-existing levels of distrust (BES; Fieldhouse, Green, Evans, Schmitt and Van der Eijk, 2015). When the salience of distrust-based heuristic thinking is raised (lowered), respondents' pre-existing levels of distrust have a stronger (weaker) effect on their evaluations of pledge fulfilment. We discuss how our research design and findings shed light on the intricate ways in which distinct aspects of trust impact on citizens' evaluations.

Two aspects of trust: The effects of healthy vigilance and heuristic thinking on citizens' evaluations

The dependent variable is citizens' evaluations of the fulfilment of specific election pledges. If citizens have a healthy mistrust of governing parties, which makes them attentive to government performance when forming their assessments, then whether or not the pledge was in fact fulfilled will be one of the most important factors shaping their evaluations. Lenard (2008) argues that one component of trust, or rather the lack of trust, is mistrust, which implies a healthy vigilance on the part of citizens regarding the use and abuse of power by leaders. This idea is also found in the argument that democracy thrives best when citizens are not overly trusting of their leaders (see Maloy, 2009, for a review). This compels leaders to behave appropriately and enables citizens to detect poor performance when it occurs.

Actual policy performance in our study refers to whether or not the election pledge was in fact fulfilled. Pledges are campaign statements that are specific enough for people to make reasonably objective assessments of pledge fulfilment. For example, one of the pledges we examine is a statement in the 2010 Liberal Democrats' manifesto to raise the tax-free allowance on income to £10,000 in the 2011 tax year. The tax-free allowance was indeed raised to £10,000 within the governing period following the 2010 election, although not yet in 2011, as stated in the pledge, making it a partially fulfilled pledge.

Citizens' information resources are relevant to their evaluations of pledge fulfilment, and in particular to the extent to which their evaluations are affected by actual performance. People who are more knowledgeable about politics hold a greater amount and higher quality of factual information about politics and are better able to

identify the relevance of new information. They are able to identify the relevance of current policy developments to parties' previous pledges. This implies that citizens with greater political knowledge make more accurate evaluations of pledge fulfilment. Similarly, Zaller's (1992) model of opinion formation posits that citizens with greater political awareness are more inclined to absorb new information about politics and policies.

Existing research on citizens' information resources arrives at different conclusions regarding their ability to incorporate relevant facts about policies into their evaluations, and therefore their ability to play the role of a vigilant citizenry effectively. Some studies find that people are generally unaware of the most basic facts about their political systems and representatives (e.g. Delli Carpini and Keeter, 1996; Milner, 2001; Althaus, 2003). This suggests we should not expect actual pledge fulfilment to affect citizens' evaluations. Other studies find that while citizens forget much detailed information, they have the cognitive capacities to make informed judgments (e.g. Inglehart, 1977; Dalton, 1988; Lodge, Steenbergen and Brau, 1995). Furthermore, studies based on macro models of public opinion often conclude that aggregate differences in public opinion are associated with meaningful differences in public mood and relevant policies (Page and Shapiro, 1992; Stimson, MacKuen and Erikson, 1995; Erikson, MacKuen and Stimson, 2001). At the aggregate level, public opinion is more responsive to actual performance than individuals' lack of knowledge suggests.

Lenard (2008) defines as second aspect of trust in terms of distrust, which is distinct from "mistrust". Distrust equates with cynicism and the expectation of betrayal and disappointment. Distrustful citizens generally do not expect parties to perform well in government. Citizens who are distrustful of governing parties are

likely to say that pledges are unfulfilled, regardless of what governments actually did. Likewise, according to this definition, trustful citizens are likely to say that pledges are fulfilled regardless of actual performance. This definition of trust understands the concept as a heuristic that people use to form their evaluations (Hetherington, 2004; Rudolph and Evans, 2005).

Distrust affects citizens' evaluations through heuristic thinking, not through the absorption of information on actual performance. Distrustful citizens' evaluations are affected by the common stereotype of promise-breaking politicians. By contrast, people with high levels of trust generally expect election pledges to be carried out, because such behaviour corresponds to their positive expectations of governing parties' behaviour. One mechanism through which the trust heuristic affects citizens' evaluations is that it provides a cue in the context of limited information. Even wellinformed citizens may lack the detailed information required to evaluate the fulfilment of specific election pledges. In the absence of factual information, trust or distrust in governing parties provides a guideline for making evaluations. Another mechanism concerns the inherent room for judgement in deciding whether or not a pledge is fulfilled. Although pledges are defined as statements that are specific enough for reliable evaluations of fulfilment to be made, some room for interpretation is inevitable. Again, the trust as distrust heuristic provides citizens with a means of making those interpretations. In the absence of complete and unambiguous information on pledge fulfilment, trusting citizens are more likely to evaluate promises as kept than are distrustful citizens.

Priming people to think about parties' promises, as distinct from policy proposals in a more general sense, should raise the salience of heuristic thinking.

When relying on heuristic thinking, distrustful citizens should evaluate pledges as

unfulfilled, while trusting citizens should evaluate pledges as fulfilled, regardless of actual performance. To test this, we vary the wording of questions posed to respondents. This follows the classic structure of priming experiments as pioneered by Sniderman and Piazza (1993), in which question wording is used to raise or lower the salience of a frame. The survey presents two groups of randomly selected respondents with different versions of similar questions. The first set of questions mentions "parties" and "promises", and then goes on to give details of six specific election pledges. The second set of questions simply refers to "proposals" and then goes on to refer to the contents of the same six election pledges, but without stating that these proposals were in fact parties' election pledges. We expect the "partypromise" version of the questions to prime citizens to use heuristic thinking when making their evaluations. We therefore expect distrustful citizens who receive the "party promise" treatment to give more negative evaluations compared to both more trusting citizens who receive the same treatment and to other distrustful citizens who receive the alternative "proposal" treatment. Varying the question wording is not intended to alter respondents' level of distrust. Rather, the question wording alters the salience of distrust as a heuristic for informing their evaluations.

