Misperceived Social Norms and Political Accountability: Evidence and Theory*

Shuhei Kitamura† Ryo Takahashi‡ Katsunori Yamada§

August 25, 2025

Abstract

Electoral accountability represents a fundamental mechanism of democratic governance, yet whether voters effectively punish corruption remains theoretically and empirically contested. While canonical models predict that corruption revelations reduce incumbent support, strategic voting considerations—particularly beliefs about others' electoral behavior—can generate ambiguous accountability effects. We exploit a major corruption scandal involving Japan's ruling party during a national election to examine how social information shapes electoral responses to misconduct. Our pre-registered field experiment randomly provided voters with information about prevailing social norms regarding intolerance of the scandal. The intervention significantly increased overall turnout and challenger support, especially among swing voters, consistent with enhanced accountability. However, the same treatment increased incumbent support among ruling party loyalists. We demonstrate that these heterogeneous treatment effects are systematically explained by voters' prior beliefs about others' voting intentions. Voters who initially expected others to punish corruption responded by increasing their own sanctioning behavior, while those who did not expect others to sanction exhibited defensive behavior. These findings reconcile conflicting evidence on electoral accountability by highlighting how strategic considerations fundamentally alter democratic sanctioning. Our results suggest that information campaigns can either

*We are especially grateful to Ricardo Perez-Truglia, whose insightful and invaluable feedback on an earlier draft greatly improved this manuscript. We also thank Gwen-Jiro Clochard, Hülya Eraslan, Tetsuro Kobayashi, Takeshi Murooka, Pablo Querubin, Teppei Yamamoto, and Leeat Yariv for their helpful comments and suggestions. We are also grateful to seminar participants at Osaka and Waseda, and to conference participants at the 2025 Asia-Pacific Meeting of the ESA and the 2025 Frontiers in Behavioral and Experimental Economics. The experiment and pre-analysis plan (PAP) were pre-registered in the American Economic Association RCT Registry (RCT ID: AEARCTR-0014635). The experiment was approved by the Ethics Review Committee on Research with Human Subjects of Waseda University (ID: 2024-402). Takahashi acknowledges the financial support from Waseda University. The authors have no conflicts of interest to disclose. Corresponding author: Shuhei Kitamura (kitamura@iser.osaka-u.ac.jp).

†Associate Professor. Institute of Social and Economic Research, The University of Osaka, 6-1 Mihogaoka, Ibaraki, Osaka, 567-0047 Japan. Email: kitamura@iser.osaka-u.ac.jp.

‡Associate Professor. Graduate School of Economics, Waseda University. 1-6-1 Nishi-Waseda, Shinjuku, Tokyo, 169-0051 Japan. Email: ryo@waseda.jp.

§Professor. Faculty of Economics, Kindai University. 228-3, Shin-Kamikosaka, Higashi-Osaka city, Osaka, 577-0813 Japan. Email: kyamada@kindai.ac.jp.

strengthen or weaken accountability depending on the distribution of voter expectations, with important implications for anti-corruption interventions.
Keywords: Political accountability, corruption, misperception, social norm, strategic voting, moral, partisan identity, information campaign

1 Introduction

Democracy has played a central role in modern economic development (Acemoglu et al., 2019). Elections, a core institution of democracy, enable voters to hold politicians accountable by removing corrupt or underperforming incumbents (Barro, 1973; Ferejohn, 1986; Persson and Tabellini, 2000; Besley, 2006). When this mechanism works as intended, it disciplines leaders and enhances governance quality.

However, whether democratic institutions fulfill this function in practice remains unclear. While canonical models of electoral accountability predict that voters will punish corruption, real-world outcomes often defy these expectations. In Italy, Silvio Berlusconi faced repeated corruption charges and convictions yet continued to return to power (BBC News, 2019). In the 2024 U.S. presidential election, Donald Trump was convicted on 34 felony counts, yet his support remained largely unchanged (Associated Press, 2024). In India, 46 percent of newly elected Lok Sabha MPs in 2024 had declared criminal cases against them (India Today, 2024). These patterns raise important questions about why, despite opportunities for accountability, voters frequently return corrupt politicians to office.[1]

The literature offers potential explanations, including lack of credible information (e.g., Ferraz and Finan, 2011; Weitz-Shapiro and Winters, 2017), corruption-competence trade-offs (e.g., Winters and Weitz-Shapiro, 2013; Vera, 2019), group loyalty (e.g., Banerjee and Pande, 2009; Anduiza et al., 2013; Solaz et al., 2019), absence of credible clean alternatives (e.g., Pavão, 2018; Agerberg, 2020), clientelism and vote-buying (e.g., Bøttkjær and Justesen, 2021), and cynicism or tolerance of corrupt norms (e.g., Klašnja and Tucker, 2013; Vera, 2019). However, most of these accounts analyze voting decisions in isolation, raising the question of how strategic considerations about others' behavior may promote or undermine political accountability.

[1] The reelection of corrupt politicians is well-documented in countries such as Italy, Japan, and the United States (see, e.g., Rundquist et al., 1977; Reed, 1996; Chang et al., 2010).

Such strategic considerations are crucial for understanding political accountability. Party loyalists, for example, may reelect a corrupt incumbent rather than risk an opposition victory, or abstain if they expect fellow loyalists to secure the incumbent's win. Strategic voting—where individuals condition their choices on expected behaviors of others—is common across electoral settings (see, e.g., Cox, 1997; Kawai and Watanabe, 2013; Eggers and Vivyan, 2020). Yet, while strategic voting is well studied in other contexts, workhorse models of electoral accountability typically assume sincere voting and abstract from such considerations of others' behavior. This theoretical gap is significant because strategic incentives can either amplify or attenuate accountability pressures, making the net effect on democratic sanctioning theoretically ambiguous.

Strategic voting necessarily involves making inferences about others' behavior. Yet voters may misjudge these behaviors—over- or underestimating, for instance, general tolerance for corruption and thus the likelihood that others will punish a corrupt incumbent. Recently, a growing body of research shows that individuals often misperceive others across diverse contexts.[2] In voting contexts, for instance, Gerber et al. (2020) find that voters substantially overestimate the likelihood of close elections, while Carlson and Hill (2022) show that individuals hold biased expectations about how a hypothetical, randomly selected peer would vote.[3] Such misperceptions can distort expectations about collective behavior and, in turn, shape individual decisions about whether to vote or whom to support. Building on these insights, we examine how voters' perceptions of others' attitudes toward corruption interact with strategic considerations to affect electoral accountability.

We examine this research question through an online survey experiment conducted in Japan in the run-up to the October 2024 general election for the Lower House, during one of the most prominent political corruption scandals implicating the ruling party in recent decades. The study exploits variation in voters' perceived social norms: individuals in the treatment group received information about the aggregate level of public disapproval toward the scandal, while those in the control group did not. The treatment was designed to shift beliefs about prevailing norms while holding information about the scandal itself constant.

[2] Bursztyn and Yang (2022) provide a recent review.

[3] Beyond voting behavior, prior research has documented misperceptions about polarization (Levendusky and Malhotra, 2016), policy disagreement (Chambers et al., 2005), public opinion (Mildenberger and Tingley, 2019), ideological extremity (Ahler, 2014), and partisan stereotyping (Ahler and Sood, 2018).

We analyze how this information updates beliefs and influences voting behavior.

We find that our information treatment increased turnout and support for challengers, particularly among swing voters. Among ruling party loyalists, however, the same information had the opposite effect: it increased support for incumbents. We also find that these heterogeneous reactions stem from differences in initial beliefs about others' voting behavior.

The estimated treatment effect on turnout is sizable—around six percentage points. The figure is comparable to the benchmark effects from large-scale voter mobilization campaigns in the United States (e.g., Gerber and Green, 2000; Gerber et al., 2008). This effect is especially notable given the non-personalized and non-coercive nature of our intervention. We discuss potential explanations for this result.

Interestingly, heterogeneity analysis reveals that the effects on turnout and challenger support are observed in districts where the incumbent is not implicated in the scandal, suggesting a spillover effect of the treatment on incumbents who are themselves clean but from the same party as implicated candidates. In addition, we find strong punishment, especially by swing voters, in districts with implicated incumbents.

To interpret these patterns, we propose a simple model of electoral accountability with strategic considerations about others and empirically assess its predictions. First, we find that the increase in turnout is not driven by an average increase in individual perceptions of pivotality. Rather, we find that the increase in the *precision* of such perceptions seems to explain it.

These findings have implications for studies examining the relationship between pivotality and turnout. Lab experiments (e.g., Levine and Palfrey, 2007; Duffy and Tavits, 2008; Agranov et al., 2018) find that expected closeness increases participation. Bursztyn et al. (2024) show that tighter polls boost turnout in Swiss referenda. However, field experiments yield more mixed results: Enos and Fowler (2014) and Gerber et al. (2020) find limited evidence that emphasizing closeness increases turnout. Our information treatment increased turnout not by raising average perceived pivotality, but by reducing the dispersion of pivotality beliefs. This precision-based channel offers a new perspective on how information influences voter turnout.

Beyond the mechanisms related to perceived pivotality, we consider two key mechanisms. First, we introduce *moral reinforcement*, which connects to the literature on the role of

morals in voting (e.g., Enke, 2020; Walter and Redlawsk, 2019). A key novelty of our model is that we explicitly incorporate the *moral costs* of supporting corrupt politicians. Second, we introduce *identity reinforcement*, which arises when norm signals are perceived as a threat to a voter's partisan identity, prompting loyalists to defensively support the incumbent. While previous studies show that partisan loyalty leads citizens to judge corruption more leniently when it affects their own party (e.g., Anduiza et al., 2013; Solaz et al., 2019), our analysis shows that corruption information can trigger defensive behavior when loyalists expect others to punish the corrupt politician, thereby counteracting electoral accountability. Crucially, both the defensive response among loyalists and the mobilization response among swing voters stem from our information treatment conveying *others'* perceptions rather than information about corruption *per se*. We find evidence consistent with both mechanisms.

