Greater Expectations: A Field Experiment to Improve Accountability in Mali

Author(s): Jessica Gottlieb

Source: *American Journal of Political Science*, JANUARY 2016, Vol. 60, No. 1 (JANUARY 2016), pp. 143-157

Published by: Midwest Political Science Association

Stable URL: https://www.jstor.org/stable/24583055

Greater Expectations: A Field Experiment to Improve Accountability in Mali

Jessica Gottlieb Texas A&M University

I argue that if citizens systematically underestimate what their government can and should do for them, then they will hold politicians to a lower standard and sanction poor performers less often. A field experiment across 95 localities in Mali in which randomly assigned localities receive a civics course identifies the effect of raising voter expectations of government on their willingness to hold leaders accountable. The course provides information about local government capacity and responsibility as well as how local politicians perform relative to others, effectively raising voter expectations of what local governments can and should do. Survey experiments among individuals in treated and control communities ($N = 5{,}560$) suggest that people in treated villages are indeed more likely to sanction poor performers and vote based on performance more often. A behavioral outcome—the likelihood that villagers challenge local leaders at a town hall meeting—adds external validity to survey findings.

The recent emergence of free and fair elections in many developing countries has frequently failed to produce improvements in government performance. Cross-country studies show the expected positive correlation between democratic institutions and public goods provision breaks down in poor places (Boix 2001; Collier and Rohner 2008; Ross 2006).[1] Some existing explanations are that poverty and inequality facilitate patronage politics (Pande 2007) and reduce public-sector wages (Montinola and Jackman 2002), improving opportunities for political corruption. So are poor and unequal societies stuck with failing democratic institutions? Motivated by insights from formal theory and a puzzle presented by the empirical literature, this article offers one novel explanation for the failure of democratic accountability—low voter expectations of what governments can do for them. It discusses findings of a randomized intervention in one African country designed to raise voter standards of politician performance.

There is consensus in the theoretical literature that information about government performance is a necessary condition for political accountability (Barro 1973; Besley 2006; Ferejohn 1986). When voters are uncertain about the actions of their politicians, good performers cannot guarantee reelection, so it is harder for voters to motivate them to perform well or respond to their interests. It is thus puzzling why some recent attempts to increase voter information have failed to improve politician accountability (Banerjee, Green et al. 2010; Chong et al. 2012; de Figueiredo, Hidalgo, and Kasahara 2011; Lieberman, Posner, and Tsai 2014). I argue that performance information on its own is not sufficient; voters also require an appropriate reference point against which they can evaluate or benchmark their own politician's performance.

Salient reference points for voters include general *performance standards* or the *relative performance* of politicians in comparable polities. Due to low levels of civic education and poor access to credible media outlets, the

Jessica Gottlieb is Assistant Professor, Bush School of Government & Public Service, Texas A&M University, 4220 TAMU, College Station, TX 77843 (jgottlieb@tamu.edu).

I thank the National Science Foundation, the International Growth Centre, the Freeman Spogli Institute, and Stanford University for financial support, without which this project would not have been possible. I am grateful for invaluable guidance from James Fearon, Beatriz Magaloni, Jeremy Weinstein, and Saumitra Jha, and for feedback from Guy Grossman, Mike Tomz, two anonymous reviewers, and members of the IGC Working Group on Political Economy, FSI Action Fund seminar, and Stanford Comparative Politics Workshop. This research would not have been possible without the superior research assistance of Sidi Zeda and Peter LeFrancois, support from Marc Dabou at the DNCT and Fabio Germano at GIZ, as well as the tireless efforts of the Malian instructors and enumerators who crossed rivers and mountains to collect data. All errors are my own. Replication materials can be found at http://dx.doi.org/10.7910/DVN/28809, and the supporting information is posted on the *AJPS* website.

[1] After heralding the recent progress of political reform on the African continent, one of its foremost scholars of democratization admits, "standards of democratic performance tend to be low, even in the best of cases" (Gyimah-Boadi 2004, 10).

American Journal of Political Science, Vol. 60, No. 1, January 2016, Pp. 143–157

DOI: 10.1111/ajps.12186

information necessary to construct such reference points is particularly lacking in developing countries. Further, voters in young democracies are more likely to underestimate rather than overestimate politician performance when priors are informed by recent experience with weak, unaccountable governments and politicians have yet to form reputations that help weed out poor candidates (Keefer 2007; Svolik 2013). Indeed, Gyimah-Boadi blames poor performance on the "self-serving excuse of newly installed democratic regimes that ... they only have to perform slightly better or no worse than a previous autocratic regime" (2004, 10). Together, this suggests that voters in developing democracies will hold politicians to low performance standards and fail to sanction poor performers at election time, providing insufficient incentives for political accountability.

Low voter expectations of politicians have negative implications for both government accountability and the democratic system itself. Public welfare suffers in the short term when politicians perform poorly, or just meet low voter expectations, without fear of electoral retribution. In the long term, this bad equilibrium of voters failing to punish politicians and politicians performing poorly can fuel disillusionment with the democratic system itself.[2] Svolik (2013) argues that this expectation-driven failure of accountability, which he calls the "trap of pessimistic expectations," can eventually result in the breakdown of democracies. When the public has little esteem not only for individual politicians but also for the democratic system, citizens fail to defend their democracy from autocratic takeover.[3] I provide evidence for one way of breaking out of this trap of pessimistic expectations. Improving voter reference points for evaluating politician performance can increase expectations about what politicians can and should do for them, thus incentivizing better government performance.

Identifying the relationship between voter reference points and electoral behavior poses an inference problem. Citizens who are more informed are likely to be different in other respects from those who are less informed. I thus manipulate voter reference points by randomly assigning a civics course in 95 rural localities in one developing country, which ensures that treatment and control groups are equivalent, on average, save for the experimental treatment. Political decentralization further allows me to exploit variation across independently governed *communes* while controlling for cultural and institutional factors at the national level. The experimental treatment to manipulate voter reference points is a civics course provided to a subset of villages in treated communes. The course improves knowledge of performance standards for local government through information about budget size and mandated responsibilities. A second treatment arm additionally improves voter ability to benchmark politician performance by providing information about how the local government performs relative to neighboring ones.

Survey evidence evaluating the first step of the argument, that informing voter reference points increases expectations, reveals the civics course effectively raised participant expectations of local government performance. Evidence supporting the second step, that greater expectations lead to increased sanctioning of poor performers, comes from survey experiments embedded in candidate vignettes. I find that treatment improved voter willingness to (1) sanction poor-performing candidates and (2) vote on the basis of performance rather than kinship or gift giving. While these survey experiments are designed to mitigate bias, they remain artifactual and may not translate into real-world behavior. I thus examine impacts on one behavioral outcome: the likelihood that citizens challenge local leaders at a town hall meeting.