We focus on trust in governing parties, rather than government or politicians in general. Other researchers have defined the object of trust more broadly. For instance, according to Miller "[p]olitical trust can be thought of as a basic evaluative or affective orientation toward the government" (1974, p. 952; see also Rudolph and Popp, 2009, p. 335). Our narrower definition is appropriate given our interest in citizens' evaluations of campaign commitments made by particular parties, rather than the government's policy performance in general.

Our main propositions concerning the effects of trust, defined in terms of mistrust and distrust, on citizens' evaluations are as follows.

Trust defined in terms of mistrust and healthy vigilance: Actual performance and political knowledge

H1a: Citizens are more likely to say that a pledge was fulfilled if the pledge was actually fulfilled.

H1b: The more knowledge citizens have about politics, the more their evaluations of pledge fulfilment are informed by whether or not the pledge was actually fulfilled.

Trust defined in terms of distrust and heuristic thinking: Question wording and respondents' self-reported trust in governing parties.

H2a: Citizens are more likely to say that a pledge was fulfilled if they are presented with questions about the adoption of proposals rather than the keeping of promises by parties.

H2b: The more citizens trust governing parties, the more likely they are to say that a pledge was fulfilled.

H2c: This effect of citizens' trust in governing parties on their evaluations is significantly greater when they are presented with questions about the keeping of promises by parties than with questions about the adoption of proposals.

While these propositions are our main theoretical interest, we recognise that other factors also matter, notably, party identification, personal circumstances and media attention. We control for these other explanations, but the available data do not support a detailed analysis of them all. Party identification is likely to affect citizens' evaluations of pledge fulfilment, whether defined in terms of the traditional or revisionist views that frame many recent studies of the concept (e.g. Bartels *et al.*, 2011). According to the classical view, party identification is an affective bond

between voters and their preferred parties that it is formed in childhood, and functions as a perceptual screen through which voters view and interpret the political world (Campbell et al., 1960, p. 133). According to the revisionist view of party identification, voters' party identification may change over time as a consequence of their evaluations of their parties' policy performance (Fiorina, 1981). The observable implications of these two views for citizens' evaluations of pledge fulfilment are the same. Voters who identify with governing parties are more likely to give positive evaluations of those parties' pledge fulfilment than non-identifiers, because this is most consistent with their existing identifications. Likewise, for voters who identify with opposition parties, their existing identifications are most consistent with negative evaluations of governing parties' pledge fulfilment. The classical and revisionist views differ with respect to the consequences of citizens' evaluations of policy performance on their future levels of party identification, a topic that is beyond the scope of our present study. Since our data are collected at one time point, we cannot discount the possibility that at least some of the party identification we observe is a consequence rather than a cause of citizens' evaluations of pledge fulfilment, as the revisionist view would suggest. We therefore include party identification as a control variable.

Respondents' personal circumstances may be another set of explanatory variables (Funk and Garcia-Monet, 1997; Duch, Palmer and Anderson, 2000, p. 638). To the extent that personal circumstances are relevant to the pledges we examine, these are associated with the demographic variables for which we control: occupation, education, age and gender. We include pledge controls for each of the pledges, since some of the pledges receives more media attention than others, which is likely to affect respondents' evaluations.

Research design

Our data concern citizens' evaluations of pledge fulfilment during the Conservative-Liberal Democrat coalition that governed from 2010 to 2015, which was the first coalition government to have held office in the UK since 1945. We expect the British public to give negative evaluations of pledge fulfilment in this context. Coalition governments are relatively rare in the UK and are generally viewed with scepticism.² Data from the British Election Study show that a majority of respondents thought that coalition government "makes it harder to decide who to blame", is a "less efficient form of government than single party government", and is less "in tune with public opinion than single party government". Two thirds agreed with the statement that "parties cannot deliver on their promises when they govern in coalition" (Fieldhouse, Green, Evans, Schmitt and Van der Eijk, 2015).

The data we examine come from the British Election Study, 2014-17 (ibid.), and refer to British citizens' evaluations of the fulfilment of six specific election pledges that were made in the campaign prior to the 2010 British General Election.

² Ipsos-Mori has asked a question about whether repondents think it "will be a good thing or a bad thing for the country if no party achieves an overall majority" since 1978, and only briefly in 1987 and again in 2010 did the number of those considering it a "bad thing" drop below 50 percent. Usually between 55 and 65 percent see it as a "bad thing", while only a third or less think it might be a "good thing". By 2014, the number thinking no party winning an overall majority in the next election would be a "good thing" had dropped to 26 percent, the lowest figure since 1983 (Ipsos-Mori, 2011).

Citizens' evaluations of the fulfilment of these pledges were measured in February-March 2014 in the first wave of the British Election Study's Internet Panel. We selected six election pledges from the 2010 election campaign to put to the public in 2014 to obtain their evaluations of fulfilment. Three of the pledges were made by the Conservatives and four by the Liberal Democrats (one was made by both parties). The selected pledges vary with respect to fulfilment: two were not fulfilled and four were at least partially fulfilled (two of which fully and two partially fulfilled). The six pledges were chosen to cover policy areas that respondents report as the most important: economy, health, education, immigration, domestic security and pensions.³

The first two pledges were unfulfilled. The first pledge is the Liberal Democrats' promise to "scrap tuition fees" in higher education. As part of the coalition agreement, the party dropped this manifesto promise and instead agreed with the Conservatives to treble tuition fees to £9,000 per annum, which resulted in widespread student demonstrations. Given the large amount of media attention given to this unfulfilled pledge, including a public apology by the Liberal Democrats for

³ Details of the primary sources consulted (the election platforms and relevant government actions) are reported in the Online Appendix. Our own assessment of the fulfillment of these pledges follows the procedure of an established research group that has assessed the reliability of coders' assessments of pledge fulfillment comparattively (see Thomson *et al.*, 2017). We conducted a robustness test in which we excluded the two pledges we code as partially fulfilled, since we acknowledge that there is more room for interpretation in these cases than the other four pledges. The test is reported in the Online Appendix and shows that our main findings hold wihtout these partially fulfilled pledges.

breaking it, we expect the overwhelming majority of people to identify this pledge as unfulfilled.