A growing literature examines how to improve political accountability. Prior research has demonstrated that increasing transparency through audits can reduce corruption and lower the re-election rates of corrupt incumbents (Ferraz and Finan, 2008; Bobonis et al., 2016; Avis et al., 2018). A related line of work studies how information about politicians' past performance shapes voter behavior through field and survey experiments, yielding mixed results (e.g., Banerjee et al., 2011; Chong et al., 2015; Weitz-Shapiro and Winters, 2017; Buntaine et al., 2018; Dunning et al., 2019; Adida et al., 2020; Cruz et al., 2020; Arias et al., 2022; de Figueiredo et al., 2023). For instance, Banerjee et al. (2011) find that providing information about candidate performance and qualifications increased turnout. By contrast, Chong et al. (2015) find reductions in both turnout and challenger support due to their information treatment. Dunning et al. (2019) report null average effects across seven randomized experiments in six countries. These mixed findings highlight the need for further evidence on how voters process political information—a gap this study addresses.

Despite these advances, relatively little research examines how voters' beliefs about others shape political behavior. For example, Arias et al. (2022) show that prior beliefs about political malfeasance moderate voters' responses to new information, but focus exclusively on individuals' own beliefs about politicians. We extend this literature by demonstrating that voters' beliefs about prevailing intolerance norms can influence turnout and vote choice. Unlike Chong et al. (2015), we find that providing intolerance norms information did not crowd out turnout or challenger support. Rather, our treatment increased them, especially

among swing voters. These effects are sizable, highlighting the potential effectiveness of information treatments that heighten the moral costs of supporting corrupt politicians.

We also contribute to a growing literature on how perceptions of others shape political behavior. Prior work shows that beliefs about others affect political contributions (Perez-Truglia and Cruces, 2017), protest (Cantoni et al., 2019), activism (Hager et al., 2023), and climate action (Andre et al., 2024). We extend this literature by demonstrating that perceptions of others also influence electoral behavior.

2 Background

Since its founding in 1955, the Liberal Democratic Party (LDP) has dominated Japanese politics, losing power only briefly in 1993–1994 and 2009–2012. Between 2023 and 2024, the LDP was at the center of a major political scandal. The scandal involved allegations that three major party factions—the Abe, Nikai, and Kishida factions—failed to properly disclose over 970 million yen (about 6.5 million USD) in political funds, violating mandated finance disclosure regulations and raising suspicions that the funds were used as personal slush funds. The scandal is widely regarded as the most serious case of political corruption in modern Japanese politics (e.g. Yomiuri Shimbun, 2024; NPR, 2023).

Public backlash was swift. By February 2024, the LDP's approval rating had dropped to 21%—its lowest since returning to power in 2012 (Asahi Shimbun, 2024). In the October 2024 general election, the LDP and its coalition partner Komeito lost their Lower House majority for the first time since 2009. The LDP's seats fell from 237 to 191, and 28 of the 46 politicians implicated in the scandal lost their races. The result dealt a significant blow to Prime Minister Shigeru Ishiba, who had taken office just weeks earlier on October 1, 2024.

3 Experimental Design

3.1 Overview

Our pre-registered experiment (AEARCTR-0014635) took place between October 18th and 28th, 2024, while the election occurred on October 27th. Our study consisted of three

surveys: a baseline survey, a post-treatment survey, and an endline survey. Figure 1 presents an overview of the survey timeline. The baseline survey was conducted between the 18th and 23rd, the information intervention and post-treatment survey took place between the 25th and 26th, and the endline survey was administered between the 28th and 31st. The election occurred on the 27th (Sunday).

The primary objective of this study was to investigate how voters' perceptions of others' attitudes influence their own voting behavior. To this end, we implemented a randomized information intervention between the baseline and post-treatment surveys. The core feature of this intervention was the provision of information about the prevailing social norm—specifically, participants were shown aggregate attitudes toward corruption, as measured in the baseline survey.

From a political-economy perspective, voters act strategically under uncertainty. They form beliefs about how intolerant their fellow citizens are of corruption and adjust their turnout and ballot choices accordingly. Our information treatment targets precisely those beliefs: by revealing the aggregate norm, it changes voters' perceptions of others' attitudes. That updated view, in turn, produces the observed shifts in turnout and vote choice. Examining this process is essential for understanding how voters' considerations of others shape electoral outcomes.

Figure 2 reproduces the intervention screen shown to participants. Combining text with a simple graphic, the screen reported the baseline social norm: on average, respondents believed that 67% of other voters regarded the scandal as unacceptable.

Further details of the overall survey design are provided below.

3.2 Survey procedure and questionnaires

Participants were informed at the outset that the surveys were related to the upcoming election, conducted for academic research by university researchers, and that no moral judgment would be made based on their responses. To avoid the experimenter demand effect, the general but specific purpose of the surveys was disclosed. Participant anonymity was guaranteed, and they were free to withdraw from the experiment at any time.

To facilitate completion, we limited the number of survey questions as much as possible

and kept the questions clear and straightforward to avoid complications or confusion.[4] Furthermore, in addition to a baseline participation reward, payments increased toward the end of the experiment. By completing all surveys, participants could earn up to 65 points (one point is equivalent to one JPY).

We also incentivized certain questions to encourage truth-telling and minimize the experimenter demand effect, following the approach of Krupka and Weber (2013). Questions were incentivized by informing participants in advance that questions with a special icon were bonus questions, and 100 participants who answered them correctly would receive an additional 100 points.[5]

Screening & Baseline survey (October 18th). Before conducting the baseline survey, screening questions were administered. The screening questions included postal code, whether the respondent had already voted in early voting, and an attention-check question. Those who did not live in the target constituencies, had already voted, or did not pass the attention-check question did not continue to the baseline survey. For those who proceeded to the baseline survey, we asked about their demographics and questions related to politics and policy. The baseline survey contained 11 questions in total.

- *Measuring prior beliefs for the intervention screen.* In the baseline survey, one question asked about respondents' perceptions regarding the proportion of other voters who consider the slush fund issue unacceptable:

> This survey project brings together a diverse group of participants, whose responses can be said to represent the views and attitudes of the average voter in Japan.
>
> Regarding the following question, how many out of 100 participants do you think answered "Completely unacceptable" or "Somewhat unacceptable"?

[4] The baseline survey, which includes the most questions, contains only 11 questions and is designed to be completed within 5 minutes.

[5] The decision to limit incentives to 100 participants was made solely for practical reasons. Specifically, when contracting with the survey firm, it was necessary to determine the total payment amount, including participant compensation, in advance. By capping the number of incentivized respondents at 100, we were able to fix the overall cost. Although participants were not informed of this operational constraint, the incentive structure remained intact in that each respondent's compensation was still contingent on their truthful reporting.

<Question> To what extent can you tolerate issues such as omissions in political fund balance reports or kickbacks from political funds?

<Answer> 1. Completely unacceptable 2. Somewhat unacceptable 3. Somewhat acceptable 4. Completely acceptable

To the above question, I believe that ( ) out of 100 people answered "1. Completely unacceptable" or "2. Somewhat unacceptable."

Respondents entered a number from 0 to 100, and these values were used to create the intervention screen.

We intentionally designed the survey question to incentivize respondents to provide their best estimate of social norms, regardless of their personal attitudes toward corruption—that is, respondents were encouraged to make their best guess even if they personally viewed the slush fund issue as acceptable or unacceptable. We then averaged these perceptions to define the prevailing social norm (67%) used in the subsequent analyses.

Intervention & Post-treatment survey (October 25th). We first showed the treatment group our intervention screen explained above. In contrast, the control group proceeded without receiving any additional information, as it is called passive design by Haaland et al. (2023). The post-treatment survey was administered immediately afterward and asked about the intended voting behavior and posterior beliefs. There were a total of eight questions. Both the treatment and control groups were asked the same set of questions.

Endline survey (October 28th). One day after the election, the endline survey asked both groups two questions about actual voting behavior: whether respondents voted, and if so, for which political party candidate.

Table A.2 in the Online Appendix shows the variables used in the analysis based on surveys administered after the treatment. All survey questions (originally in Japanese) and their English translations are also provided in the Online Appendix.

3.3 Sample

We conducted the experiment in collaboration with *Myvoice Inc.*. The target population consisted of 2,800 registered individuals aged between their 20s and 60s in target constituencies in Japan.

These constituencies fall into two types, both of which had incumbent candidates from the LDP: *fraud districts*, where the incumbent politicians were involved in the slush fund issue, and *clean districts*, where the incumbent politicians were not involved.[6]

We determined this target sample size through a power calculation described in the PAP. The final sample includes 10 fraud districts and 15 clean districts. The details of the selection process for the target constituencies are provided in Section A.1 of the Online Appendix. As shown in Figure A.1 in the Online Appendix, these constituencies are distributed across various regions. We use the combined sample from both types of constituencies in the main analysis. Heterogeneous effects between these constituencies are examined in Section 7.

Due to limitations in the demographic characteristics of the survey company's registered sample pool, we conducted stratified sampling based on constituency types. Table A.1 in the Online Appendix compares these constituencies with others, demonstrating that they are comparable to each other as well as to non-targeted constituencies where incumbent candidates were implicated in the slush fund issue or to non-targeted constituencies in general. It is also worth noting that the vote shares for the LDP and the Constitutional Democratic Party (DP, the largest opposition party) are each approximately one quarter, with the LDP holding a slightly larger share.

3.4 Balance and attrition

We conducted a statistical test (t-test) to check the balance between the treatment and control groups. The variables used to check the balance were pre-specified in the PAP. Table A.3 in the Online Appendix shows that all variables, except for educational attainment, are balanced at a false discovery rate of 0.05. Following the PAP, we included this variable as a baseline control in the subsequent regression analyses.