The main contribution of the article is demonstrating the relevance of a new type of information in voter decision making. I argue that voter uncertainty over the government's potential to perform constrains voters' ability to evaluate politicians, thus undermining government accountability. Extending classic moral hazard models of electoral accountability (Barro 1973; Besley 2006; Fearon 1999; Ferejohn 1986), a formalization of the argument models an information asymmetry between the voter and politician about the availability of a budget to provide public goods in a particular village. From the model, I derive testable implications about how voters will act when faced with new information about what they can expect of their governments.

This research also addresses mixed findings from previous studies on the impact of information on voter behavior. Recent information experiments in India (Banerjee, Green, et al. 2010), Mexico (Chong et al. 2012), and Brazil (de Figueiredo, Hidalgo, and Kasahara 2011) fail to improve voter sanctioning of poor performers. In these experiments, only performance information about candidates in the voter's district is provided. If voters already believe candidates are corrupt or think little is at stake in local government, then it

[2]Indeed, only 70% of respondents in Africa (Afrobarometer 2008), 60% in Latin America (AmericasBarometer 2008), and 56% in East Asia (AsiaBarometer 2005–8) say democracy is preferable to other regimes. Ireland was the only advanced industrial democracy that had rates as low as these; 88% of American respondents, for example, strictly prefer democracy (World Values Survey).

[3]This was recently illustrated in Mali when a disenchanted public failed to turn out in meaningful numbers to protest the 2012 military coup.

is not surprising that providing additional information about government misconduct has little or no effect on voter behavior (or even increases voter apathy in some cases). In contrast, a publicly released audit of Brazilian municipalities allowing voters to make *relative* performance evaluations results in improved sanctioning of corrupt incumbents (Ferraz and Finan 2008b).[4]

The finding that a civics course can inform voter beliefs about expected politician performance, as a potential input to increasing government accountability, is important for students of democracy in Africa and beyond. Where the transition to democracy is rapid and led by narrow segments of the population, the remaining majority has weak claims on the new regime. I show, in one such country, that this manifests as low expectations of politician performance and that a civics course can lead citizens to usefully update these expectations. With increased international attention to governance in developing nations, this study also has potential practical implications for policy makers about how to improve democratic accountability. While it is unlikely that incumbent governments will endogenously decide to provide information that constrains their own behavior, other actors could realistically do so. External donors and local nongovernmental organizations (NGOs) engage in promoting civic education and disseminating political information, and in decentralized countries like Mali, the federal government may be sufficiently interested in federal transfers being spent well to support an intervention aimed at improving local government performance.

Voter Expectations and Politician Accountability

In classic moral hazard models of electoral control (Barro 1973; Besley 2006; Fearon 1999; Ferejohn 1986), voters motivate politicians by conditioning electoral decisions on government performance in the previous term. They use a decision rule such as a cut point: If perceived performance exceeds a certain level, voters favor the incumbent; otherwise, the challenger wins. Voters face a trade–off—they want to set the cut point as high as possible to extract maximum utility from the incumbent, but setting it too high will discourage the incumbent, giving him or her incentive to shirk or underperform. The more precision with which voters know actual government policy, the higher they will set the cut point. In equilibrium, incumbents in a high-information setting will perform better due to this increased threat from voters. Predictions of these models substantiate empirical findings that increased access to information improves government accountability (Besley and Burgess 2002; Ferraz and Finan 2008a; Reinikka and Svensson 2005).

Modeling the problem in this way makes two implicit assumptions that while innocuous in some contexts may be problematic in others. First, the above setup assumes the primary dimension along which voters make decisions is a performance criterion, which is not necessarily the case in many developing country settings where gift giving or ethnic ties can prevail. Second, it is implicit that voters have a reasonable reference point against which to evaluate past performance. For voters to generate an appropriate cut point, they must know what good performance looks like.

I argue that voters, particularly those in developing democracies, may not have sufficient information to establish appropriate reference points. Especially where there are multiple levels of government, voters may be uncertain about the level of government that is responsible for providing a public service—and that uncertainty can benefit elected officials at the ballot box. This was the case in Afghanistan, where elected local councils failed to improve the equity of aid distribution when assignment of responsibility was unclear (Beath, Christia, and Enikopolov 2013). Voters may also be uncertain about the size of the government budget or the types of projects the government is capable of implementing. In an experiment in Mexico, reporting information about how much local governments spent *relative* to the total budget caused voters to turn out more and sanction incumbents more often when the percentage of funds spent was low (Chong et al. 2012). Finally, voters who are uninformed about their own politician's performance will be even more unlikely to know about the performance of other politicians. "Yardstick" competition, using the performance of comparable politicians as an indicator of a politician's potential (Besley and Case 1995; Shleifer 1985),[5] helps explain the success of an information experiment in New Delhi slums in which report cards of multiple candidates were distributed to voters (Banerjee, Kumar, et al. 2010).

I extend the classic moral hazard model to show that this new type of information asymmetry affects the voter's

[4]A similar intuition holds in nonelectoral contexts: Providing parents with information about student performance in Kenya had no discernible effect on parental engagement (Lieberman et al. 2014), and, in Peru, providing parents with information about performance standards increased parental involvement, closing the gap between educational performance and benchmarks (World Bank 2009).

[5]Benchmarking against comparable economies has also been shown to influence voters' decisions (Kayser and Peress 2012).

ability to hold politicians accountable. Dal Bó and Powell (2009) similarly extend asymmetric-information models of international conflict by introducing uncertainty over the size of state spoils in contrast to the standard uncertainty over the cost of fighting or distribution of power. In their model, an information asymmetry between government and opposition about "the size of the spoils or 'pie'" leads to costly and inefficient conflict over state resource allocation. In our case, voter uncertainty over the size of the pie leads to less accountable politicians.

The formalization (see supporting information [SI] A) explores the consequences of an information asymmetry between the voter and politician about the availability of a budget to provide public goods in a particular village. The voter can condition reelection of the incumbent either on the receipt of private campaign gifts or on gifts *and* the provision of public goods. If the probability of there being a public goods budget is low, or if the voter thinks it low, then the politician can be sanctioned for failing to provide public goods even when he or she is acting responsibly. This makes it more attractive for the incumbent to shirk, leaving the voter with nothing (including gifts). From the model, I derive observable implications about how voters will act when faced with new information about what they can expect of their governments.

Hypotheses

Shown graphically in Figure S1, increasing the village's belief p that the local government has a budget to provide a public goods project impacts the village's equilibrium behavior. First, marginal increases in p result in marginal increases in the maximum cut point at which the village can condition reelection and avoid shirking by the politician ($\overline{k}$). Second, if an increase in p crosses a certain threshold ($\frac{1}{\delta} - w$), then the village moves from always voting on private campaign gifts to voting on gifts *and* public goods in equilibrium.

While an increase in p may result from local changes such as an increase in transfers to the government budget or an increase in the competitiveness of an election, p can also be exogenously manipulated. I posit that improving information to villages about what good government performance looks like will result in an increase in p. (This is assuming voters have downwardly biased beliefs about potential government performance, which I show in the next section is true in Mali).

From this discussion, I derive the following testable predictions:

H0: Improving information about potential government performance will increase voter expectations of politician performance, p.