Second, we include the Conservatives unfulfilled pledge to "take steps to take net migration back to the levels of the 1990s – tens of thousands a year, not hundreds of thousands". There was no systematic decline in net migration since 2010, but rather an increase over the course of the subsequent governing period as a whole. Net migration first declined from over 200,000 in 2010 to 150,000 in the middle of 2012, and then rose again to over 300,000 in 2014. A considerable amount of media attention was given to this issue and the Conservative Party's promise on it.

The third pledge is the fully fulfilled promise made separately by both parties in their 2010 manifestos, to "scrap ID cards and the next generation of biometric passports". The introduction of ID cards was set out by the previous Labour administration in the Identity Cards Act 2006 but not carried out before 2010, and both Conservatives and Liberal Democrats pledged to scrap the plan entirely when in office. They included this pledge in the coalition agreement and the Identity Card Act 2006 was indeed repealed in 2010. Fourth, the Liberal Democrats pledged to "scrap compulsory retirement ages, allowing those who wish to continue in work to do so". This election pledge was fulfilled, as default retirement ages were phased out by the end of September 2011.

The fifth pledge is the Conservatives' pledge "to increase health spending in real terms every year", which we code as partially fulfilled. Given the severe cuts to other departments, we consider the ring-fencing of Health spending and real increases in some years as a partial fulfilment of the pledge. Nonetheless, there is some room for debate on this pledge. In 2012 the Chair of the UK Statistics Authority wrote to the Health Secretary stating that a real-terms increase had not been achieved in 2011-

12, and indeed that expenditure in 2011-12 had been lower than 2009-10, but that "given the small size of the changes and the uncertainties associated with them, it might also be fair to say that real terms expenditure had changed little over this period". Sixth and finally, we selected the Liberal Democrats' promise to raise the tax-free personal allowance to £10,000 for the start of the financial year 2011-2012. This is a partially fulfilled promise, because it took two years longer than stated to raise the tax-free allowance to the promised level. It was raised to £10,000 by the beginning of the new tax year in 2014.

The British Election Study includes a survey experiment embedded in a split-sample design to elicit respondents' evaluations of pledge fulfilment. Half of the respondents received questions that referred to "parties" and "promises", while the other half received questions that referred to "proposals". As explained above, the different question wordings are intended to raise and lower the salience of distrust as a heuristic when respondents give their evaluations. Since we only test for one specific frame, we avoid the trap of many survey experiments that lack a control group (see Gaines, Kuklinski and Quirk, 2006).

The "party-promise" questions in the survey experiment, which were put to just over 2,600 randomly selected respondents, read:

Before the 2010 General Election, the following promises were made by one or both of the parties that afterwards formed the government. For each of these, do you think the promise was fully kept, partially kept, or not kept at all?

- 1) A promise to increase health spending in real terms every year
- 2) A promise to scrap ID cards and the next generation of biometric passports
- 3) A promise to reduce net migration to less than 100,000 per year.
- 4) A promise to scrap compulsory retirement ages

- 5) A promise to raise the tax-free personal allowance to £10,000 for the start of the financial year 2011-2012
- 6) A promise to scrap University tuition fees

The "proposal" questions, which were also asked of just over 2,600 respondents, read:

The following policy proposals were made in 2010. For each of these, do you

think the proposal was fully, partially or not adopted at all?

- 1) A proposal to increase health spending in real terms every year
- 2) A proposal to scrap ID cards and the next generation of biometric passports
- 3) A proposal to reduce net migration to less than 100,000 per year.
- 4) A proposal to scrap compulsory retirement ages
- 5) A proposal to raise the tax-free personal allowance to £10,000 for the start of the financial year 2011-2012
- 6) A proposal to scrap University tuition fees

Each of the party promise-versions of the questions was answered by at least 1,892 respondents (excluding "don't know" answers). Each of the proposal-versions of questions was answered by at least 1,943 respondents. In total, we have 24,785 responses from 4,825 respondents with evaluations of pledge fulfilment. This reduces to 16,448 responses from 3,175 respondents if we exclude cases for which we miss information on one or more of our preferred measures of our explanatory or control variables. Most of our analyses are performed on this "stacked" dataset of 16,448 observations. This is the most appropriate design, given that we are interested in the effect of variation in actual pledge fulfilment on respondents' evaluations, which varies at the level of the respondent-item dyad. We apply sample weights in all of our models to ensure that the sample is representative of the population.

We measure respondents' political knowledge with a set of twelve knowledge questions. Respondents were first asked to identify their local MP, with randomized

false candidate names offered together with the correct one. In other questions, respondents were asked to match UK politicians and leaders of other countries to their correct political offices. Given that these knowledge questions were spread across waves 1-5, not all respondents received all twelve questions (although 90 percent did). We constructed a standardised 0-10 knowledge scale ranging from 0 (no correct answers) to 10 (all correct answers), regardless of the number of knowledge questions received. "Don't know" answers are counted the same as incorrect answers.

To measure trust defined in terms of distrust and heuristic thinking we use a variable to which the BES refers as "government trust". Recall that our conceptual definition of trust in relation to Lenard's (2008) use of the term distrust is that it is a general expectation on the part of a citizen that the party will meet his or her positive expectations when governing. The relevant BES item asks: "How much would you expect each of the following political parties to do a good job or a bad job if they are in government after the General Election (either by themselves or as part of a coalition)?" For each political party, respondents are asked to give a number from 1 (for "would do a bad job") to 7 (for "would do a good job"). For the pledges made by the Liberal Democrats, we match respondents' trust in the Liberal Democrats, and likewise for the Conservatives. For the pledge made by both parties we take the

⁴ Scholars of public opinion in the US commonly use the "do what is right" question in the US National Election Study to measure trust, by which respondents are asked how much of the time they thought they could trust the government in Washington to do what is right (e.g. Rudolph and Popp, 2009, p. 339). The BES does not contain a comparable instrument, but our measure is more in line with our concept of trust, which focuses on trust in governing parties, rather than government as a whole.

maximum values. We rescaled the values to range from 0-6, which facilitates the interpretation of the interaction terms.