[6] Japan is divided into 289 single-member constituencies, each with roughly 364,000 registered voters on average (as of October 1st, 2024). A constituency is smaller than a prefecture (47 in total) but larger than a municipality (1,741).

We also checked the balance between attritors and non-attritors using the same set of variables. Table A.4 in the Online Appendix shows the result. First, the number of attritors after the treatment is impressively low (95/2280 or 4%). Second, the table shows that all variables, except gender and marital status, are balanced at a false discovery rate of 0.05. As pre-specified in the PAP, we computed bounds using the method proposed by Lee (2009).

The summary statistics of the final sample are provided in Table A.5 in the Online Appendix.

4 Empirical Strategy

In the main regression analyses, we estimate the following model using Ordinary Least Squares (OLS):

$$y_i = \alpha + \beta T_i + \mathbf{X}_i \gamma + \varepsilon_i, \tag{1}$$

for individual i, where y_i is the outcome, T_i is the treatment indicator, $\mathbf{X}_i$ is a vector of control variables, and ε_i is the error term. The control variables include the baseline control variable and prefecture fixed effects. We use a randomized experiment to estimate the causal effect β. As specified in the PAP, we use robust standard errors in the main analyses and conduct robustness checks using the Wild bootstrap and randomization inference.

5 Empirical Results

5.1 Perceptions of intolerance norms

Panel (a) in Figure 3 shows the actual distribution of respondents' prior beliefs about other voters' corruption intolerance. The average value—our definition of intolerance norms—is 67%. The distribution is wide and markedly skewed, with many respondents perceiving other voters as less tolerant of corruption.

Panel (b) displays the cumulative distribution functions (CDFs) for the three groups. The CDFs for swing voters and opposition-party supporters first-order stochastically dominate that of LDP supporters, indicating that, at every level of prior belief, a greater share of swing voters and opposition supporters believe that a larger proportion of others are intolerant of

corruption, relative to LDP supporters.

Overall, although some LDP supporters hold stringent views about the scandal, such views are more commonly found among swing voters and opposition-party supporters, whereas LDP supporters tend to hold more lenient views. We later exploit this important variation in prior beliefs across political alignment to examine heterogeneous treatment effects.

5.2 Effects of the information treatment

5.2.1 Effects on voting outcomes

Figure 4 summarizes the effects of the information treatment on voting outcomes. Our main analyses focus on voting outcomes reported in the post-election survey. Table A.6 in the Online Appendix reports regression results based on these outcomes as well as on the pre-election intention measures.

First, the treatment increased voter turnout by 6.4 percentage points (left panel). This effect appears to be driven by a 7.3-percentage-point increase in the likelihood of voting for a challenger (right panel). By contrast, we find no evidence that the treatment affected the likelihood of voting for the incumbent candidate (middle panel). Taken together, these results suggest that providing information about actual intolerance norms mobilized voters and increased support toward the challenger.

These findings connect to a broader literature on voter information campaigns. Banerjee et al. (2011) show that providing information on candidate performance and qualifications increased turnout in India. Arias et al. (2022) document a non-monotonic effect of malfeasance revelations on turnout in Mexico, with turnout increasing at both low and high levels of malfeasance. By contrast, Chong et al. (2015) find that exposing corruption reduced both turnout and challenger support. Unlike Chong et al. (2015), our results indicate that information about intolerance norms not only mobilized voters but also increased challenger support.

Importantly, voter responses varied by political alignment. In Table A.7 in the Online Appendix, we find that swing voters were mobilized and became more likely to vote for a challenger. In contrast, while the treatment did not affect turnout among LDP supporters, it significantly increased their support for the incumbent. Finally, we find no significant effects

among supporters of opposition parties. The results for swing voters and LDP supporters are summarized in Figure 5.

One possible interpretation is as follows. First, recall that swing voters tend to hold stringent views on corruption, whereas LDP supporters tend to hold lenient views, as shown in Figure 3. Upon learning that others are *more* tolerant of corruption, swing voters may have felt a stronger motivation to support the challenger, leading to increased turnout and a higher likelihood of casting a challenger ballot. In contrast, LDP supporters appear less responsive to the information treatment. Nonetheless, when informed that others are *less* tolerant of corruption, they became more likely to vote for the incumbent.

To assess the plausibility of this interpretation, we further divide voters by their prior beliefs—specifically, whether they hold lenient or stringent views. As shown in Table A.8 in the Online Appendix, the results are consistent with this interpretation. We find that it was primarily stringent swing voters who were mobilized and chose to vote for a challenger. In contrast, lenient LDP supporters were more likely to support the incumbent and were also significantly less likely to vote for a challenger. This pattern effectively offsets any net effect on turnout among LDP supporters (see Table A.8 in the Online Appendix).

Table A.9 in the Online Appendix summarizes the results discussed above. In summary, swing voters tend to hold stringent prior beliefs—expecting that others are intolerant of the corruption scandal—whereas LDP supporters tend to hold more lenient views. Upon learning that others are less tolerant (for swing voters) or more tolerant (for LDP supporters), they become more likely to vote for a challenger or the incumbent, respectively. However, the overall effects are stronger among swing voters.

5.2.2 Robustness checks

We conduct various robustness checks in the Online Appendix, as specified in the PAP. Table A.15 reports p-values using randomization inference, considering multiple hypothesis testing (Young, 2019), and applying the Wild bootstrap (e.g., Cameron et al., 2008). Table A.16 presents Lee bounds (Lee, 2009) that account for attrition. Table A.17 shows estimation results that include additional controls (gender and marital status) which were statistically significant in the attrition check. Finally, Table A.18 presents estimates using

Poisson pseudo-maximum likelihood (Correia et al., 2020). Overall, the qualitative results are similar across these exercises.

5.2.3 Effects across district types

Recall that our sample includes both constituencies where the incumbent was implicated in corruption and those where no wrongdoing was detected. A central question is whether the information treatment increases turnout and challenger support across all districts, or only within those directly tainted by corruption. Put differently, we ask whether voters punish incumbents only when they themselves are implicated in corruption, or whether they also punish candidates from the ruling LDP more generally, regardless of direct involvement.

According to Panel A in Table 1, the treatment significantly increases voter turnout in *clean districts* (Column (4)), an effect driven largely by swing voters (Column (5)). By contrast, the estimates for fraud districts are small and statistically insignificant (Column (1)). Decomposing the sample helps explain why. Recall that turnout requires voting for either the incumbent or a challenger. As shown in Column (1) of Panel B, treated voters—especially swing voters—*do* punish tainted incumbents. This incumbent penalty offsets the positive increase in challenger support documented in Panel C, yielding only muted effects on turnout (Panel A). In clean districts, by contrast, no such incumbent penalty arises, so the increase in challenger support translates directly into higher turnout. Interestingly, challenger support rises in both district types, suggesting a spillover effect of the treatment: even when the incumbent is not implicated, swing voters become more likely to back challengers, which produces a strong turnout effect in clean districts.

In addition, we observe a different pattern among LDP loyalists. Treated loyalists increase their support for the incumbent, especially in clean districts. They are also somewhat less likely to support challengers, though this effect is statistically insignificant.

6 Mechanism

To further understand the empirical patterns documented above, we outline a simple voting model that extends canonical frameworks of electoral accountability (Barro, 1973; Ferejohn, 1986; Persson and Tabellini, 2000; Besley, 2006). A complete formalization and all proofs

are relegated to the Online Appendix. In what follows, the terms "loyalists" and "LDP supporters" are used interchangeably.

6.1 Model setup

Setting. A unit mass of voters chooses whether to vote (incurring a fixed cost $c > 0$) and, if so, whether to support a scandal-tainted incumbent I or a (clean) challenger C.

Voter i's ex-post utility is

$$U_i(x, t_i) = u_i(x) \; - \; c t_i,$$

where $t_i \in \{0,1\}$ is an indicator for turning out and

$$u_i(x) = \begin{cases} -\theta_i + \eta_i \phi, & x = I, \\ 0, & x = C, \end{cases}$$

where $\theta_i > 0$ is the moral cost of re-electing the corrupt incumbent; $\eta_i \in \{0,1\}$ is a party-loyalty indicator ($\eta_i = 1$ for loyalists, 0 for swing voters); and $\phi > 0$ is the expressive benefit a loyalist enjoys when the incumbent remains in office. Normalizing $u_i(C)$ to zero is without loss of generality.

Thus, the net utility gain from electing the challenger rather than the incumbent is

$$\Delta_i \; := \; u_i(C) - u_i(I) \; = \; \theta_i - \eta_i \phi.$$

Therefore, each voter i balances two forces: *moral cost* $\theta_i > 0$ from re-electing a corrupt politician and *expressive benefit* $\eta_i \phi$ from keeping the loyalists' preferred party in office. The latter benefit is zero for swing voters.

Signal and voting behavior. Voters form prior beliefs $\hat{p}_i$ about the unknown challenger vote share P. Before the election, voters receive a public signal stating that a large share of the electorate finds the slush-fund scandal unacceptable (e.g. "67% consider the scandal unacceptable"). This signal shifts their beliefs about the challenger's vote share.

Let $p \in (0,1)$ be the voter's expected challenger vote share. Let $\rho(p)$ denote the probability that a single ballot is pivotal in a large electorate when the expected challenger share

is p. Under the standard Poisson (or asymptotic normal) approximation, $\rho(p)$ is continuous, symmetric around $p = 1/2$, and single-peaked with its maximum at $p = 1/2$.

Given the net utility $|\Delta_i|$, the *instrumental benefit* of turning out is

$$B_i(p) := \rho(p)\,|\Delta_i|,$$

and voter i turns out whenever $B_i(p) > c$.

Consequently, voting behavior is summarized by the following rules.