H1: Improving information about potential government performance will increase the cut point, $\overline{k}$, at which poor-performing candidates are sanctioned.

H2: Since voters will never condition reelection on public goods provision for small enough p, improving information about potential government performance will increase the likelihood of voting along the performance dimension.

General Equilibrium Effects

The comparative statics derived from the model imply that higher levels of information should lead voters to sanction poor-performing incumbents more often at election time. The change in voter behavior, however, is not an end in itself, but rather a means to an end: better government performance. Over time, politicians should respond to an increased real or perceived threat of sanctions by improving their performance. This is not a foregone conclusion but a question for further research.

How will politicians respond to changes in voter behavior? On the one hand, elected leaders might act more responsibly, knowing their actions are now being scrutinized more acutely by voters. On the other hand, leaders might make more of an attempt to hide their misbehavior from more watchful citizens.[6]

Institutional Context

Democratic since 1992 (except between the March 2012 military coup and presidential election of August 2013), Mali remains politically corrupt and economically underdeveloped. Freedom House (2014) cites corruption as a major problem in government, procurement, and contracting, and even a potential instigator of the 2012 Islamist takeover in the North. Malians are twice as poor and half as literate as those in the average sub-Saharan African country, with a literacy rate of 24% and gross national income per capita at US$500. This apparent failure of democracy to improve development outcomes cannot be attributed to weaknesses in formal democratic

[6]See, for example, Humphreys and Weinstein (2012). I similarly find that leaders appear less transparent where voters receive an information treatment (see SI H).

institutions themselves. At the time of this study, Malian elections had all been deemed free and fair, there were two peaceful transitions of power between political parties, and there was relatively high media freedom and freedom of association.

Political decentralization in Mali provides an opportunity to examine variation across independent and locally governed units. Each of the 703 communes democratically elects a multimember council with autonomous control over the local budget for public goods and services. Though commune governments can and do levy taxes, revenue is small relative to public expenditures that fall under their purview (e.g., clean water, primary schooling, and primary health care). To increase revenue, a para-statal agency transfers funds for annual development projects to each commune that make up more than half of the average government budget (locally generated funds comprise only 15%).

In spite of local government's mandate to provide public goods and its capacity (if limited) to do so, Malian voters appear to have downwardly biased beliefs about government performance. The household survey conducted in conjunction with the experiment reveals that more than a third of respondents do not know the local government has a budget to invest in the commune. About half of the respondents are under the false impression that the local budget is insufficient to finance even small development projects. Mali's nationally-representative Afrobarometer survey (2008) demonstrates widespread confusion about what falls within the local government's jurisdiction: Only a third of respondents say the local government is responsible for managing schools and health clinics, whereas half name the central government.

Consistent with the model's prediction about the effect of low expectations of politician performance, respondents in my survey often prioritize dimensions other than government performance in electoral decision making. While 64% of respondents say they prioritize performance themselves when voting, 48% say the receipt of gifts from candidates is the primary criteria *other* people use. Given the potential stigma attached to admitting the importance of vote buying, the latter figure is likely a better approximation of the truth. When asked why people would not vote on performance, about half of the respondents cite lack of information, a quarter blame lack of government capacity, and another quarter say people vote based on personal needs and interests.

Mali serves as a relatively easy test of the hypotheses in that voters are particularly underinformed and have demonstrated low expectations of local government. As a result, the generalizability of the findings is limited to places that, like Mali, exhibit low levels of information about potential government performance. This might include countries where only a narrow fraction of the population was involved in the democratic transition, where civic education is scant and political information hard to access, or where voters are uncertain about the division of responsibilities between levels of government due to recent decentralization. Moehler's work (2008) demonstrates that the curse of low expectations can also vary *within* a country by one's level of information and exposure to politics. Ugandan citizens become more distrusting of government (and can be assumed to have higher expectations) the more they learn about and participate in the political system.

Research Design

A field experiment in 95 rural communes in Mali tests whether and how a randomly assigned information intervention that improves voter expectations affects electoral behavior. In the absence of randomization, the treatment effect would likely be biased upward. Voters may have higher expectations of government performance precisely because their government is better behaved. Or some third factor such as modernization or political competition may produce both informed citizens and well-behaved governments. Because treatment is conditional on random assignment, these endogeneity concerns are alleviated.

Treatments: "Capacity" and "Capacity + Performance"

Two information treatments in the form of a civics course were conducted in 32 communes each in early 2011. Communes, the unit of randomization, are subdivided into villages (18 on average) with about 1,000 inhabitants each. Rather than at the commune level, the civics course treatment is conducted at the village level to allow for participation by a more representative sample of commune residents. Provided at one-week intervals, the course consists of two or three half-day sessions conducted in the local language by a trained Malian instructor.

The curriculum was developed in collaboration with the Malian government, local NGOs, and a national civic education program. Course sessions struck a balance between audiovisual materials to maximize homogeneity and interactive exercises to maximize comprehension (details in SI B). Much of the information provided in the course was likely new to participants. There is no

systematic civic education in Mali.[7] The government occasionally supports civic education campaigns around elections, but these are brief and under-funded. And while students receive some information about government in high school, fewer than 3% of respondents in the survey report having finished secondary school.

Communes are randomly assigned to receive one of two treatment arms. The Capacity treatment (T1) provides information about local government capacity and responsibility, whereas the Capacity + Performance treatment (T2) provides the same information as T1 with an additional component on relative local government performance.

- *Capacity component*: Both T1 and T2 villages receive two course sessions on local government capacity as a way of improving voter reference points through increased knowledge of *performance standards*. Participants learn about the responsibility of local governments to provide a menu of local public goods, such as clean water, primary health care, and primary schooling. They learn about the size of the local government budget with concrete examples of public services that can be realized with that amount of funding. Basic information about democracy and decentralization in the Malian context is also provided. While voters are not instructed on how not to vote (e.g., to avoid being swayed by kinship ties or gifts), voting based on performance is critically emphasized.
- *Performance component*: T2 villages receive a third course session on actual government performance that aims to additionally improve voter reference points through increased knowledge of *relative performance*. This session provides information specific to each village about how the local commune government performs relative to neighboring communes. Performance indicators include the number of locally funded development projects in the village, the distribution of projects between the commune seat and outlying villages, and the number of public meetings held by the commune council. To allow for benchmarking, each performance indicator is presented as an index comparing the local commune government to other communes in the sample. This has the advantage of giving participants information about the incumbent's performance record while at the same time improving knowledge about what is in the feasible set of government actions.

[7] Civic education was officially eliminated during the Traoré regime in 1972 and reintroduced in 2009.

If voter behavior is constrained by uncertainty about potential government performance as well as uncertainty about actual government performance, then the capacity component is not sufficient to produce changes in voter behavior. Even if voters learn about the standards they can hold governments to, they cannot properly evaluate politicians without knowing the incumbent's performance record. This is the motivation for providing participants in the Capacity + Performance treatment information about how their government performed.