The control variable for party identification uses respondents' answers to the question, "Do you usually think of yourself as close to any political party? [If so] Which party is that?" Responses to this question determined whether respondents were coded as identifying with a governing party or an opposition party. This operationalization is in line with Miller's (1991) view of party identification as a categorical rather than a continuous variable. It is also in line with previous operationalizations of party identification in the UK (Marsh and Tilley, 2010).

The demographic control variables consist of age, gender, education and occupation. For education, we distinguish between those with and without college education. The variable for occupation is based on the categories defined in the British Election Study, and distinguishes between higher (including professional and supervisory roles), intermediate and lower (including routine jobs) occupations.

In addition to our preferred model, which we present in the following section, we ran a series of robustness tests, which we also discuss at the end of the Analysis section and present in the Online Appendix. These robustness tests confirm the main findings from the model. They include a test using an alternative operationalisation of trust that approximates Hardin's (2002) conception of trust as encapsulated interest. They also include models that examine the answers to each pledge question separately, which are relevant given the differences between the pledges in, among other things, the amount of media attention they received. We run a model excluding the pledges we code as partially fulfilled, since there is arguably more ambiguity in these cases. We also present a model without the control variables. None of these robustness tests give substantially different results to those we now present.

Analysis

The distributions of respondents' answers to the twelve items about parties' election pledges suggests two noteworthy points (Figure 1). First, citizens' evaluations broadly reflect actual performance (Hypothesis 1a). This is true regardless of whether the pledges are presented as parties' promises or simply as proposals. We find that the overwhelming majority of respondents (80 percent) correctly identified the pledge on tuition fees as not fulfilled when they were asked if this promise was kept. This is unsurprising given the prominence of this broken promise in the media. Likewise, 61 percent of respondents correctly answered that the promise to reduce yearly immigration to under 100,000 was not kept. By contrast, people were more likely to say that pledges were at least partially fulfilled if they were in fact at least partially fulfilled. Majorities of respondents (54 and 60 percent) said that the promises to scrap compulsory retirement ages and ID cards were at least partially fulfilled. We coded these two pledges are fully fulfilled. For the pledge to raise the tax-free allowance, which we coded as partially fulfilled, the majority of respondents (54 percent) thought the pledge was either partially or fully fulfilled. Despite the obvious association between respondents' evaluations and actual performance, people give negative evaluations more readily than positive ones. For the pledge on health spending, which we coded as partially fulfilled, "not fulfilled" was the most common response category from the public, regardless of the question wording.

<Figure 1>

The second noteworthy point from Figure 1 is that people appear more inclined to say that pledges were fulfilled in response to questions that refer to the

adoption of proposals than the fulfilment of parties' pledges (Hypothesis 2a). This is the case for five of the six pledge topics. For instance, regarding the partially fulfilled pledge to increase health spending each year, 48 percent of people who were asked the party-promise version of the question said it was not fulfilled, compared to 38 percent of those who were asked the proposal version. The pledge to scrap ID cards and biometric passports is a partial exception to this pattern, since the evaluations in response to the party-promise question appear to be more positive than those in response to the proposal question. It is possible that the two components of the pledge – "ID cards" and "biometric passports" – had some distorting effect on the responses in this case. While a considerable amount of attention was given to the issue of ID cards in the media, less attention was given to biometric passports.

We now turn to the multiple regression analysis. The data are stacked, so that each row refers to a single respondent in relation to one of the twelve pledge items in the survey. Each respondent appears in the dataset a maximum of six times if they answered all six of the promise or proposal questions. This gives 16,448 observations after accounting for missing values on our variables of interest. The dependent variable in the main analysis presented here is coded 0 for respondents who answered that the promise was not fulfilled or the proposal not adopted and 1 for respondents who answered that the promise was partially or fully fulfilled or that the proposal was partially or fully adopted. We exclude "don't know" answers. We analyse this dichotomous dependent variable with a logit model, in which the standard errors are clustered by respondent. We prefer the dichotomous dependent variable since some of the answer categories are sparsely populated, particularly the "Fully fulfilled" category in response to pledges that were not fulfilled. These robustness tests (see

Online Appendix) include a multinomial model in which we use the three-category indicator of evaluations, which gives the same results.

The model in Table 1 indicates that people are more likely to say that a pledge was fulfilled it is was in fact fulfilled (Hypothesis 1a). This is true even for people who know very little about politics. However, the evaluations of more knowledgeable respondents are significantly more accurate – i.e. are more strongly affected by actual fulfilment - than are the evaluations of less knowledgeable citizens (Hypothesis 1b). The first three coefficients in Table 1 are relevant here: those associated with the variables Actual fulfilment, Knowledge, and the interaction between the two. The coefficient associated with Actual fulfilment is positive (1.86) and significant (p<.00), which indicates that Actual fulfilment has a positive effect on citizens' evaluations when Knowledge has a value of zero. So even respondents who gave incorrect answers to all of the knowledge questions are significantly more likely to give positive evaluations for pledges that were at least partially fulfilled. The exponent of the coefficient, 6.41, is the odds ratio and indicates that this is a large effect. The odds of people with a knowledge score of zero giving a positive evaluation are 6.41 times greater for pledges that were at least partially fulfilled than for pledges that were not fulfilled.

<Table 1>

The level of political knowledge held by respondents strengthens the impact of actual performance on their evaluations. The direct effect of Knowledge is negative and significant, which indicates that for pledges that were in fact unfulfilled, the more knowledgeable are less likely to say that those pledges were fulfilled. The exponent of the coefficient, .88, indicates that for every one-point increase in respondents' knowledge score (on the 0-10 scale), the odds that they give an "incorrect" positive

evaluation of an unfulfilled pledge decreases by 12 percent. The significant interaction between the variables Actual fulfilment and Knowledge indicates that the more knowledgeable are more likely to give positive evaluations of fulfilled pledges than are the less knowledgeable. Respondents' knowledge has a positive effect on their evaluations of fulfilled pledges; combining the main and interaction effects of the variable Knowledge gives a coefficient estimate of .07 (=-.12+.19), which is significant (s.e. .02, p=.00). The exponent of .07, which is 1.07, indicates that every one-point increase in respondents' knowledge score is associated with a 7 percent increase in the odds that they give a "correct" positive evaluation of a pledge that was in fact fulfilled at least partially.