- **Turnout rule**

$$t_i = \mathbb{1}\{B_i(p) > c\}.$$

  The voter turns out if and only if the instrumental benefit evaluated at the posterior vote-share belief p strictly exceeds the fixed cost c; if $B_i(p) = c$ she abstains.

- **Ballot rule** (given turnout)

$$x_i = \begin{cases} C & \text{if } \Delta_i \geq 0, \\ I & \text{if } \Delta_i < 0. \end{cases}$$

  Conditional on turning out, the voter selects the option that yields higher utility; An indifferent voter ($\Delta_i = 0$) casts a challenger ballot C.

Timeline. The timeline of the model is as follows:

1. Nature draws the latent challenger vote share P. Each voter i independently draws a triple $(\theta_i, \eta_i, \hat{p}_i)$ from the common-knowledge joint distribution.
2. The public signal is disclosed.
3. The election takes place. Voters compare the instrumental benefit $B_i(p) = \rho(p)|\Delta_i|$ with the fixed cost c. Those who turn out vote C if $\Delta_i \geq 0$ and I otherwise. The incumbent stays in office if the vote share for I is at least one-half; otherwise, the challenger wins. Finally, payoffs are realized.

This game admits at least one pure-strategy Bayesian-Nash equilibrium; the formal statement and proof appear in the Online Appendix.

The next section presents the model's key comparative-statics results.

6.2 Comparative statistics

The public signal updates voters' expected payoffs through several mechanisms. We examine the four most plausible channels below: *mean update*, *variance shrink*, *moral reinforcement*, and *identity reinforcement*. Throughout, a superscript 0 denotes a quantity under the baseline (pre-signal) information set, whereas a superscript 1 denotes its counterpart after the signal is received.

Mean update. The first channel is the *mean-update* channel; that is, the signal moves mean beliefs from a prior p^0 to a posterior p^1 such that $|p^1 - 0.5| < |p^0 - 0.5|$. After receiving the signal, voters therefore believe that the election is more competitive than they initially thought. This update can occur from either above or below. For voters with a low prior—those who initially believe that the challenger's vote share is low—the update comes from below, raising their estimate of the challenger's support (and thus the perceived competitiveness of the race). As noted in Section 5, loyalists tend to hold lenient views on corruption, suggesting that they are more likely to form such low priors.

Conversely, for voters with a high prior—those who initially believe that the challenger's vote share is high—the update comes from above, lowering their estimate of the challenger's support. Because swing voters tend to hold stringent views on corruption (see Section 5), they are more likely to start with such high priors.

In both cases, the shift in beliefs increases perceptions of pivotality and, therefore, the incentive to turn out. Since we observe higher turnout among swing voters, a mean-update mechanism implies that the information treatment should also increase their perceived pivotality.

To test this implication, we use responses to the post-treatment survey question, "How much influence do you think your vote has on the election outcome?" (1 = Not at all, 2 = Not much, 3 = Some, 4 = A great deal) as the outcome variable. According to Table A.10 in the Online Appendix, this measure is strongly positively correlated with actual turnout.

However, the data do not perfectly fit this channel. As shown in Figure 6 (and Table A.11 in the Online Appendix), the estimated treatment effect on perceived pivotality is positive but not statistically significant for the full sample or for swing voters. By contrast,

we find a significant positive effect among loyalists—despite no corresponding increase in their turnout. Table A.12 in the Online Appendix indicates that this effect among loyalists is driven by respondents with lenient views. We return to this point later.

We also examine an alternative proxy for perceived pivotality: respondents' expectations about the share of voters in their constituency who will support the incumbent. Although this variable is a more direct measure of the challenger's expected vote share in the model—even though the question asks about the incumbent—its interpretation carries an important caveat. Because the item explicitly references the incumbent's electoral performance, responses are susceptible to *identity reinforcement* described below: loyalists may report higher incumbent support simply to affirm their partisan identity. In that case, any estimated treatment effect would blend belief updating with strengthened loyalty rather than capture pure changes in perceived pivotality.

Bearing this caveat in mind, Panel A of Table A.13 in the Online Appendix shows results mirroring our earlier pivotality measure: a strong treatment effect for loyalists. Yet the sign is positive: treated loyalists believe that more voters will back the incumbent than do their control-group counterparts—contrary to the model's mean-update prediction. A plausible interpretation is that the treatment heightens partisan identity, making loyalists even more inclined toward their party. Panel B of the same table supports this view. Using respondents' beliefs about the share of voters who *should* back the incumbent, we find a positive treatment effect among loyalists. This indicates that the intervention strengthens their normative support for the incumbent. Consequently, they form a wishful belief that more voters *will* vote for her.

In sum, the evidence provides little support for the mean-update channel.

Variance shrink. Thus far, we find that the mean-update channel is unlikely to be the primary mechanism. A natural next step is to consider whether the treatment instead reduced the perceived *variance* of the vote-share distribution while keeping its mean fixed—a mechanism we label the *variance-shrink* channel. In the Online Appendix, we extend the model to allow for non-degenerate beliefs with variance. The extension shows that, when the variance of beliefs about the challenger's vote share is reduced, while the perceived mean vote share is unchanged, the expected pivotality rises. The downstream implication mirrors

that of the mean-update channel: this raises the instrumental benefit, inducing voters who clear the cost threshold to turn out.

Testing this mechanism would ideally require an individual measure of *perceived accuracy (or precision)* in one's own beliefs, which our data lack. We therefore investigate whether the treatment reduces the variance of perceived pivotality *across* respondents in the treatment group relative to the control group. We note that this analysis is exploratory and was not pre-registered in the PAP.

We find evidence consistent with this channel. Among swing voters with stringent priors—the subgroup for which we observe a strong turnout effect—we apply the Brown-Forsythe test to the residuals from a regression of perceived pivotality on baseline controls and prefecture fixed effects. The median-based statistic is $W_{50} = 4.82$ ($p = 0.028$), indicating that the treatment group exhibits a significantly lower residual variance ($s = 0.725$) than the control group ($s = 0.798$).[7] Using the alternative pivotality measure—the perceived incumbent vote share—produces a similar result ($W_{50} = 7.09$, $p = 0.01$).

By contrast, we find no significant difference in residual dispersion between treated and control groups among lenient loyalists ($W_{50} = 1.35$, $p = 0.25$ for perceived pivotality; $W_{50} = 0.88$, $p = 0.35$ for the incumbent-vote-share measure).

Thus, while the treatment leaves the perceived mean vote share unchanged, it reduces its dispersion. The resulting rise in expected pivotality can prompt additional voters—especially swing voters—to turn out, suggesting that a variance-shrink channel, rather than a mean-update channel, explains the observed turnout increase.

Moral reinforcement. Although the variance-shrink channel is plausible and finds some support in the data, the same turnout pattern can also arise from a *moral-reinforcement* channel, which connects to the literature on the role of morals in voting (e.g., Enke, 2020; Walter and Redlawsk, 2019). Under this alternative mechanism, the signal increases each voter's moral cost of re-electing a corrupt incumbent:

$$\theta_i^0 \longrightarrow \theta_i^1 > \theta_i^0 \quad \text{(moral reinforcement)}.$$

[7] Test statistics are computed with Stata's `robvar` command. The statistic without controls yields $W_{50} = 5.11$ ($p = 0.024$).

We model this as a common multiplicative shift, $\theta_i^1 = g\,\theta_i^0$ with $g > 1$.

Moral reinforcement predicts higher turnout, because a larger moral cost scales up voters' instrumental benefit, even though, like the variance-shrink channel, it leaves perceived pivotality unchanged.[8]

Our survey cannot test this channel directly, but we proxy an individual's moral cost with her support for measures that increase transparency in political funding. Specifically, we look at (i) stated support for pro-transparency policies and (ii) the intention to donate to NGOs that promote transparency.

Table 2 provides evidence consistent with this channel. Among swing voters, the information treatment raises support for pro-transparency policies, whereas no significant change appears for party loyalists. By contrast, the treatment effect on intended donations is statistically insignificant for both groups.

Loyalty reinforcement. Finally, among lenient loyalists, we observe a non-trivial fraction switching their ballot from the challenger to the incumbent, while the information treatment *does not* increase turnout. Since none of the previously discussed mechanisms can account for this pattern, we introduce an *identity-reinforcement* channel, relating to studies that emphasize partisan loyalty as a key reason for the weak electoral punishment of corrupt politicians (e.g., Anduiza et al., 2013; Solaz et al., 2019).

According to this mechanism, loyalists ($\eta_i = 1$) receive an expressive payoff ϕ from supporting their party. The signal reinforces this identity:

$$\phi^0 \longrightarrow \phi^1 > \phi^0 \quad \text{(identity reinforcement)}.$$

Because of this change, for always voters ($B_i^0(p), B_i^1(p) \geq c$), turnout is unaffected. The treatment merely raises the identity payoff ϕ enough to flip $\Delta_i = \theta_i - \phi$ and hence the ballot from C to I. In contrast to previous studies, our framework explicitly captures this *reinforcement* of party loyalty, triggering defensive behavior. We proxy individual party loyalty using perceptions of the incumbent's competence and trustworthiness in the empirical

[8] Moral reinforcement is closely related to the concept of "ethical voting," which posits that individuals make voting decisions based on ethical considerations. In contrast to the formulation in Feddersen and Sandroni (2006), however, moral reinforcement may also stem from self-interested motives grounded in one's own moral values.

analysis.

Evidence in Table 3 is consistent with this mechanism. While the treatment has no effect on swing voters' perceptions of competence or trustworthiness, we find strong positive effects on both outcomes among LDP supporters, especially those with initially lenient views.

