

# Horizontal But Not Vertical: Accountability Institutions and Electoral Sanctioning in Northeast Brazil

Taylor C. Boas<sup>1</sup>, F. Daniel Hidalgo<sup>2</sup>, and Marcus A. Melo<sup>3</sup>

<sup>1</sup>Boston University

<sup>2</sup>Massachusetts Institute of Technology

<sup>3</sup>Federal University of Pernambuco

January 2, 2018

Institutions of horizontal accountability often aim to sanction malfeasant or corrupt officeholders before they have an opportunity to seek reelection. In Italy, Silvio Berlusconi was expelled from the Senate and barred from public office under an anticorruption law after being convicted of tax fraud in 2013. In Peru's 2016 election, electoral authorities disqualified César Acuña as a candidate in the presidential race for vote-buying activities when he was mayor of Trujillo. Though decisions to bar candidates are sometimes controversial and seen as politically motivated—as in the case of Caracas mayor Leopoldo López, a leading opponent of Venezuelan president Hugo Chávez—they can often further the cause of good governance by preventing dishonest politicians from perpetuating their hold on power.

Yet horizontal accountability institutions do not always succeed in their efforts to block malfeasant officeholders from seeking reelection. In Brazil, the 2010 Clean Slate Law allowed candidates to be disqualified if a government auditing agency had charged them with corruption or financial irregularities during prior terms in office. However, a 2016 Supreme Court ruling held that, in the case of candidates for executive office, disqualification required that the charge be confirmed by the corresponding legislature—which, at the state and municipal levels, typically does the bidding of the governor or mayor.

When horizontal accountability is stymied by legal obstructions or candidates' political connections, citizens have the potential to step in and exercise vertical accountability, voting against incumbents who have been charged with corruption or malfeasance but managed to remain on the ballot. Electoral sanctioning requires, first and foremost, that voters be made aware of incumbents' transgressions while in office. Towards this end, auditing agencies often seek to disseminate their decisions as part of a broad public education mission. Yet it also requires that voters condemn malfeasance by elected officials, and that they be willing to act upon this norm when they go to the polls.

Our Metaketa project aimed to test whether horizontal accountability institutions could induce vertical accountability by informing citizens of significant wrongdoing, or lack thereof, by incumbent politicians running for reelection. Partnering with the State Accounts Court of Pernambuco, a government auditing agency in Brazil, our intervention told voters whether the Court's annual audit of municipal accounts had found substantial evidence of malfeasance attributable to the mayor. We examine the effect of this treatment on self-reported vote for the mayor, measured via a secret ballot question in a post-electoral wave of the panel study.

Our study found that informing voters of the approval or rejection of their mayor's accounts has no significant effect on the decision to vote for his or her reelection. This null effect also applies to evaluations of the mayor's performance and to levels of certainty regarding this evaluation. We argue that the divergence between norms and action explains these null effects. While Brazilians strongly condemn corruption in the abstract, their behavior in the real world is constrained by factors such as loyalty to local political dynasties and the greater salience of more pressing concerns like employment and health services.

## **1 The Politics of Horizontal Accountability in Pernambuco**

A key feature of our study is that we collaborated with one of Brazil's State Accounts Courts (Tribunais de Contas dos Estados, or TCEs), the main institutions of horizontal accountability charged with monitoring state and municipal governments' compliance with the law.<sup>1</sup> Brazilian TCEs are key actors in state politics and policy making because their decisions provide the

---

1. Most TCEs are charged with auditing both municipal governments and the state governments. In a few states, there is a separate Municipal Accounts Court that handles only the municipal audits.

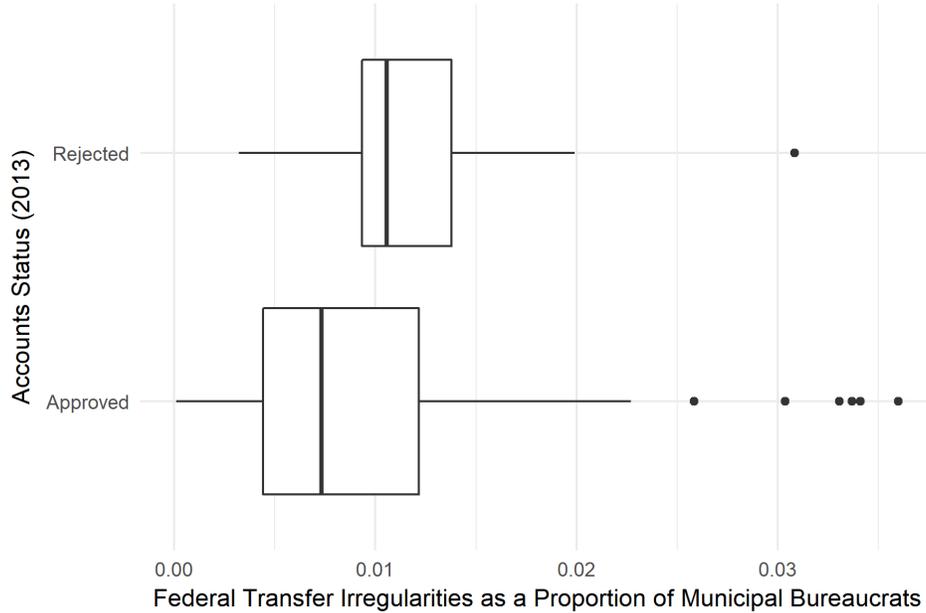
primary legal and political basis for sanctioning local and state governments. Relying on a large and highly trained staff, the courts engage in routine annual audits of all government bodies, as well as conducting *ad hoc* audits of specific programs and governments. The process culminates in an overall recommendation that audited accounts be "approved," "approved with reservations," or "rejected." In the case of executive branch audits, the recommendation is then sent to the corresponding legislature for a final decision. At the municipal level, the court's recommendation regarding a mayor's accounts can only be overturned by a two-thirds vote of the city council.