Figure 2 depicts the effects of the variables Actual fulfilment and Knowledge as predicted probabilities, holding other variables of interest constant at their means (for scale variables) or modes (for categorical variables). The downward sloping lines at the bottom of the figure give the probabilities (and relevant confidence intervals) that respondents with different levels of knowledge offer positive evaluations of unfulfilled pledges. As the variable Knowledge increases from its minimum value of 0 to its maximum value of 10, the probability that a respondent gives an incorrect positive evaluation of an unfulfilled pledge decreases from .16 (95% confidence interval (CI): .08, .24) to .06 (95% CI: .04, .06). The ascending line indicates that as Knowledge increases from its minimum to its maximum, the probability that a respondent gives a correct positive evaluation of a fulfilled pledge increases from .55 (95% CI .46, .64) to .71 (95% CI .67, .75). The fact that these two sets of confidence intervals do not overlap, even for citizens with low levels of knowledge, reinforces the point that actual performance has a large effect on citizens' evaluations, even for those with low levels of political knowledge.

<Figure 2>

The next three coefficients refer to our experimental treatment variable and respondents' trust in politicians. People are significantly more likely to say that pledges were fulfilled in response to the proposal wordings of the questions (Hypothesis 2a). Moreover, trust has a marked positive effect on people's evaluations (Hypothesis 2b), particularly when they are presented with questions about parties' promises (Hypothesis 2c). The coefficient associated with the variable Proposal treatment refers to the effect of the proposal treatment compared to the party-promise treatment on respondents with the lowest level of trust in governing parties. The odds of these respondents giving a positive evaluation are 98 percent greater (p=.00) when they are presented with the proposal version of the question than the party-promise version of the question.

Respondents' trust, which is defined in terms of distrust and heuristic thinking, affects their evaluations positively, particularly when they are presented with questions that refer to parties' promises. The coefficient associated with the variable Trust is positive and significant. The coefficient is .24, of which the exponent is 1.27, and this indicates that every one-unit increase in respondents' trust on the 0-6 scale increases the odds that they give a positive evaluation by 27 percent. The significant negative coefficient of -.12 associated with the interaction between the variables Trust and Proposal treatment indicates that the effect of trust is significantly lower when respondents are presented with the proposal versions of the questions than the party-promise versions. However, the net effect of trust is still positive at .12 (=.24+(-.12)) and significant (s.e. .02; p=.00). So the effect of trust on people's evaluations, while still positive and significant, is much weaker when the proposal treatment is given compared to the party-promise treatment. The significant negative interaction effect

also indicates that the proposal treatment has a weaker effect on more trusting people. For respondents with relatively high levels of trust (or low levels of distrust), the proposal treatment has no effect.

The effects of the variables Proposal treatment and Trust are depicted as predicted probabilities in Figure 3. Like the predicted probabilities in Figure 2, these hold other variables constant at their means or modes. Since four of our six pledges were at least partially fulfilled, these probabilities refer to respondents' evaluations of pledges that were at least partially fulfilled. The unbroken line that ascends most steeply from left to right gives the predicted probabilities that respondents with different levels of trust offer positive evaluations in response to the party-promise questions. As trust increases from its minimum to its maximum values, the probability of a positive evaluation increases from .54 (95% CI .49, .59) to .83 (95% CI .80, .86). The flatter line refers to the effect of trust on the probability of a positive evaluation in response to the proposal versions of the questions. The effect is weaker, as the predicted probability increases from .70 (95% CI .66, .74) to .82 (95% CI .79, .86). The fact that the confidence intervals overlap at medium and high levels of trust indicates that the proposal treatment has no effect on respondents with medium and high levels of trust in politicians. For respondents with a score of 3 or higher on the variable Trust, the confidence intervals overlap. The bars, which show the distribution of observations, clearly show that many respondents have low levels of trust. In total 50 percent of the observations have values of 0 to 2 on the 0-6 scale, at which points there are significant differences between respondents who receive the party-promise and proposal treatments.

<Figure 3>

Of the control variables, Party identification is worth mentioning. Regardless of whether a pledge is actually fulfilled, people who identify with one of the governing parties are more likely to say that pledges were fulfilled than people who do not identify with a party. However, people who identify with an opposition party are not less likely to give positive evaluations than people who do not identify with a party. It could be that the non-identifiers tend to be more disengaged from politics than identifiers and that this disengagement is associated with more negative evaluations.

We conducted a series of tests to check the robustness of the results, which are detailed in the Online Appendix. Among these, we ran an additional analysis with an alternative measure of trust as encapsulated interest, according to which "I trust you because I think it is in your interest to take my interests in the relevant matter seriously ... you encapsulate my interests in your own interests" (Harding 2002, 1). Although based on a distinct definition and measure of trust, the results are very similar. We also ran a model that explored whether there is an interaction between trust and knowledge, which there is not. We ran a multinomial model with a threecategory dependent variable, which gave substantively the same results. Other robustness tests involved separate analyses of subsets of the data: a model that excluded the partially fulfilled pledges and models that examined each topic separately. We ran separate analyses of each topic with an additional variable to identify respondents who were personally affected by each pledge. We also ran a model without the control variables. The findings of these robustness tests are consistent with those presented here, although as would be expected not all coefficients are statistically significant when we limit the analyses to relatively small subsets of the data. Our key findings are therefore robust to different coding of

citizens' evaluations and actual fulfilment, are not driven by the particular topics of the pledges we examine, and are robust to different model specifications.