Changing the identity parameter ϕ should leave the pivotal probability $\rho(p)$ unchanged. Yet we observe higher reported pivotality among lenient loyalists. A possible interpretation is that respondents implicitly value pivotality as

$$\text{reported pivotality} \propto \rho(p) \times |\Delta_i|.$$

Under identity reinforcement, an increase in ϕ reduces Δ_i; for loyalists near the margin, this can flip their choice from challenger to incumbent, increasing the distance from indifference. As a result, $|\Delta_i|$—and thus perceived pivotality—rises, even though $\rho(p)$ itself is unchanged. By contrast, for swing voters ($\eta_i = 0$), the gap $|\Delta_i|$ is unaffected, so no analogous rise is expected.

Taken together, the evidence points to two distinct mechanisms: for swing voters the treatment most plausibly worked through *variance shrink* and *moral reinforcement*, whereas for loyalists it operated primarily via *identity reinforcement*.

7 Discussion

7.1 Effect size

A seminal field experiment by Gerber and Green (2000) delivered messages stressing civic duty, electoral closeness, and neighborhood solidarity through door-to-door canvassing, direct mail, and commercial phone calls. The largest effect—8.7 percentage points—came from personal canvassing. Subsequent work confirms that face-to-face contact is typically the most effective mobilization tactic (Green et al., 2013; Gerber and Green, 2019). Even without personal contact, however, Gerber et al. (2008) achieved a similar impact using mailers that leveraged social pressure—threatening to "publicize" who does and does not vote—which

increased participation by 8.1 percentage points.[9]

Our design involved neither personal canvassing nor social pressure, yet still produced a sizable effect: six percentage points.[10] A plausible explanation lies in the nature of our experimental setting. By design, the treatment provides voters with precise information about others' tolerance of corruption. For stringent swing voters, the message revealed that many fellow citizens were surprisingly lenient toward the scandal. This discrepancy appears to have sharpened their own moral aversion to re-electing a corrupt incumbent and motivated them to turn out. Because no social sanctions were at stake, the increase in turnout is more plausibly driven by personal moral considerations than by fear of external punishment.

7.2 Alternative explanations

The main findings of the paper are that informing voters about intolerance norms increases turnout and support for challengers, especially among swing voters, while it increases support for incumbents among ruling party loyalists. Although we find evidence consistent with mechanisms such as moral reinforcement and identity reinforcement, other explanations are possible.

First, issue salience. Boas et al. (2019) argue that one possible reason for the lack of punishment in their information treatment is the greater salience of other policy concerns. Our treatment might have made the corruption issue more salient compared to other policy issues. In our setting, however, the scandal was already a major issue in the election, with high media coverage beforehand. Given this existing salience, it is unlikely that our treatment had an additional salience effect. Moreover, this mechanism *per se* does not explain the heterogeneous results between swing voters and ruling party loyalists.

To test this, we examine whether treatment effects differ based on whether respondents listed the slush fund issue as one of the three most important issues in the October election during the baseline survey. Columns (1) and (2) in Table A.14 in the Online Appendix show significant effects among *both* those who did not consider the slush fund issue important and

[9] Other studies, such as Funk (2010) and Dellavigna et al. (2016), also show that social concerns affect voter turnout.

[10] This is a relatively large effect, even when compared to the two percentage point increase in turnout found by Banerjee et al. (2011), who provided voters with report cards detailing the incumbent's performance and candidate qualifications.

those who did. Moreover, the estimate is slightly larger for those who *already* considered the issue important. Thus, the salience channel is unlikely to be an alternative explanation.

Second, corruption prevalence perceptions. Previous studies suggest that (perceptions of) the prevalence of corruption affect whether voters punish corruption (Klašnja and Tucker, 2013; Vera, 2019; Arias et al., 2022). Since our treatment concerns others' tolerance of corruption while holding information about the scandal itself constant, it should not affect voters' perceptions about the prevalence of corruption.

We test this by splitting the sample based on respondents' baseline perceptions of the percentage of politicians involved in the slush fund, using the median as a cutoff. Columns (3) and (4) in Table A.14 show no systematic differences between those who perceive high versus low corruption prevalence.

In sum, these alternative explanations do not account for our main findings.

8 Conclusion

This study revisits a fundamental question of democratic accountability: Do voters punish corrupt incumbents, even when strategic considerations are at play?

Using a pre-registered online survey experiment conducted during Japan's 2023–2024 political funding scandal, we examine how strategic considerations about others, in particular, perceptions of intolerance norms against corruption, affect voting behavior. By providing information about the intolerance norms against the scandal, we find that it increased turnout and support for challengers. In contrast, we find that the same treatment increased support for incumbents among ruling party loyalists.

These findings underscore the importance of considering not only voters' individual preferences, but also their beliefs about others. Perceptions of prevailing norms shape how individuals interpret political information and respond to it—especially when strategic considerations are involved.

To explain these patterns, we develop a simple model of electoral accountability incorporating strategic considerations about others. In addition to the mechanisms related to perceived pivotality, the model includes two key mechanisms: *moral reinforcement*, in which norm information raises the moral cost of supporting a corrupt incumbent; and *identity*

reinforcement, in which the same information threatens partisan identity and triggers defensive alignment. The observed heterogeneity in treatment effects is consistent with these mechanisms.

Overall, our results suggest that the accountability mechanisms emphasized in workhorse models can operate through norm-based mobilization—particularly via swing voters—but their effectiveness depends on the social and partisan context in which information is received. By highlighting how perceptions of prevailing norms shape voter behavior in the presence of strategic considerations, this study contributes to understanding how democratic institutions can remain resilient in the face of political misconduct.

As with any experimental study, these findings should be interpreted with care. The external validity of the effects may depend on features of the political context, the salience of the scandal, and the nature of the norm signal. Future research could explore whether similar mechanisms operate in other democratic settings, particularly in environments characterized by high polarization or institutional distrust.

References

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2019). Democracy Does Cause Growth. *Journal of Political Economy 127*(1), 47–100.

Adida, C., J. Gottlieb, E. Kramon, and G. McClendon (2020). When Does Information Influence Voters? The Joint Importance of Salience and Coordination. *Comparative Political Studies 53*(6), 851–891.

Agerberg, M. (2020). The lesser evil? corruption voting and the importance of clean alternatives. *Comparative Political Studies 53*(2), 253–287.

Agranov, M., J. K. Goeree, J. Romero, and L. Yariv (2018). Equilibrium selection in games with entry: An experimental study. *Journal of the European Economic Association 16*(4), 1091–1135.

Ahler, D. J. (2014). Self-Fulfilling Misperceptions of Public Polarization. *The Journal of Politics 76*(3), 607–620.

Ahler, D. J. and G. Sood (2018). The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences. *The Journal of Politics 80*(3), 964–981.

Andre, P., T. Boneva, F. Chopra, and A. Falk (2024). Misperceived Social Norms and Willingness to Act Against Climate Change. *Review of Economics and Statistics forthcoming*.

Anduiza, E., A. Gallego, and J. Muñoz (2013). Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption. *Comparative Political Studies 46*(12), 1664–1692.

Arias, E., H. Larreguy, J. Marshall, and P. Querubín (2022). Priors Rule: When Do Malfeasance Revelations Help Or Hurt Incumbent Parties? *Journal of the European Economic Association 20*(4), 1433–1477.

Asahi Shimbun (2024). Cabinet Support Drops to 21%, Lowest Since LDP's Return to Power; Party Support Also at 21%, Asahi Poll Finds. https://www.asahi.com/articles/ASS2L7FLZS2HUZPS004.html (published 2024-2-18; accessed: 2024-12-26).

Associated Press (2024). Donald trump is convicted of a felony. here's how that affects the 2024 presidential race. `https://apnews.com/article/donald-trump-verdict-republican-voters-2024-86f481bb54753ffb2dc3f6971d89da1d`. Accessed: 2025-08-14.

Avis, E., C. Ferraz, and F. Finan (2018). Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians. *Journal of Political Economy 126*(5), 1912–1964.

Banerjee, A. and R. Pande (2009). Parochial politics: Ethnic preferences and politician corruption. Unpublished manuscript, Harvard University.

Banerjee, A. V., S. Kumar, R. Pande, and F. Su (2011). Do Informed Voters Make Better Choices? Experimental Evidence from Urban India. *Working paper*.

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice 14*(1), 19–42.

BBC News (2019). Silvio berlusconi: Italy ex-pm cleared of corruption charges. `https://www.bbc.com/news/world-europe-48745162`. Accessed: 2025-08-14.

Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford: Oxford University Press.

Boas, T. C., F. D. Hidalgo, and M. A. Melo (2019). Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil. *American Journal of Political Science 63*(2), 385–400.

Bobonis, G. J., L. R. C. Fuertes, and R. Schwabe (2016). Monitoring Corruptible Politicians. *American Economic Review 106*(8), 2371–2405.

Buntaine, M. T., R. Jablonski, D. L. Nielson, and P. M. Pickering (2018). SMS texts on corruption help Ugandan voters hold elected councillors accountable at the polls. *Proceedings of the National Academy of Sciences 115*(26), 6668–6673.

Bursztyn, L., D. Cantoni, P. Funk, F. Schönenberger, and N. Yuchtman (2024). Identifying the effect of election closeness on voter turnout: Evidence from swiss referenda. *Journal of the European Economic Association 22*(2), 876–914.

Bursztyn, L. and D. Y. Yang (2022). Misperceptions About Others. *Annual Review of Economics 14*(1), 425–452.

Bøttkjær, L. and M. K. Justesen (2021). Why do voters support corrupt politicians? experimental evidence from south africa. *The Journal of Politics 83*(2), 788–793.

Cameron, A., J. Gelbach, and D. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cantoni, D., D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2019). Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement*. *The Quarterly Journal of Economics 134*(2), 1021–1077.

Carlson, T. N. and S. J. Hill (2022). Experimental Measurement of Misperception in Political Beliefs. *Journal of Experimental Political Science 9*(2), 241–254.

Chambers, J. R., R. S. Baron, and M. L. Inman (2005). Misperceptions in Intergroup Conflict. *Psychological Science 17*(1), 38–45.