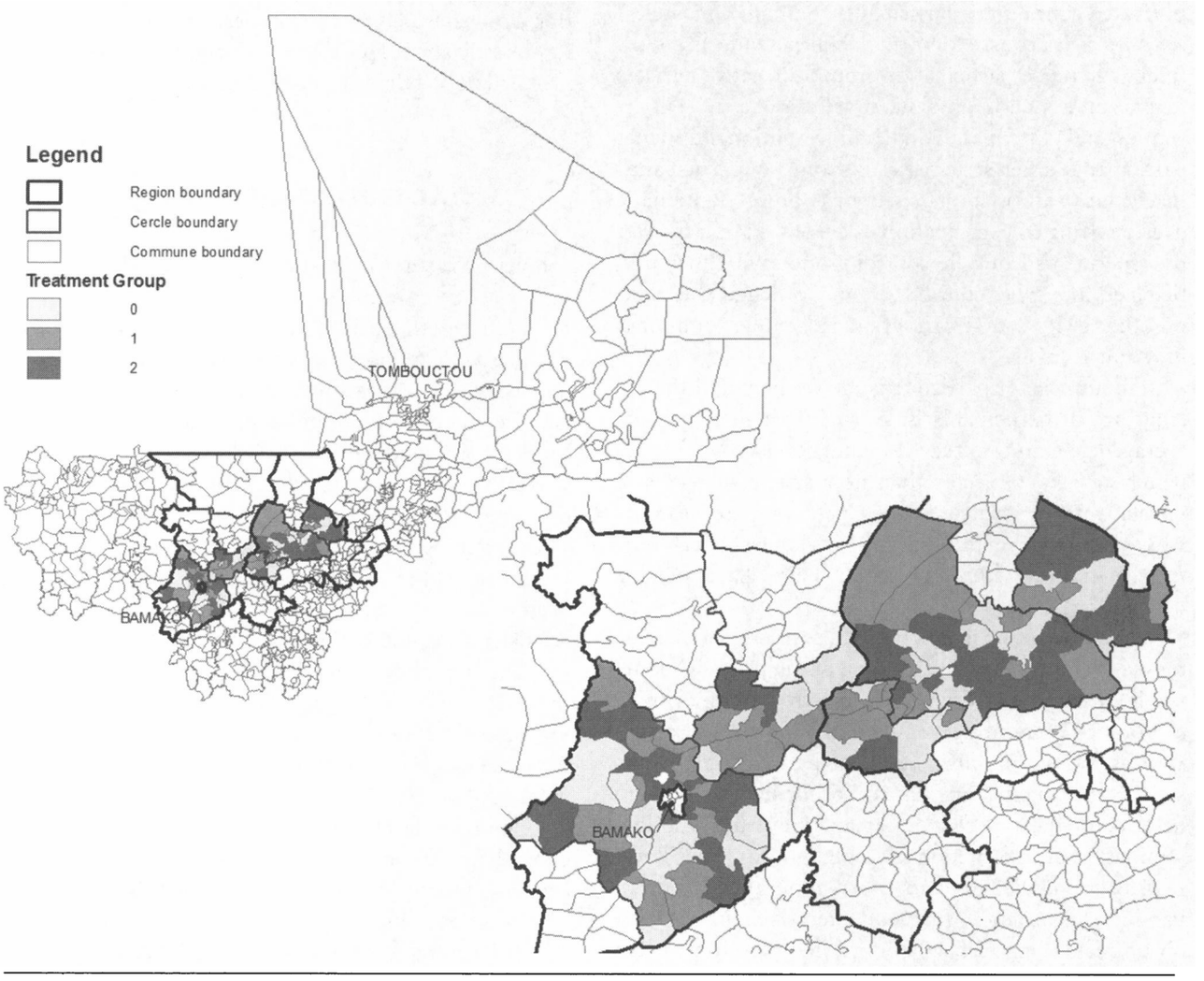
There are thus two key differences between T1 and T2 with respect to intended effects on voting behavior. First, there are differences in how each treatment informs voter reference points: T1 provides information about local government capacity as a standard against which to evaluate their own government, and T2 provides another, perhaps more salient, benchmark—the performance of candidates in neighboring districts. The second difference is how well voters can perform evaluations of their local governments. While voters in T1 must rely on existing knowledge of government performance, voters in T2 receive this information, making more accurate evaluations possible.

This design can answer the following questions: Is the capacity component sufficient to change political behavior? Are the capacity and performance components sufficient to change political behavior? Does the combined treatment produce effects that are significantly different from the Capacity treatment on its own? Because no group received a three-session treatment with only basic civics information (and no relative performance information), we cannot rigorously distinguish the effect of relative performance information from the effect of a longer treatment.

Sample and Timeline

The experimental sample consists of the 95 rural communes in five *cercles* or districts: Kati, Koulikoro, Segou, Macina and Baraoueli (see Figure 1). The sample was chosen to coincide with the intervention zone of the German aid agency (GIZ), which finances democracy programs such as annual town hall meetings. These meetings are a critical part of the measurement strategy, used to evaluate treatment effects on how voters interact with elected leaders. The sample communes have the same average population size and wealth level as the average Malian commune and are not differently politically competitive. However,

FIGURE 1 Map of Experimental Sample



located along the Niger River, these sample communes have more villages and are less remote than the national average.

Each commune in the sample is randomly assigned to one of three groups: the Control group (T0), the Capacity treatment condition (T1), or the Capacity + Performance treatment condition (T2). The control group receives no intervention and is visited for the first time during the survey. The intervention in the 370 treated villages took about two months to complete. There is a three-month lag between treatment and the household, chief, and politician surveys.

Using a block randomization design, the sample is stratified on three variables related to information and accountability: region, mayor incumbency, and a commune-level development index. By chance, 32 communes were assigned to the two treatment conditions and 31 to control. Budget and time constraints precluded studying every village in the sample communes. Sufficient power was attained by selecting six villages at random[8] from each commune in the sample, for a total of 556.

Treatment and control communes are neither significantly different on key indicators (see balance test in SI D) nor clustered geographically (see Figure 1). This latter observation makes spillovers from treatment into control communes more feasible. However, vast terrain and low population density combined with poor road and communication infrastructure limit this concern. Furthermore, information spillovers would bias the estimated treatment effect downward.

[8]The commune seat is always included because it would be politically infeasible not to do so.

Compliance

In treated communes, participants voluntarily self-select into the civics course following a village-wide assembly. Inferring course participation from affirmative answers to two survey questions about awareness of a civics intervention and participation in that intervention, I find that participants differ systematically from nonparticipants. Likely due to affirmation bias, there is, however, dramatic overreporting of participation as evidenced by a quarter of respondents from the control group saying they participated in a civics course. To my knowledge, there was no other civics course carried out in sample communes in recent years.

About 43% of participants are women, and the average age of participants is 45, slightly higher than the average age among survey respondents. Participants are better socially connected than nonparticipants: 48% of participants are related to the chief, compared to only 34% of nonparticipants. Participants are less likely to be from a minority ethnic group in their village and are better educated.