Municipalities and the mayors that lead them are central political actors in Brazil. Akin to a U.S. county, municipalities are largely responsible for the provision of basic social services such as primary education and health care, in addition to local services like garbage collection, housing, and water provision. Comparative analyses of inter-governmental relations classify Brazil as among the most decentralized polities in the world, with very high levels of fiscal and political decentralization, although taxation capacity at the municipal level remains generally low (Falleti 2010, p. 150). Directly elected mayors are the most important local political actors because they control the local municipal apparatus and also function as important intermediaries between citizens and the state and federal governments. Political competition is largely non-ideological, and parties tend to be weak (Novaes 2017), especially in smaller municipalities. While many voters may have persistent loyalties to political families or other groups within the municipality (discussed further below), swing voters tend to vote on more personalistic or clientelistic grounds.

We chose the state of Pernambuco as the location for our study largely because of the professionalism and efficiency of its TCE. As discussed below, Brazilian auditing agencies vary in the degree to which they are considered independent, professional organizations free from overt political meddling; the reputation of Pernambuco's court, the TCE-PE, is among the best (Melo, Pereira, and Figueiredo 2009). In addition to the likely effect of boosting citizen confidence in the court's judgments and their potential influence on voting behavior, the TCE-PE's professionalism meant that it was more open to a partnership with academics than a more politicized agency would have been. Brazilian TCEs also vary widely in their efficiency. Some routinely take five or more years to review municipal accounts, meaning that information on a mayor's first four-year term is not available until after he or she has stood for reelection. In Pernambuco, the TCE typically completes its review in three years or less, meaning that for the vast majority of mayors, a judgment of their first year's accounts is issued prior to the next election and is available to communicate to voters. By the time our intervention began, the TCE-PE had reviewed the accounts from 2013, the first year of the current mayoral term, in 95% of the state's municipalities.

Rejection of a municipality's accounts occurs when the TCE finds that the municipal government failed to comply with regulations and laws that govern local government expenditures, such as procurement legislation, constitutionally mandated spending, and hiring procedures. The court issues a report (*parecer prévio*) that describes any violations, recommendations for remediation, and recommended punishments. Not every violation is sufficient basis for the rejection of accounts, as the court has the option of recommending "Approval with Reservations" when improprieties are less serious. In the municipalities included in our study, the rejection of a mayor's accounts occurred for a variety of reasons. In the municipality of Flores, the court cited excessive personnel expenditures, municipal debt that far surpassed legal limits, and the failure to properly report details about government spending. In Santa Filomena, the court highlighted failure to spend required amounts on education and to transfer employees' pension contributions to the state pension fund, among other violations. In Bom Conselho, the court charged that the mayor had incurred substantial debt without approval from the local legislature,

among other infractions. Reports for other municipalities with rejected accounts described similar violations.



**Figure 1** – Number of Federal Transfer Irregularities by Accounts Status. These irregularities are reported by the federal government auditing agency CGU. Total number of irregularities has been normalized by number of bureaucrats employed in 2008.

Reassuringly, the decisions of the TCE-PE are correlated with other, independent measures of government irregularities. The most well-known auditing agency in Brazil, the Comptroller General of the Union (*Controladoria-Geral da União* or CGU) performs regular audits of federal transfers to municipalities—an area outside the purview of the TCE—and publicizes the names of public servants responsible for irregularities. Using these data, we computed the number of public servants named in CGU audits in all municipalities in the state, normalized by the total number of municipal bureaucrats employed in 2008.<sup>2</sup> To assess the extent to which the TCE-PE’s overall judgment correlates with the CGU’s audit findings, Figure 1 plots the distribution of CGU irregularities by the TCE-PE’s approval or rejection of the mayor’s 2013 accounts. On average, municipalities whose accounts had been rejected by TCE-PE had many more public servants found to have misspent federal funds. This finding suggests that the TCE-PE’s judgments do reflect broader differences in governance among Brazilian municipalities.