Chang, E. C. C., M. A. Golden, and S. J. Hill (2010). Legislative malfeasance and political accountability. *World Politics 62*(2), 177–220.

Chong, A., A. L. De La O, D. Karlan, and L. Wantchekon (2015). Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification. *The Journal of Politics 77*(1), 55–71.

Correia, S., P. Guimarães, and T. Zylkin (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal 20*(1), 95–115.

Cox, G. W. (1997). *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.

Cruz, C., P. Keefer, and J. Labonne (2020). Buying Informed Voters: New Effects of Information on Voters and Candidates. *The Economic Journal 131*(635), 1105–1134.

de Figueiredo, M. F., F. D. Hidalgo, and Y. Kasahara (2023). When Do Voters Punish Corrupt Politicians? Experimental Evidence from a Field and Survey Experiment. *British Journal of Political Science 53*(2), 728–739.

Dellavigna, S., J. A. List, U. Malmendier, and G. Rao (2016). Voting to tell others. *The Review of Economic Studies 84*(1), 143–181.

Duffy, J. and M. Tavits (2008). Beliefs and voting decisions: A test of the pivotal voter model. *American Journal of Political Science 52*(3), 603–618.

Dunning, T., G. Grossman, M. Humphreys, S. D. Hyde, C. McIntosh, G. Nellis, C. L. Adida, E. Arias, C. Bicalho, T. C. Boas, M. T. Buntaine, S. Chauchard, A. Chowdhury, J. Gottlieb, F. D. Hidalgo, M. Holmlund, R. Jablonski, E. Kramon, H. Larreguy, M. Lierl, J. Marshall, G. McClendon, M. A. Melo, D. L. Nielson, P. M. Pickering, M. R. Platas, P. Querubín, P. Raffler, and N. Sircar (2019). Voter information campaigns and political accountability: Cumulative findings from a preregistered meta-analysis of coordinated trials. *Science Advances 5*(7), eaaw2612.

Eggers, A. C. and N. Vivyan (2020). Who votes more strategically? *American Political Science Review 114*(2), 470–485.

Enke, B. (2020). Moral values and voting. *Journal of Political Economy 128*(10), 3679–3729.

Enos, R. D. and A. Fowler (2014). Pivotality and turnout: Evidence from a field experiment in the aftermath of a tied election. *Political Science Research and Methods 2*(2), 309–319.

Feddersen, T. and A. Sandroni (2006). A theory of participation in elections. *American Economic Review 96*(4), 1271–1282.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1-3), 5–25.

Ferraz, C. and F. Finan (2008). EXPOSING CORRUPT POLITICIANS: THE EFFECTS OF BRAZIL'S PUBLICLY RELEASED AUDITSON ELECTORAL OUTCOMES. *The Quarterly Journal of Economics 123*, 703–745.

Ferraz, C. and F. Finan (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review 101*(4), 1274–1311.

Funk, P. (2010). Social incentives and voter turnout: Evidence from the swiss mail ballot system. *Journal of the European Economic Association 8*(5), 1077–1103.

Gerber, A., M. Hoffman, J. Morgan, and C. Raymond (2020). One in a Million: Field Experiments on Perceived Closeness of the Election and Voter Turnout. *American Economic Journal: Applied Economics 12*(3), 287–325.

Gerber, A. S. and D. P. Green (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review 94*(3), 653–663.

Gerber, A. S. and D. P. Green (2019). *Get Out the Vote: How to Increase Voter Turnout* (4th ed.). Washington, D.C.: Brookings Institution Press.

Gerber, A. S., D. P. Green, and C. W. Larimer (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review 102*(1), 33–48.

Green, D. P., M. C. McGrath, and P. M. Aronow (2013). Field experiments and the study of voter turnout. *Journal of Elections, Public Opinion and Parties 23*(1), 27–48.

Haaland, I., C. Roth, and J. Wohlfart (2023). Designing information provision experiments. *Journal of Economic Literature 61*(1), 3–40.

Hager, A., L. Hensel, J. Hermle, and C. Roth (2023). Political Activists as Free Riders: Evidence from a Natural Field Experiment. *The Economic Journal 133*(653), 2068–2084.

India Today (2024). Glass half empty: Nearly half of lok sabha mps face criminal charges. `https://www.indiatoday.in/diu/story/glass-half-empty-nearly-half-of-lok-sabha-mps-face-criminal-charges-2553578-2024-06-15`. Accessed: 2025-08-14.

Kawai, K. and Y. Watanabe (2013). Inferring Strategic Voting. *American Economic Review 103*(2), 624–662.

Klašnja, M. and J. A. Tucker (2013). The economy, corruption, and the vote: Evidence from experiments in sweden and moldova. *Electoral Studies 32*(3), 536–543.

Krupka, E. L. and R. A. Weber (2013). IDENTIFYING SOCIAL NORMS USING COORDINATION GAMES: WHY DOES DICTATOR GAME SHARING VARY? *Journal of the European Economic Association 11*(3), 495–524.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies 76*(3), 1071–1102.

Levendusky, M. S. and N. Malhotra (2016). (Mis)perceptions of Partisan Polarization in the American Public. *Public Opinion Quarterly 80*(S1), 378–391.

Levine, D. K. and T. R. Palfrey (2007). The paradox of voter participation? a laboratory study. *American Political Science Review 101*(1), 143–158.

Mildenberger, M. and D. Tingley (2019). Beliefs about Climate Beliefs: The Importance of Second-Order Opinions for Climate Politics. *British Journal of Political Science 49*(4), 1279–1307.

NPR (2023). Party bosses fall in Japan's worst political corruption scandal in decades. https://www.npr.org/2023/12/22/1221230635/japan-alleged-political-corruption-ldp-slush-fund (published 2023-12-22; accessed: 2025-6-3).

Pavão, N. (2018). Corruption as the only option: The limits to electoral accountability. *The Journal of Politics 80*(3), 996–1010.

Perez-Truglia, R. and G. Cruces (2017). Partisan Interactions: Evidence from a Field Experiment in the United States. *Journal of Political Economy 125*(4), 1208–1243.

Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

Reed, S. R. (1996). Political corruption in japan. *International Social Science Journal 48*(149), 395–405.

Rundquist, B. S., G. S. Strom, and J. G. Peters (1977). Corrupt politicians and their electoral support: Some experimental observations. *American Political Science Review 71*(3), 954–963.

Solaz, H., C. E. D. Vries, and R. A. d. Geus (2019). In-group loyalty and the punishment of corruption. *Comparative Political Studies 52*(6), 896–926.

Vera, S. B. (2019). Accepting or resisting? citizen responses to corruption across varying levels of competence and corruption prevalence. *Political Studies 68*(3), 653–670.

Walter, A. S. and D. P. Redlawsk (2019). Voters' partisan responses to politicians' immoral behavior. *Political Psychology 40*(5), 1075–1097.

Weitz-Shapiro, R. and M. S. Winters (2017). Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil. *The Journal of Politics 79*(1), 60–74.

Winters, M. S. and R. Weitz-Shapiro (2013). Lacking information or condoning corruption? when do voters support corrupt politicians? *Comparative Politics 45*(4), 418–436.

Yomiuri Shimbun (2024). Accounting Officers of Three Factions Indicted for Failing to Report 970 Million Yen in Total; Indictment of Abe Faction Executives Dropped. https://www.yomiuri.co.jp/national/20240119-OYT1T50211/ (published: 2024-1-13; accessed: 2024-12-26).

Young, A. (2019). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics 134*(2), 557–598.

Figures and Tables

Figure 1: Timeline of the Experiment

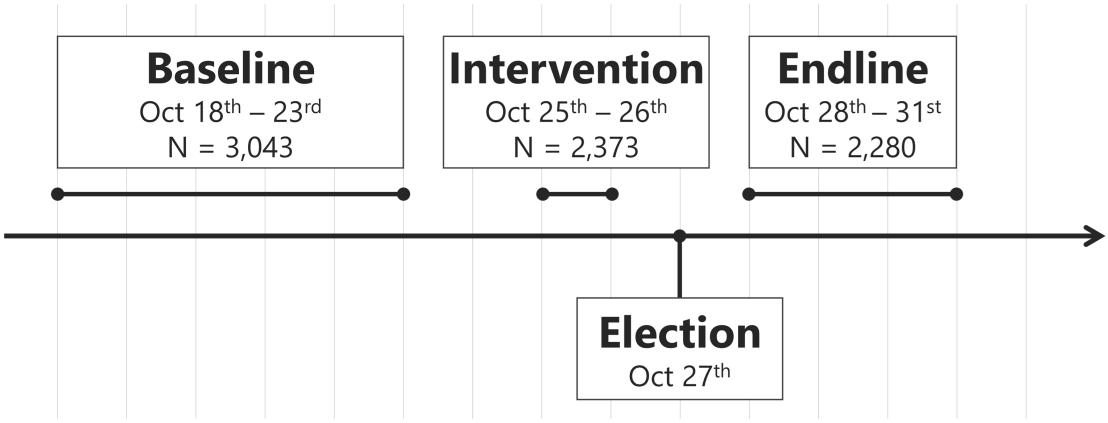


Figure 2: Information Treatment Screen

This survey project brings together a diverse group of participants, whose responses can be said to represent the views and attitudes of the average voter in Japan. In the previous survey, participants were asked how many of the other respondents considered that "omissions in political funding reports or kickbacks from political funds" are "completely unacceptable" or "somewhat unacceptable."

As a result, it was found that people believe that out of 100 participants, 67 considered "omissions in political funding reports or kickbacks from political funds" to be either "completely unacceptable" or "somewhat unacceptable."