Due to selection issues, I was more interested in treatment effects on the village than on participants. The survey is thus conducted among a mostly random sample of households regardless of participation in the course. With an average of only 30 villagers participating, it may seem surprising that any impacts of the intervention are detected by the survey. One explanation is the fluid dissemination of information through tight-knit social networks within villages. In an information intervention in Pakistan, for instance, treatment effects were just as large on untreated female neighbors as they were on treated women (Gine and Mansuri 2011). The Malian *concession,* or enclosed grouping of households of the same extended family, facilitates the spread of information, as women of a concession jointly prepare food and men farm the same plots of land.

To measure within-village spillover of information from participant to nonparticipant households, I conducted a follow-up survey in a random sample of treated villages. On average, 18 unique concessions per village sent at least one member to participate in the course. An average of 193 adults in treated villages thus share a concession with at least one course participant. Considering that about half of the village residents are under 18, an average of two-fifths of adults in treated villages either attended the civics course or share a concession with an adult who did. Further, about half of the participants said they spoke about the course to other members of the village, and about a third of nonparticipants said they learned something about the course from people in their community. These data suggest it is feasible that information traveled widely within villages—a promising finding for the efficiency of providing information to only a small proportion of a tight-knit community.

Measurement Strategy

A household survey measures levels of civic knowledge, beliefs about government capacity, and respondent behavior in hypothetical voting simulations. While behavior in these simulations will not perfectly predict actual voting behavior, the measures are designed to minimize bias and the ability of the respondent to game the questions. The survey is conducted with 10 different households, six of which are randomly selected. Stratifying on gender, individuals within these households are randomly selected. The remaining four surveys are conducted with targeted local leaders: the women's leader, the youth leader, the head of the village association, and the village assistant elected during the civics course. These leaders are targeted because they are more likely to have participated in the civics course and to permit evaluation of differential impacts of the course across these household types.

Analysis of survey outcomes treats the individual as the unit of observation. Because random assignment is at the commune level, I use a mixed model to account for dependence among individual observations within villages and among village observations within communes. The model is fit using restricted maximum likelihood.[9]

Accounting for block randomization, average treatment effects on individual-level outcomes y_{ivc} for individual i in village v in commune c are estimated using the following equation:

$$y_{ivc} = \beta_0 + \beta_1 T1_c + \beta_2 T2_c + W_c'\Gamma + Z_e'\Lambda + \alpha_c + \gamma_{vc} + \varepsilon_{ivc}, \quad (1)$$

where $T1$ and $T2$ are indicators of whether the commune received Treatment 1 or 2; W_c is a fixed effect for block, the unit on which randomization was stratified; Z_e are enumerator fixed effects; α_c are random effects for commune; γ_{vc} are random effects for village; ε_{ivc} is an error term; and β_1 and β_2 are the parameters of interest, average treatment effects for T1 and T2.

[9] I use the xtmixed command in Stata.

TABLE 1 Treatment Effect on Expectations Index

	(1)	(2)
T1: Capacity	0.034*	0.035*
	(0.014)	(0.014)
T2: Capacity + Performance	0.022	
	(0.014)	
T2 High		0.040*
		(0.018)
T2 Low		0.006
		(0.017)
Intercept	0.044	0.043
	(0.043)	(0.044)
N	5545	5545

Notes: Standard errors in parentheses. *p < .05, **p < .01.

Data Analysis

To evaluate the first stage of the argument, I assess whether treatment has an effect on voter expectations of government performance. Following the data analysis plan (see SI C), I create a summary index of relevant survey items to estimate average treatment effects on voter expectations. Following Kling, Liebman, and Katz (2007), equally weighted component outcomes are demeaned by subtracting the control group mean and then converted into effect sizes by dividing by the control group standard deviation. Component questions include whether the local government is responsible for providing a range of public goods, beliefs about the size of the local budget, the number of future projects expected to be implemented by the council, and beliefs about how democracy works.[10]

Table 1 reports average treatment effects on the expectations index. Model 1 is the basic specification. Because T2 (Capacity + Performance) provided information to participants about relative government performance, it is effectively a heterogeneous treatment. Participants receiving good news about the local government's performance will be differentially affected compared to participants who receive bad news. In Model 2, I examine differential responses to treatment in the data by constructing heterogeneous treatment indicators. *T2 High* is an indicator that takes a value of 1 for local governments performing above the mean, and *T2 Low* indicates communes where governments performed below the mean.[11]

The Capacity treatment is responsible for about a one-tenth standard deviation increase in the expectations index. Whereas the effect of the Capacity + Performance treatment (T2) is not significant in the aggregate, the coefficient on T2 High reaches conventional levels of significance and the coefficient on T2 Low is estimated to be zero. These results accord with how we would expect individuals to update their priors. When participants learn their governments perform well relative to others, expectations of local government increase. What is perhaps surprising is that information about government capacity and responsibility results in the same type of updating as information about actual government performance.

Voting Simulations

I use evidence from two novel voting simulations conducted during the survey to evaluate hypotheses about the effect of information on voter behavior. One advantage of the simulations is the ability to manipulate specific candidate characteristics, isolating effects on behavior. Candidates in actual elections are different across multiple dimensions, making it difficult to identify the effect of any single characteristic. A second advantage is minimizing survey bias. Voting simulations are constructed so that the respondent cannot easily game the question or infer what is socially desirable. Survey items about vote choice may be subject to social desirability bias because respondents in the treatment group have adopted new ideas about normatively "correct" behavior. Treatment effects would thus reflect changes in norms about what is socially desirable rather than changes in past or future behavior.

One example of a treatment effect on norms is the self-reporting of campaign gifts. The survey asks whether the respondent has ever received a gift from a candidate during an election. About 57% of individuals in the control group say "never," compared to significantly more in treated groups: 61% in T1 and 65% in T2. Randomization, and the fact that all previous elections occurred prior to treatment, precludes significant differences in actual gift giving between treated and control communes. The difference in self-reporting implies treatment strengthened

[10] When I exclude questions about how democracy works (specified in the preanalysis plan), the magnitude of the effect is about twice as large.

[11] Performance measures are generated during the civics course, so they are only available for T2 communes. Because performance levels are not randomly assigned, endogeneity is a concern. Results hold with commune-, village-, and individual-level controls to reduce potential bias.

the norm against vote buying, causing people to be less likely to admit to receiving gifts from candidates.

Hypothesis 1: Improving Information Will Increase the Cut Point at Which Candidates Are Sanctioned. In the model, the cut point k at which voters sanction incumbents is increasing in p, or the likelihood there is a budget for public goods, in equilibrium. If treatment increases voter knowledge with respect to local government capacity, essentially an exogenous shock to p, we should expect voters to sanction at higher cut points in the treated groups. A voting simulation gives respondents the choice between a high-performing candidate and a low-performing candidate in a hypothetical election (see SI F for simulation wording). If the respondent initially chooses the high-performing candidate, the low-performing one offers a monetary gift in exchange for a vote. To minimize social undesirability of accepting candidate gifts, the respondent is offered an increasing scale of prices at which she can sell her vote. Because the respondent is not forewarned of the elements or limits of the scale, intuiting the most socially desirable response is difficult.