Conclusions

The evidence shows that British citizens are able to give accurate evaluations of policy performance in terms of governing parties' fulfilment of specific pledges that were made in previous election campaigns. The capacity to form evaluations based on performance is part of the healthy vigilance that Lenard (2008) identifies in mistrustful, critical citizens, who control their governments effectively. In line with our expectations, the evidence shows that knowledge improves the accuracy of citizens' evaluations. However, the level of knowledge required to make accurate evaluations of pledge fulfilment is well within most people's grasp, and does not require unrealistic demands of their cognitive capacities or interest in politics.

This key finding, together with a growing body of evidence from related research, points to the relevance of promissory representation in the practice of democratic politics. Promissory representation holds that candidates make promises to voters during election campaigns and seek to keep those promises if they subsequently enter office (Mansbridge, 2003, p. 515). Existing research focuses mainly on the extent to which parties they keep their election pledges, and generally finds high rates of pledge fulfilment depending on institutional constraints and economic conditions (Pomper and Lederman, 1980; Rallings, 1987; Royed, 1996; Thomson *et al.*, 2017). Little attention has been given to citizens' evaluations of promise keeping and breaking by governing parties, despite the fact that a responsible electorate, which rewards and punishes parties for variation in performance is an

essential counterpart to responsible parties (VO Key, 1966). Our key finding that citizens' evaluations of promise keeping reflect actual performance in terms of promise keeping contrasts with their answers to general questions about the extent to which politicians keep their promises, which are generally overwhelmingly negative (e.g. ISSP, 2008). Our finding is, however, in line with the two previous studies that examined this topic, which focused on citizens' evaluations of pledge fulfilment in Ireland (Thomson 2011) and Sweden (Naurin and Oscarsson, 2017, p.9). This research indicates that people hold far more nuanced and accurate views on promise keeping and breaking than suggested by their responses to general questions. The discrepancy between citizens' responses to general questions about promise keeping and the results of pledge research by academics may be due to different definitions of pledges. Citizens' may adhere at least implicitly to broader definitions of campaign promises than those used in research on pledge fulfilment (see also Naurin 2011). This is part of the explanation of the puzzle of the discrepancy between findings from research on pledge fulfilment and citizens' responses to general questions about promise breaking by politicians.

Another part of the explanation of the puzzle is that citizens' evaluations are influenced not only by actual policy performance, in terms of whether promises were actually kept, but also individual characteristics, some of which lead them to make negative evaluations of promise keeping, even for promises that were kept. This conclusion resonates with research on citizens' evaluations of governments' general performance on the economy and other broad policy areas (Lewis-Beck, Nadeau and Elias, 2008; Duch, Palmer and Anderson, 2000; Evans and Anderson, 2006; Marsh and Tilley, 2010). Like Kramer (1983) and Duch, Palmer and Anderson (2000), our explanation of evaluations of pledge fulfilment combines actual performance with

individual-level characteristics. Our study shows that this explanatory approach can be extended beyond evaluations of policy performance to evaluations of promise keeping.

Of these individual characteristics, we focused on trust as a heuristic, which

Lenard (2008) defines as distrust or cynicism and the expectation of betrayal. When

defined in these terms, trust is a heuristic for making evaluations in the context of
incomplete information and ambiguity (Hetherington, 2004; Rudolph and Evans,
2005). We argued that the evaluations of untrusting citizens are more negative
because they are informed by the stereotype of promise-breaking politicians. Our
survey experiment found evidence for this mechanism by changing the salience of
trust as relevant heuristic for respondents when they gave their assessments.

Respondents with low to average levels of trust in parties, and who were primed to
think about parties' promises when they answered, were more likely to say that
pledges were unfulfilled. Their answers were significantly more negative than similar
respondents who were not primed to think about parties' promises when they
answered. This significant difference indicates that respondents' evaluations of pledge
fulfilment are affected by their pre-existing levels of trust or distrust, independently of
whether or not the pledges were actually fulfilled.

The effect of distrust, which appears to shape citizens' evaluations of performance irrespective of actual performance, is of great concern. Many scholars see trust as a necessary component of the political culture that supports healthy democratic systems and lament declining levels of trust in recent decades (e.g. Diamond, 1998, p. 208). An alternative view is that "in perhaps a strange and counterintuitive way, representative democracy and distrust go together in political theory" (Hardin, 2002, p. 107; see also Hart, 1978). While we recognise the

importance of mistrust as healthy vigilance with respect to representative democracy, we see no benefit of distrust as endemic cynicism. We found it useful to distinguish between the normatively positive and negative aspects of trust and recommend this approach to other researchers. We also look forward to future research on the relations between these different aspects of trust as well as their causes and consequences. For instance, while mistrustful citizens are able to detect broken and kept promises, research is yet to identify the conditions under which they punish and reward governing parties for their performance in this respect.

References

Althaus, Scott L. 2003. *Collective preferences in democratic politics: Opinion surveys and the will of the people*. New York: Cambridge University Press.

Anderson, Christopher J. 2000. Economic voting and political context: a comparative perspective. *Electoral Studies* 19: 151-70.

Bartels, Brandon L., Janet M. Box-Steffensmeier, Corwin D. Smidt and Renée M.
Smith. 2011. The dynamic properties of individual-level party identification in the
United States. *Electoral Studies* 30: 210-22.

Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Dalton, Russell J. 1988. Citizen Politics in Western Democracies: Public Opinion and Political Parties in the United States, Great Britain, West Germany and France (Chatham: Chatham House, 1988).

Delli Carpini, Michael and Scott Keeter. 1996. What Americans know about politics and why it matters. New Haven, CT: Yale University Press.