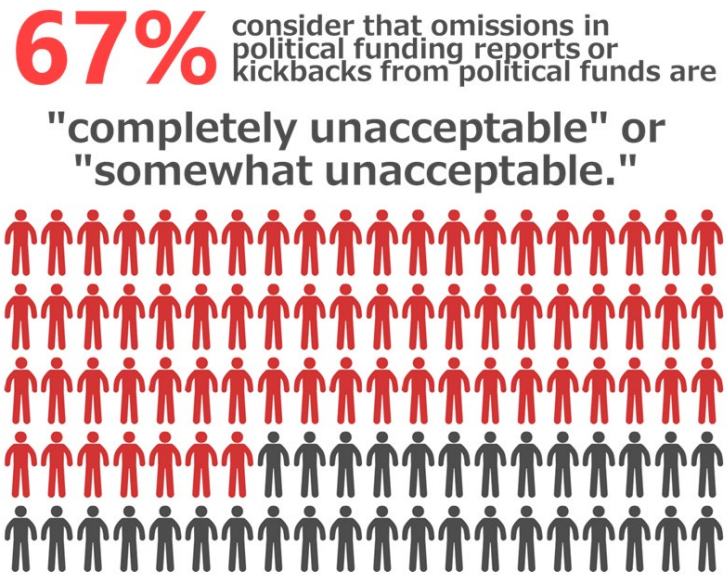
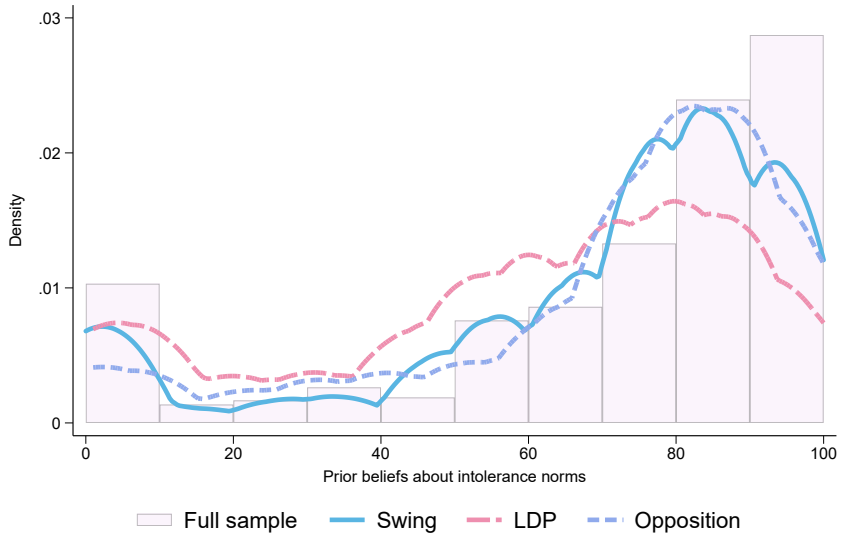
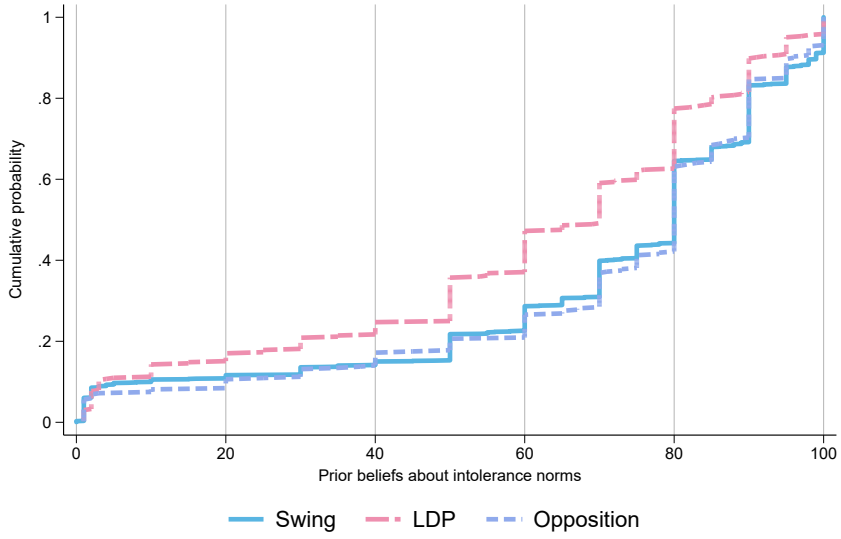


Figure 3: Distribution of Prior Beliefs about Intolerance Norms



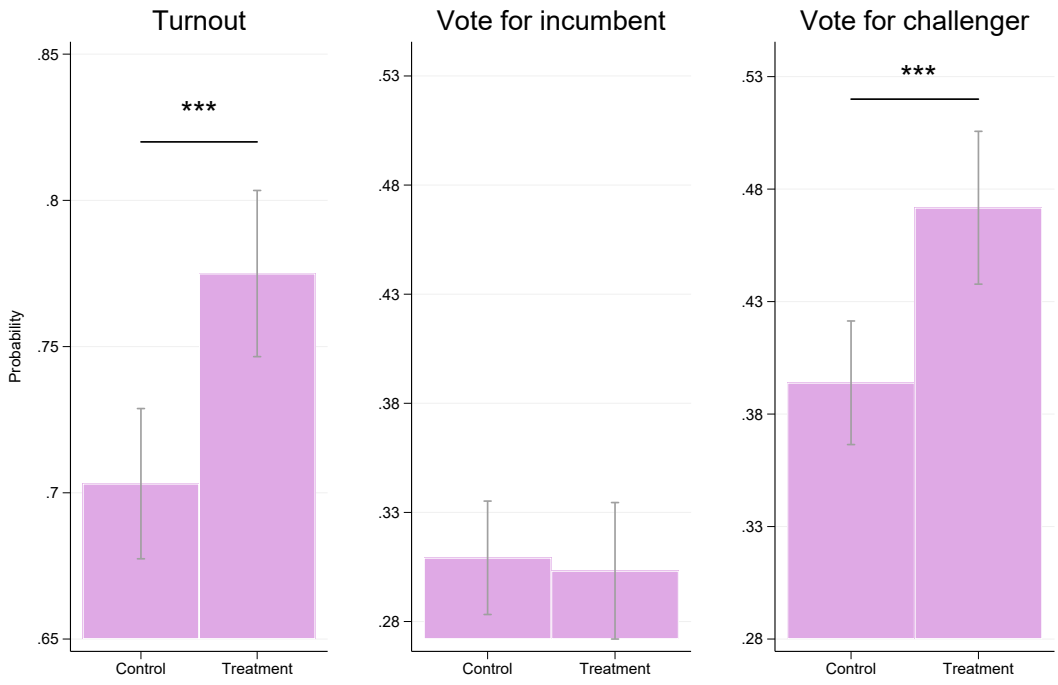
(a) Kernel density



(b) Cumulative distribution

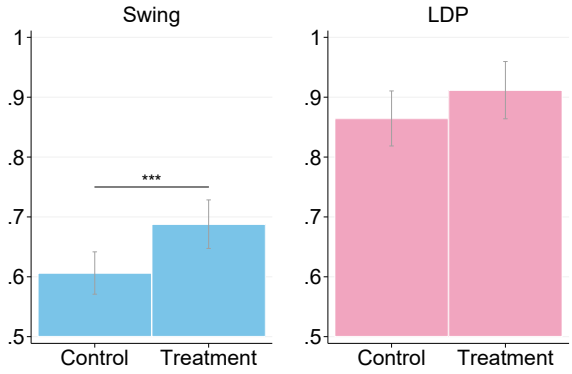
Note: The figures show the distributions of prior beliefs about intolerance norms. Panel (a) presents the kernel density estimates for swing voters, defined as those who do not support any particular party, LDP supporters, and opposition-party supporters. The background histogram reflects the full sample distribution. Panel (b) displays the cumulative distribution function for each group.

Figure 4: Main Effects of the Information Treatment on Voting Outcomes

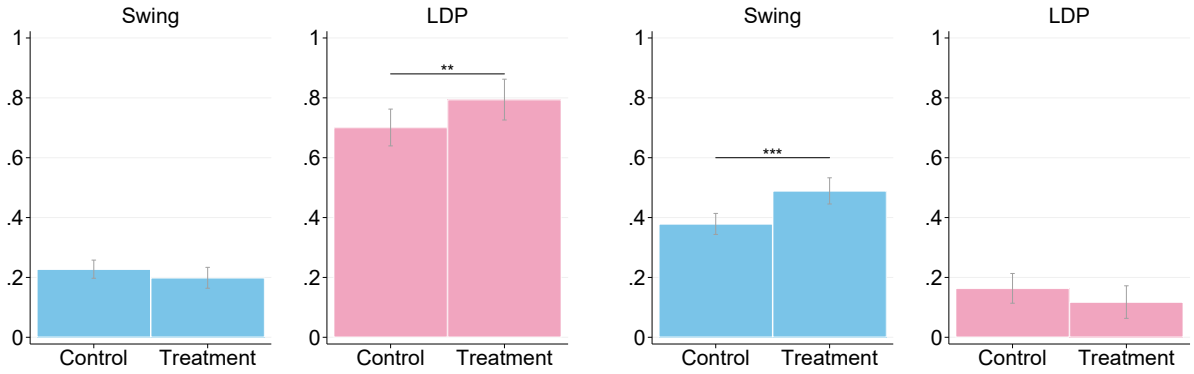


Note: Each bar represents the group average with a 95% confidence interval. The leftmost bars use a turnout dummy as the outcome; the middle bars use a dummy equal to one if the respondent voted for the incumbent; and the rightmost bars use a dummy equal to one if the respondent voted for the challenger. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Figure 5: Heterogeneous Effects of the Information Treatment on Voting Outcomes (Swing Voters and LDP Supporters)