To evaluate Hypothesis 1, I compare the average price at which respondents are willing to sell their vote across treated and control groups. Choosing the high-performing candidate at higher offers against the low-performing candidate is equivalent to sanctioning the low-performing candidate more often, or under a wider range of parameters. Thus, I generate support for the hypothesis if treatment causes the voters to choose the high-performing candidate at higher offers against the low-performing candidate.[12]

Seventy-four percent of respondents initially choose the high-performing candidate. Subsequently, these respondents are asked whether they would switch their vote to the other (low-performing) candidate if he offered a gift of about US$1. If the respondent refuses, there are three consecutive amounts offered up to US$20. On average, 22% of these respondents agree to switch their vote for some amount of money. Regressing an indicator of whether the respondent ever switches her vote on treatment indicators, Table 2 shows that, conditional on voting for Candidate A, slightly but insignificantly fewer people ever switch their vote in the treatment groups compared to control.

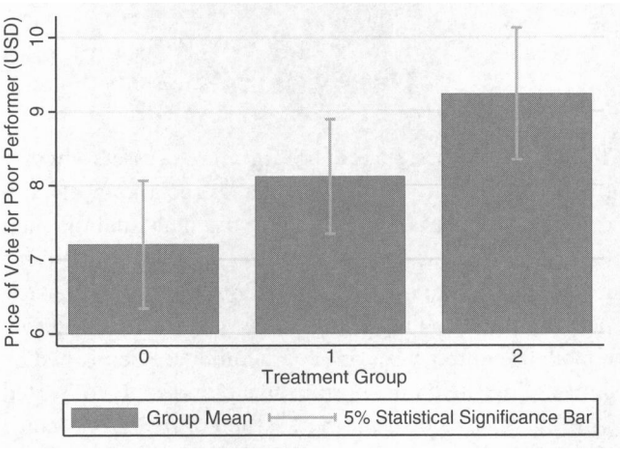
[12]The order of the two voting simulations was randomly assigned to the beginning or end of the survey. Evidence of priming, it is significantly more expensive to buy votes when respondents receive the voting simulation at the end of the survey. I thus analyze the half of the respondents who receive this voting simulation first.

TABLE 2 Treatment Effect on Switching One's Vote

	(1)
T1: Capacity	−0.011
	(0.023)
T2: Capacity + Performance	−0.034
	(0.024)
Intercept	0.214**
	(0.079)
N	2005

Notes: Standard errors in parentheses. *p < .05, **p < .01.

FIGURE 2 Votes More Expensive to Buy in Treated Groups



Note: Means and standard deviations calculated using commune-level averages. Results hold with the mixed model using individual-level data.

The following analysis of the treatment effect on vote price is performed only among respondents who say they are willing to sell their vote at some price. Because there is no significant effect on whether a respondent agrees to sell his vote, this should be an unbiased test of the price at which a vote is sold, conditional on being willing to sell it. However, among certain subgroups of the population, this is not the case. See SI G for tests of heterogeneous effects that impute values for respondents who were unwilling to sell their vote at any price offered.

For each treatment group, Figure 2 shows the average price the low-performing candidate would have to pay to win the respondent's vote.[13] There is evidence of a significant positive effect of the Capacity + Performance

[13]Means and confidence intervals are calculated using the commune, or level of randomization, as the unit of analysis. Results hold if I instead conduct the individual-level analysis in Equation 1.

treatment (T2) on price. The effect of the Capacity treatment (T1) is also positive but not significant at conventional levels. These findings are not substantively different when high prices are imputed for respondents who never sell their vote (see SI G), allowing the test to be performed on a much larger sample.

While the low-performing candidate would have to pay almost a dollar more in T2 compared to T1, the coefficients are not statistically different. T2 increases the price by about half of a standard deviation, or US$2. For the more than half of rural Malians living on less than a dollar a day, this is not an insignificant sum.

Hypothesis 2: Improving Information Will Increase the Likelihood of Voting along the Performance Dimension. The comparative statics of the model predict that low values of p, or a low enough likelihood that there are sufficient funds in the budget for a public good project in their village, will make it less feasible for voters to condition electoral decisions on public goods provision. The information intervention should thus increase the likelihood that a voter conditions her vote on the provision of public goods rather than relative to private transfers.

A second simulated election between new Candidates A and B examines whether the respondents in the treated group are more likely to vote based on performance criteria relative to other nonperformance dimensions. The two nonperformance dimensions salient in this exercise are kinship and support of the village chief. Shared lineage between the voter and politician as well as a village chief who serves as a local intermediary for a particular candidate both signal the possibility of greater private transfers.

Candidate performance records also differ. Candidate A built a well in one village during his tenure as mayor—a concrete but relatively meager public goods record—whereas Candidate B held annual public budget debates during his tenure as mayor. These descriptions are purposefully ambiguous on the performance dimension and were pretested to ensure substantial variation in preferences. The survey respondents are given a binary choice between Candidates A and B to reveal their preference for a particular type of performance.

To minimize the effect of social desirability bias, a survey experiment randomly assigns respondents to receive one of three versions of the candidate descriptions: Baseline, Kin, and Chief. The only difference between versions is that in the Kin condition, Candidate A is given the same last name as the respondent to signal kinship. In the Chief condition, Candidate A is given the support of the village chief. The Baseline condition provides only performance information. The Kin and Chief cues are designed to be subtle, and since each respondent receives only one version, he or she is unlikely to be aware of the experimental manipulation.

TABLE 3 Effect of Treatment on Voting Criteria

A. First Differences: Mean Vote Share for Candidate A

	Control	T1	T2
Difference (Kin–Baseline)	0.120**	0.077*	−0.015
	(0.025)	(0.033)	(0.028)
Difference (Chief–Baseline)	0.080*	0.021	−0.033
	(0.030)	(0.027)	(0.035)

B. Difference-in-Difference

	T1–C	T2–C	T2–T1
Kin condition	−0.045	−0.135**	−0.092*
	(0.045)	(0.037)	(0.045)
Chief condition	−0.062	−0.116*	−0.054*
	(0.040)	(0.047)	(0.039)

Notes: Standard errors in parentheses. *p < .05, **p < .01.

Table 3A analyzes the vote share for Candidate A (the candidate who was manipulated) across groups. In the control group, Candidate A receives a significantly higher share of the votes in the Kin and Chief conditions compared to the Baseline condition. Within treatment, the only significant difference is between the Kin and Baseline conditions for T1.

Because treatment may have affected voter preferences over candidates in the Baseline condition, the relevant test is whether the difference in vote share for Candidate A between conditions (Kin/Chief versus Baseline) is different in the treatment groups compared to the control. If treatment successfully decreases the salience of the kin and chief dimensions relative to the performance dimension, then the difference in vote share for Candidate A across versions of the survey should be smaller in the treatment group than in the control. A difference-in-difference estimation analyzes treatment effects on whether the change in Candidate A's vote share across survey conditions varies significantly with treatment. Hypothesis 2 is confirmed if the difference in vote share between the Baseline and other conditions is smaller for the treatment groups than for the control.