- Diamond, Larry. 1998. *Developing Democracy: Toward Consolidation*. Baltimore: John Hopkins University Press.
- Downs, Anthony. 1957. An Economic Theory of Democracy (New York: Harper.
- Duch, Raymond M. Harvey D. Palmer and Christopher J. Anderson. 2000.
 Heterogeneity in perceptions of national economic conditions. *American Journal of Political Science* 44: 635-52.
- Duch, Raymond M. and Randolph T. Stevenson. 2008. *The Economic Vote. How Political and Economic Institutions condition Election Results*. New York:

 Cambridge University Press
- Erikson, Robert S., Michael B. MacKuen and James A. Stimson. 2001. *The Macro-Polity*. New York: Cambridge University Press.
- Evans, Geoffrey and Robert Andersen. 2006. The political conditioning of economic perceptions. *Journal of Politics* 68: 194-207.
- Fieldhouse, Ed, Jane Green., Geoffrey Evans., Hermann Schmitt, and Cees van der

 Eijk. 2015. *British Election Study Internet Panel Waves 1-6*http://www.britishelectionstudy.com/data-objects/panel-study-data/
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven CT: Yale University Press.
- Funk, Carolyn L. and Patricia Garcia-Monet. 1997. The Relationship Between Personal and National Concerns in Public Perceptions About the Economy. Political Research Quarterly 50: 317-42.
- Gaines, Brian J., James H. Kuklinski and Paul J. Quirk. 2006. The Logic of the Survey Experiment Reexamined. *Political Analysis* 15(1): 1-20.

- Grossback, Lawrence J., David A.M. Peterson and James A. Stimson. 2005.Comparing Competing Theories on the Causes of Mandate Perceptions. *American Journal of Political Science* 49: 406-419.
- Hardin, Russell 2002. *Trust and Trustworthiness*. New York: Russell Sage Foundation.
- Hart, Vivien. 1978. *Distrust and Democracy: Political Distrust in Britain and America*. Cambridge: Cambridge University Press.
- Hetherington, Marc, J. 2004. Why Trust Matters: Declining Political Trust and the Demise of American Liberalism. Princeton: Princeton University Press.
- Inglehart, Ronald. 1977. *The Silent Revolution: Changing Values and Political Styles Among Western Publics*. Princeton: Princeton University Press.
- Ipsos Mori. 2011. Attitudes to a Hung Parliament. Available at www.ipsos-mori.com/researchpublications/researcharchive/2630/Attitudes-to-a-Hung-Parliament.aspx (Accessed 16.05.2016)
- ISSP Research Group 2008. International Social Survey Programme: Role of Government IV - ISSP 2006. GESIS Data Archive, Cologne. ZA4700 Data file Version 1.0.0.
- Key, V.O. 1966. The Responsible Electorate. New York: Vintage.
- Klingemann, Hans-Dieter, Richard I. Hofferbert and Ian Budge. 1994. *Parties, Policies and Democracy*. Boulder: Westview Press.
- Kramer, Gerald H. 1983. The Ecological Fallacy Revisited: Aggregate vs. Individual-Level Findings on Economics and Elections and Sociotropic Voting. *American Political Science Review* 65: 131-43.
- Lenard, Patti Tamara. 2008. Trust Your Compatriots, but Count Your Change: The Roles of Trust, Mistrust and Distrust in Democracy. *Political Studies* 56: 312-32.

- Lewis-Beck, Michael S., Richard Nadeau and Angelo Elias. 2008 Economics, Party and the Vote: Causality Issues and Panel Data. *American Journal of Political Science* 52: 84-95.
- Lodge, Milton, Marco R. Steenbergen, and Shawn Brau. 1995. The Responsible Voter: Campaign Information and the Dynamics of Candidate Evaluation.

 *American Political Science Review 89: 309-326.
- Maloy, J.S. 2009. Two Concepts of Trust. Journal of Politics 71: 492-505.
- Mansbridge, Jane. 2003. Rethinking Representation. *American Political Science Review* 97: 515-28.
- McDonald, Michael D. and Ian Budge. 2005. *Elections, Parties, Democracy:*Conferring the Median Mandate. Oxford: Oxford University Press.
- Marsh, Michael and James Tilley. 2010. The attribution of credit and blame to governments and its impact on vote choice. *British Journal of Political Science* 40: 115-34.
- Miller, Arthur H. 1974. Political Issues and Trust in Government, 1964-1970. *American Political Science Review* 68(3): 951-72.
- Miller, Warren. 1991. Party Identification, Realignment, and Party Voting: Back to the Basics. *American Political Science Review* 85: 557-68.
- Milner, Henry. 2001. *Civic literacy: How informed citizens make democracy work.*Hanover, NH: University Press of New England.
- Naurin, Elin. 2011. *Election Promises, Party Behaviour and Voter Perceptions*. New York: Palgrave Macmillan.
- Naurin, Elin and Henrik Ekengren Oscarsson. When and Why are Voters Correct in their Evaluations of Specific Government Performance? *Political Studies*. Online First, July 6, 2017.

- Page, Benjamin I. and Robert Y. Shapiro. 1992. *The Rational Public*. Chicago: University of Chicago Press.
- Pomper Gerald M. and Susan S. Lederman. 1980. *Elections in America: Control and Influence in Democratic Politics*, 2nd ed. New York: Longman.
- Rallings, Colin. 1987. The influence of election programs: Britain and Canada 19451979. In Ian Budge, David Robertson and Derek Hearl. eds. *Ideology, Strategy and Party Change: Spatial Analysis of Post-War Election Programs in Nineteen*Democracies. Cambridge: Cambridge University Press, Pp. 1-15.
- Royed, Terry J. 1996. Testing the Mandate Model in Britain and the United States:

 Evidence from the Reagan and Thatcher Eras. *British Journal of Political Science*26: 45-80.
- Rudolph, Thomas J. and Jillian Evans. 2005. Political Trust, Ideology, and Public Support for Government Spending. *American Journal of Political Science* 49(3): 660-71.
- Rudolph, Thomas J. and Elizabeth Popp. 2009. Bridging the Ideological Divide: Trust and Support for Social Security Privitization. *Political Behavior* 31(3): 331-51.
- Sniderman, Paul M. and Thomas Piazza. 1993. *The Scar of Race*. Cambridge, MA: Belknap Press.
- Stimson, James A., Michael B. MacKuen and Robert S. Erikson. 1995. Dynamic Representation. *American Political Science Review* 89: 543-65.
- Stokes, Susan. 2001. *Mandates and Democracy: Neoliberalism by Surprise in Latin America*. Cambridge: Cambridge University Press.
- Thomson, Robert. 2011. Citizens' evaluations of the fulfilment of election pledges: Evidence from Ireland. *Journal of Politics* 73: 187-201.