(a) Turnout

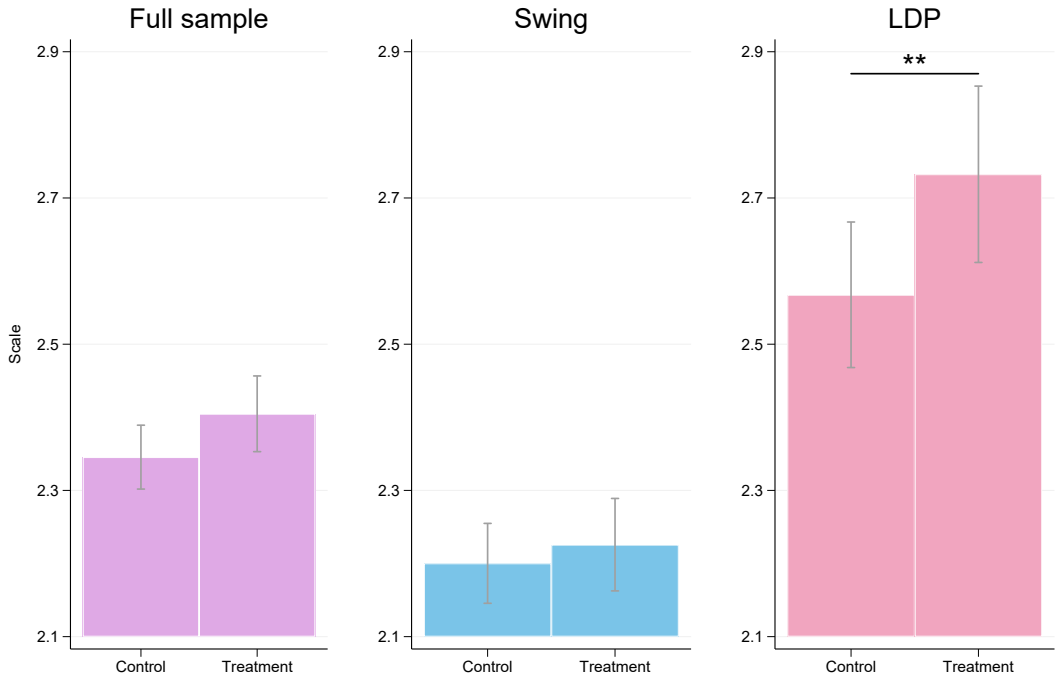


(b) Vote for incumbent

(c) Vote for challenger

Note: Each bar represents the group average with a 95% confidence interval. Panel (a) uses a turnout dummy as the outcome; Panel (b) uses a dummy equal to one if the respondent voted for the incumbent; and Panel (c) uses a dummy equal to one if the respondent voted for the challenger. In each panel, the left two bars represent swing voters, and the right two bars represent LDP supporters. *, **, and *** indicate p <0.10, p <0.05, and p <0.01, respectively.

Figure 6: Effects on Posterior Beliefs about Pivotality



Note: Each bar represents the group average with a 95% confidence interval. The outcome is posterior beliefs about pivotality, measured by responses to the question: "How much influence do you think your vote has on the election outcome?" (1 = Not at all, 2 = Not much, 3 = Some, 4 = A great deal) in the pre-election (post-treatment) survey. The leftmost bars represent the full sample, the middle bars represent swing voters, and the rightmost bars represent LDP supporters. *, **, and *** indicate p <0.10, p <0.05, and p <0.01, respectively.

Table 1: Effects of the Information Treatment on Voting Outcomes, by District Type

Sample:	Fraud district			Clean district		
		Swing	LDP		Swing	LDP
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	Dependent variable: Turnout					
Treatment	0.029	0.020	0.007	0.099	0.128	0.067
	(0.028)	(0.039)	(0.053)	(0.027)***	(0.038)***	(0.046)
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	0.735	0.667	0.857	0.670	0.545	0.872
Control std.	(0.442)	(0.472)	(0.352)	(0.471)	(0.499)	(0.336)
R^2	0.04	0.06	0.08	0.05	0.06	0.03
N	994	597	163	1053	640	187
Panel B	Dependent variable: Vote for incumbent					
Treatment	−0.058	−0.100	0.059	0.031	0.031	0.128
	(0.028)**	(0.031)***	(0.070)	(0.030)	(0.035)	(0.062)**
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	0.288	0.220	0.686	0.332	0.236	0.716
Control std.	(0.453)	(0.414)	(0.466)	(0.471)	(0.425)	(0.453)
R^2	0.02	0.03	0.12	0.02	0.03	0.08
N	994	597	163	1053	640	187
Panel C	Dependent variable: Vote for challenger					
Treatment	0.086	0.120	−0.052	0.068	0.097	−0.061
	(0.032)***	(0.041)***	(0.057)	(0.030)**	(0.038)**	(0.047)
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	0.447	0.447	0.171	0.338	0.310	0.156
Control std.	(0.498)	(0.498)	(0.379)	(0.474)	(0.463)	(0.364)
R^2	0.04	0.07	0.10	0.03	0.05	0.09
N	994	597	163	1053	640	187

Note. Robust standard errors are in parentheses. In Panel A, the dependent variable is whether the respondent turned out to vote; in Panel B, whether the respondent voted for the incumbent; and in Panel C, whether the respondent voted for a challenger. All columns use actual voting outcomes from the post-election survey. Columns (1)-(3) include individuals in fraud constituencies, while Columns (4)-(6) include individuals in clean constituencies. Columns labeled "Swing" comprise respondents who do not support a particular party, whereas columns labeled "LDP" comprise LDP supporters. Baseline controls include educational attainment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 2: Effects of the Information Treatment on Support for Transparency in Political Funding

Sample:	Swing			LDP		
	All	Stringent	Lenient	All	Stringent	Lenient
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	Dependent variable: Policy support					
Treatment	0.156	0.113	0.207	0.079	0.174	−0.077
	(0.054)***	(0.065)*	(0.098)**	(0.079)	(0.114)	(0.109)
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	2.725	2.837	2.500	2.865	2.933	2.803
Control std.	(0.999)	(1.005)	(0.952)	(0.724)	(0.737)	(0.710)
R^2	0.02	0.01	0.04	0.06	0.06	0.12
N	1392	965	427	364	187	177
Panel B	Dependent variable: Donation					
Treatment	−312.400	6.791	−1122.188	−153.346	−811.189	333.010
	(347.594)	(394.088)	(720.598)	(927.424)	(1582.643)	(1126.759)
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	2248.588	1947.600	2857.105	3260.302	3281.143	3241.598
Control std.	(7462.502)	(6712.429)	(8767.616)	(7482.803)	(8471.179)	(6506.688)
R^2	0.00	0.01	0.03	0.02	0.04	0.09
N	1392	965	427	364	187	177

Note. Robust standard errors are in parentheses. In Panel A, the dependent variable is support for greater transparency in political funding, measured in the pre-election (post-treatment) survey by responses to the question: "To what extent do you support policies that aim to make the flow of political funds more transparent, including legal reforms?" (1 = Do not support at all, 2 = Do not support much, 3 = Somewhat support, 4 = Strongly support). In Panel B, the dependent variable is the amount (in Japanese yen) the respondent would donate to an NGO that promotes transparency in political funding, measured in the same survey by the question: "Suppose you have 50,000 yen at your disposal. How much of it would you donate to an NGO working to increase the transparency of political funding?" Columns labeled "Swing" comprise respondents who do not support a particular party, whereas columns labeled "LDP" comprise LDP supporters. Columns labeled "Stringent" contain respondents whose prior beliefs about intolerance norms are above the sample mean, whereas columns labeled "Lenient" contain respondents whose prior beliefs are at or below the mean. Columns labeled "All" pool these sub-samples. Baseline controls include educational attainment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 3: Effects of the Information Treatment on Perceptions of the Incumbent's Valence (Competence and Trustworthiness)

Sample:	Swing			LDP		
	All	Stringent	Lenient	All	Stringent	Lenient
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	Dependent variable: Competence					
Treatment	1.424	1.543	0.807	6.007	1.801	10.456
	(1.132)	(1.345)	(2.126)	(2.632)**	(3.576)	(3.849)***
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	30.722	31.425	29.301	34.604	36.971	32.479
Control std.	(21.458)	(21.480)	(21.382)	(23.907)	(22.167)	(25.273)
R^2	0.01	0.01	0.02	0.05	0.05	0.10
N	1392	965	427	364	187	177
Panel B	Dependent variable: Trustworthiness					
Treatment	1.649	1.818	1.239	6.592	3.349	9.576
	(1.104)	(1.306)	(2.089)	(2.557)**	(3.540)	(3.654)***
Baseline controls	yes	yes	yes	yes	yes	yes
Prefecture F.E.	yes	yes	yes	yes	yes	yes
Control mean	29.681	29.991	29.054	34.955	36.219	33.821
Control std.	(21.261)	(21.343)	(21.120)	(23.975)	(22.638)	(25.158)
R^2	0.01	0.01	0.02	0.05	0.07	0.08
N	1392	965	427	364	187	177

Note. Robust standard errors are in parentheses. Panel A uses respondents' beliefs about the incumbent's competence as the dependent variable. Specifically, each respondent estimates the share of other voters who view the incumbent as "very competent" or "competent" in the pre-election (post-treatment) survey; higher values indicate a belief that more people consider the incumbent competent. Panel B uses an analogous measure of perceived trustworthiness. Respondents estimate the share of voters who regard the incumbent as "very trustworthy" or "trustworthy" in the same survey; higher values indicate a belief that more people consider the incumbent trustworthy. Columns labeled "Swing" comprise respondents who do not support a particular party, whereas columns labeled "LDP" comprise LDP supporters. Columns labeled "Stringent" contain respondents whose prior beliefs about intolerance norms are above the sample mean, whereas columns labeled "Lenient" contain respondents whose prior beliefs are at or below the mean. Columns labeled "All" pool these sub-samples. Baseline controls include educational attainment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.