Table 3B reveals a negative treatment effect on the likelihood of respondents to vote along the kinship or chief support dimensions.[14] Voters in the control group

[14]For simplicity, means and standard errors are calculated using the commune, the level of randomization, as the unit of analysis. Results are even stronger if I instead conduct the individual-level analysis in Equation (1).

are more likely to vote for Candidate A when he is either a kin or has the chief's backing compared to voters in the treatment groups. The difference-in-difference estimators are significant for T2 but not for T1. In addition, the difference between the effect of T2 and T1 is significant in this case. In sum, treatment caused respondents to rely less on kinship or chief support dimensions in their voting calculus, increasing the salience of the performance dimension relative to the other two.

As a robustness check, I examine whether the kin and chief dimensions are more salient for respondents who stand to benefit from membership in a local social network. Aligning with the chief's preference would only be a good strategy for someone who believes his or her well-being is linked with the chief's. As expected, the kin and chief dimensions are more salient among respondents who are well connected or say they are relatives of the chief. Among chief relatives, Candidate A receives 22% more of the votes under the Kin condition than the Baseline condition compared to only a 7% bump among nonrelatives. Re-running the difference-in-difference analysis among the subset of chief relatives (39% of the respondents), treatment effects are larger and more significant. The difference-in-difference estimators are now significant for T1. This analysis indicates that treatment has a greater effect on increasing voting along the performance dimension when nonperformance dimensions are more salient.

As another robustness check, I examine whether respondent behavior is consistent across the two voting simulations. One would expect that a respondent who values the performance dimension more should be both harder to buy off in the first voting simulation and less tempted by the kin or chief dimensions in the second. There is indeed a strong negative correlation between the price at which a respondent can be bought off in the first simulation (imputing high values for respondents who said they would not sell their vote at any price) and the likelihood of voting for the candidate described as kin or having chief support.

External Validity of Survey Outcomes

In this section, I assess whether changes in behavior in survey experiments are reflected in changes in real-world behavior—namely, on petitioning leaders in a public forum. In recent years, Mali's Ministry of Territorial Administration along with GIZ have been encouraging commune governments to hold annual town hall meetings to provide a financial and administrative account of the previous year to citizens from constituent villages. This is one of the only formal opportunities for villagers to exchange with their commune leadership.

In communes that held town hall meetings after the implementation of the civics course, I analyze treatment effects on the extent to which meeting participants challenge their elected leadership. The expectation is that treatment should increase the extent to which villagers petition leaders, especially given poor citizen participation in past meetings. GIZ attributes this to lack of information and fear of speaking in public, both of which should have improved with treatment.

Because the subsample of communes that held a town hall meeting is not random, we have to rule out the possibility of selection bias. GIZ originally budgeted for the entire experimental sample to receive a town hall meeting between March and June 2011 (shortly after the completion of the civics course). Due to budget constraints, only one-third of communes held meetings. The criteria used in selecting these communes included the recency of previous town hall meetings, organizational capacity in the geographic area, and relationships with higher-level authorities. While some of these criteria might influence our dependent variable of civic participation at the meetings, all are orthogonal to whether or not a commune was treated. A balance test (see SI D) on pretreatment characteristics reveals no significant difference in the communes that received and did not receive a town hall meeting. In the subsample of communes holding a town hall meeting, 8 are in the control group, 13 are in T1, and 12 are in T2.

For the dependent variable of interest, number of challenges to leadership, I code as 1 each time a participant at the meeting intervened with a complaint or challenge to the commune government. Trained observers who were blind to treatment status attended each meeting and recorded details of all participant interventions. Regressing the dependent variable on treatment indicators, Table 4 reports the average treatment effect on the number of participant interventions. Because of the small number of observations, exact p-values are also calculated using randomization inference and compared to the asymptotic approximations.[15]

The effect is positive in every case, with treatment increasing the number of challenges on average by almost half in both T1 and T2. The effect is only significant at the 10% level when comparing the pooled treated groups with the control group, which is not surprising given the very small sample size. While only about 5% of all meaningful interventions (i.e., questions, comments, challenges)

[15]While randomization inference tests the null hypothesis that none of the units are affected by treatment, the standard null hypothesis is that the means across treatment conditions are the same.

TABLE 4 Average Treatment Effect on Challenges to Leadership during Town Hall Meeting

Mean Number of Challenges by Group			
	Control	T1	T2
Challenges	4.125	5.93	5.83
N	8	13	12
Mean Differences			
	Difference	p-Value (Two-Sided)	Exact p†
T1–C	1.80	.125	.148
T2–C	1.71	.109	.178
T–C	1.76	.086	.103

Note: †Exact p-values are calculated using randomization inference.

came from female participants, of the three women in the sample who openly challenged their leadership during meetings, all originated from treated villages.

Discussion

Taken together, these analyses provide evidence of a positive effect of improving voter reference points on electoral sanctioning of poor performers. A randomly assigned civics course successfully increased voter expectations of local government performance, as measured by survey questions about local government capacity and responsibility to provide public goods. In one survey experiment, treatment increases the average price a poor-performing politician would have to pay for a vote. This result confirms the hypothesis that raising voter expectations of politician performance increases the cut point at which poor-performing politicians are sanctioned. A second survey experiment confirms the hypothesis that raising voter expectations will increase the salience of the performance dimension: Voters are less likely to vote along two nonperformance dimensions (i.e., kinship and chief support) relative to a performance dimension. It is particularly striking that we see these effects of treatment due to (1) the brevity of the civics intervention, (2) relatively low power of the statistical analyses, and (3) the fact that outcomes are measured using a representative sample of villagers, many of whom did not participate in treatment.

Since the above findings are generated by survey measures, one concern may be the relevance of the intervention for actual behavior. Using observations of participants at town hall meetings, I demonstrate that findings from the survey experiments are consistent with treatment effects on actual behavior. People in treated communes are more likely to challenge their local leadership at these meetings.

Findings are mixed with respect to differences in impacts between the two treatment arms. Effects of the Capacity treatment (T1) and the Capacity + Performance treatment (T2) are in the same direction, but impacts of T2 are statistically significant, whereas those of T1 generally are not. The difference between the two treatment arms is not significant in the first survey experiment, but it is in the second. There are two competing interpretations of these findings. T2 may be working primarily through its effect on raising knowledge of local public goods provision, or through its effect on changing expectations (due to the fact that the performance information is relative). The mixed findings in the literature are inconsistent with the first interpretation, and I additionally find support in my own data for the second.