Thomson, Robert, Terry Royed, Elin Naurin, Joaquin Artés, Rory Costello, Laurenz Ennser-Jedenastik, Mark Ferguson, Petia Kostadinova, Catherine Moury, Francois Pétry, Katrin Praprotnic. 2017. The Fulfillment of Parties' Election Pledges: A Comparative Study on the Impact of Power Sharing. *American Journal of Political Science* 61(3): 527-42.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.



Figures and tables

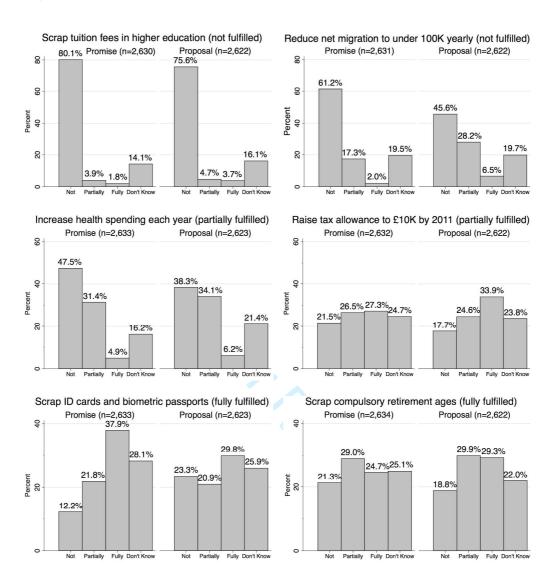


Figure 1. Citizens' evaluations of the fulfilment of six pledges

Source: British Election Study 2014-17. Note: "Promise" refers respondents' answers to question about the "fulfilment of parties' promises", while "Proposal" refers to respondents' answers to questions about the "adoption of proposals".

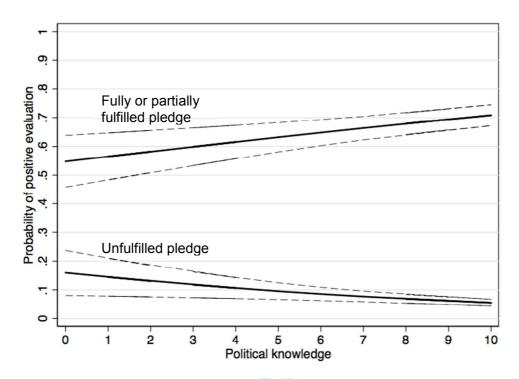


Figure 2. The effects of pledge fulfilment and knowledge on citizens' evaluations Note: Unbroken lines are point estimates of the probability that a respondent evaluates a pledge as partially or fully fulfilled; broken lines are 95% confidence intervals.

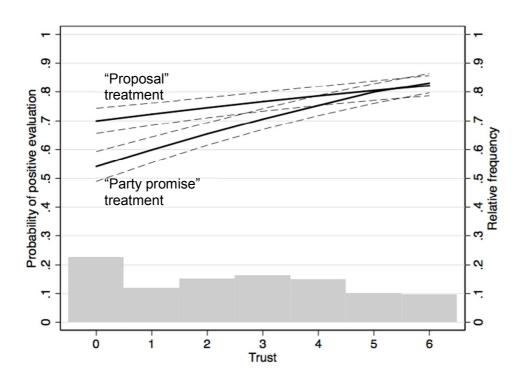


Figure 3. The effect of trust as a heuristic on citizens' evaluations

Note: Probability of a positive evaluation (i.e. evaluating a pledge as partially or fully fulfilled) of a pledge that was in fact at least partially fulfilled. Unbroken lines are point estimates; broken lines are 95% confidence intervals.

Table 1. Model of the causes of citizens' evaluations of promise keeping

Table 1. Model of the causes of citizens' evaluations of promise keeping			
	Exp(b)	b (s.e.)	p
Trust defined as mistrust and healthy vigilance:			
Actual performance and political knowledge			
Actual fulfilment (0=not; 1=partially/fully)	6.41	1.86 (.24)	.00
Knowledge	.89	12 (.03)	.00
Actual fulfilment × Knowledge	1.20	.19 (.03)	.00
Trust defined as distrust and heuristic thinking:			
Question wording and self-reported trust in			
governing parties			
Proposal treatment (0=party promise;			
1=proposal treatment)	1.98	.68 (.10)	.00
Trust	1.27	.24 (.02)	.00
Proposal treatment × Trust	.88	12 (.03)	.00
Control variables			
Party identification			
(reference category=no party identification)			
Governing party identifiers	1.60	.47 (.08)	.00
Opposition party identifiers	1.00	.00 (.08)	.96
Occupation	.99	01 (.05)	.76
Age	1.00	.00 (.00)	.77
Gender (0=male; 1=female)	.83	18 (.07)	.01
Education (0=lower than college; 1=college or	1.08	.07 (.06)	.24
higher)			
Topics (reference category=			
tuition fees and ID cards)			
Immigration numbers	4.67	1.54 (.10)	.00
Health spending	.30	-1.19 (.08)	.00
Tax-free allowance	1.23	.21 (.08)	.01
Compulsory retirement ages	1.21	.19 (.08)	.02
Constant	.09	-2.38 (.32)	.00
Log pseudolikelihood	-9996.70		
Wald χ^2 (p)	2164.95		.00
n observations	16,448		
n respondents (clusters)	3,175		
Note: Logit model Pobust standard errors cluster		dant Danandar	nt.

Note: Logit model. Robust standard errors clustered by respondent. Dependent variable: whether respondents evaluate the pledge as "not" fulfilled (0) or "partially" or "fulfy" fulfilled (1). Adjusted for sample weights.