Replicating the experiment with a third treatment arm—the performance component minus the capacity component—would be an ideal way to discriminate between these claims. Using existing data, however, I test one observable implication. For the first interpretation to be true, T2 would necessarily need to impact knowledge of local government performance. Otherwise, T2 is more likely working through its effect on expectations. I compare post-treatment knowledge of local public goods across treatment and control groups, assuming randomization ensures equal distribution of public goods across these groups.[16] Using survey data on participant knowledge of which local public goods are funded by the local government in their village, and verifying these responses with data from the village chief survey, I count the number of verifiable local public goods each participant has knowledge of in his or her village. Using the specification in Equation (1), I find that neither T1 nor T2 has a substantial or significant impact on an individual's knowledge of existing local public goods in the village (see SI E). I interpret this as evidence that T2 is not working primarily through its impact on knowledge of local government performance, but rather through its effect on changing expectations achieved by providing relative performance information about local politicians.

Whether governments are engaged in corruption or fail to respond to citizen demands, a key barrier to political progress is the inability of citizens to make informed decisions at the ballot box. This project demonstrates the importance of a new type of information above and

[16]I must also assume there are no significant treatment effects on dishonest reporting between these groups.

beyond a politician's performance record. For voters to adequately sanction politicians, they require informed reference points against which they can evaluate their own politician. This information can come in the form of improved knowledge of performance standards or information about how comparable politicians are performing to allow for benchmarking.

While improvements in politician accountability and responsiveness are of ultimate interest, this study shows progress only on intermediate outcomes: changes in voter beliefs, behavior in a hypothetical election, and participation at a town hall meeting. Further research is necessary to examine whether such civic education programs can produce longer-term impacts on politician performance.[17]

References

Afrobarometer Data, Round 4, 2008–9, available at http://www.afrobarometer.org.

Banerjee, Abhijit, Donald Green, Jennifer Green, and Rohini Pande. 2010. "Can Voters Be Primed to Choose Better Legislators? Experimental Evidence from Rural India." Presented at the Political Economics Seminar, Stanford University.

Banerjee, Abhijit, Selvan Kumar, Rohini Pande, and Felix Su. 2010. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." Unpublished manuscript.

Barro, Robert J. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14(1): 19–42.

Beath, Andrew, Fotini Christia, and Ruben Enikopolov. 2013. "Do Elected Councils Improve Governance? Experimental Evidence on Local Institutions in Afghanistan." MIT Political Science Department Research Paper No. 2013–24.

Besley, Timothy. 2006. *Principled Agents? The Political Economy of Good Government.* New York: Oxford University Press.

Besley, Timothy, and Anne Case. 1995. "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition." *American Economic Review* 85(1): 25–45.

Besley, Timothy, and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117(4): 1415–51.

Boix, Carles. 2001. "Democracy, Development, and the Public Sector." *American Journal of Political Science* 45(1): 1–17.

[17]Evidence from leader surveys suggests that, at least in the short run, more informed voters may lead politicians to become less transparent (see SI H).

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics* 127(4): 1755–1812.

Chong, Alberto, Ana DeLa O, Dean Karlan, and Leonard Wantchekon. 2012. "Looking Beyond the Incumbent: The Effects of Exposing Corruption on Electoral Outcomes Center for Economic Policy Research DP8790.".

Collier, Paul, and Dominic Rohner. 2008. "Democracy, Development and Conflict." *Journal of the European Economic Association* 6(2–3): 531–40.

Dal Bó, Ernesto, and Robert Powell. 2009. "A Model of Spoils Politics." *American Journal of Political Science* 53(1): 207–22.

de Figueiredo, Miguel F. P., F. Daniel Hidalgo, and Yuri Kasahara. 2011. "When Do Voters Punish Corrupt Politicians? Experimental Evidence from Brazil." Unpublished manuscript.

Fearon, James. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation,* ed. Adam Przeworski, Susan C. Stokes, and Bernard Manin. Cambridge: Cambridge University Press, 55–97.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50(1): 5–25.

Ferraz, Claudio, and Frederico Finan. 2008a. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." IZA Discussion Paper No. 2843.

Ferraz, Claudio, and Frederico Finan. 2008b. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2): 703–45.

Freedom House. 2014. "Freedom in the World 2014: Mali." https://freedomhouse.org/report/freedom-world/2014/mali (accessed February 6, 2014).

Gine, Xavier, and Ghazala Mansuri. 2011. "Together We Will: Evidence from a Field Experiment on Female Voter Turnout in Pakistan." World Bank Policy Research Working Paper No. 5692.

Gyimah-Boadi, Emmanuel. 2004. *Democratic Reform in Africa: The Quality of Progress.* Boulder, CO: Lynne Rienner.

Humphreys, Macartan, and Jeremy Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." Presented at the Annual Bank Conferences on Development Economics, Washington, DC.

Inoguchi, Takashi, AsiaBarometer Survey Data 2003–2008. AsiaBarometer Project http://www.asiabarometer.org. AsiaBarometer is a registered trademark of Takashi Inoguchi, President of the University of Niigata Prefecture, Japan, and Director of the AsiaBarometer Project.

Kayser, Mark Andreas, and Michael Peress. 2012. "Benchmarking across Borders: Electoral Accountability and the Necessity of Comparison." *American Political Science Review* 106(3): 661–84.

Keefer, Philip. 2007. "Clientelism, Credibility, and the Policy Choices of Young Democracies." *American Journal of Political Science* 51(4): 804–21.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83–119.

The AmericasBarometer by the Latin American Public Opinion Project (LAPOP), www.LapopSurveys.org.

Lieberman, Evan S., Daniel N. Posner, and Lily L. Tsai. 2014. "Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya." *World Development* 60: 69–83.

Moehler, Devra C. 2008. *Distrusting Democrats: Outcomes of Participatory Constitution Making.* Ann Arbor: University of Michigan Press.

Montinola, Gabriella R., and Robert W. Jackman. 2002. "Sources of Corruption: A Cross-Country Study." *British Journal of Political Science* 32(1): 147–70.

Pande, Rohini. 2007. "Understanding Political Corruption in Low Income Countries." *Handbook of Development Economics* 4: 3155–84.

Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3(2–3): 259–67.

Ross, Michael. 2006. "Is Democracy Good for the Poor?" *American Journal of Political Science* 50(4): 860–74.

Shleifer, Andrei. 1985. "A Theory of Yardstick Competition." *RAND Journal of Economics* 16(3): 319–27.

Svolik, Milan W. 2013. "Learning to Love Democracy: Electoral Accountability and the Success of Democracy." *American Journal of Political Science* 57(3): 685–702.

World Bank. 2009. "From Inertia to Participation: The Case of RECURSO in Peru." http://blogs.worldbank.org/publicsphere/inertia-participation-case-recurso-peru.

WORLD VALUES SURVEY Wave 5 2005–2008 OFFICIAL AGGREGATE v.20140429. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: Asep/JDS, Madrid SPAIN.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

A. Modeling voter expectations
B. Course curriculum outline
C. Pre-analysis plan
D. Balance tests
E. Treatment effect on knowledge of public goods
F. Voting simulations
G. Heterogeneous treatment effects on vote price
H. Impacts on leader behavior