Decisions taken by Brazilian TCEs have potentially severe consequences for politicians, yet in practice these institutions are often quite hampered in their ability to exercise horizontal accountability. A 1990 law allowed politicians to be barred from running for office for 8 years if their accounts had been rejected, the legislature had upheld the decision (in the case of executive officeholders), and all possibilities for appeal had been exhausted. However, the long, draw-out appeals process meant that incumbents with rejected accounts were typically able to run again—and even finish a second term—before a final decision on their case could be rendered (Speck

2. Specifically, we used the list of public servants whose accounts were judged "irregular" by the CGU between 2008 and 2016. The number of municipal bureaucrats was obtained from IBGE’s “Perfil dos Municípios Brasileiros”.

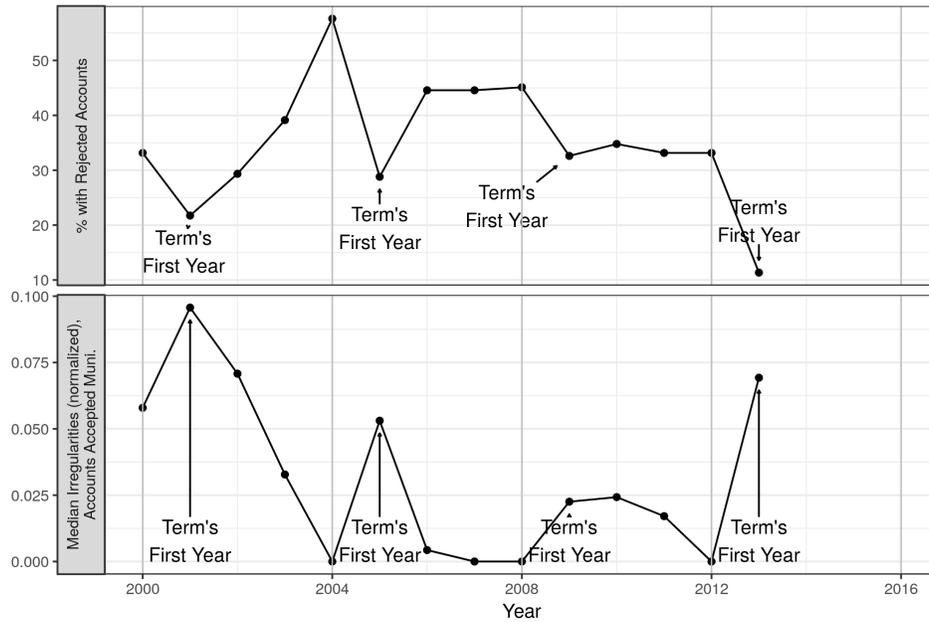
2011, p.145). In 2010, the new Clean Slate (*Ficha Limpa*) Law sought to close the judicial appeals loophole, allowing candidates to be barred based solely on the rejection of their accounts by the TCE. However, a Supreme Court decision in August 2016 significantly weakened the law by ruling that, in the case of executive branch accounts rejection, candidates could only be disqualified if the rejection had been upheld by the corresponding legislature. City councils often incur substantial delays in voting on the TCE-PE's recommendation, during which time the mayor may have run for reelection. Moreover, since city councils are usually dominated by the mayor's allies, they routinely overturn recommendations that accounts be rejected. From 1994 to 2013, 50% of rejection recommendations were overturned by the city council, versus only 5% of approval recommendations.

In addition to their efforts to exercise horizontal accountability, TCEs seek to induce vertical accountability by publicizing their auditing decisions and educating the public about their general mission. In Pernambuco, the TCE-PE employs a dedicated public outreach staff and publishes a column in the major newspaper that reports on its activities, including decisions on accounts. More generally, it seeks to directly inform the public about various aspects of municipal governance in Pernambuco. For example, its website "Tome Conta" (roughly translated as "Supervise" or "Take Notice") conveys indicators of government performance in a variety of areas, including health and education. The TCE-PE also has an outreach program, "TCEndo Cidadania" (a play on "weaving citizenship"), that involves holding public forums in municipalities around the state in order to educate citizens about local governance and help them hold elected officials accountable.

Although Brazil's TCEs project themselves as impartial arbiters and investigators, they are political institutions by design. TCEs are led by a panel of seven "councilors" (*conselheiros*), three of whom are appointed by the governor and four by the state legislature. The governor is fairly constrained in two of his three choices—one must be an Accounts Court career auditor, the other must be a career public prosecutor, and both have to be chosen from a list of three candidates compiled by either the Accounts Tribunal or the Public Prosecutor's Office. However, the governor's third choice, and all four of the legislature's choices, are essentially unrestricted, meaning that political criteria often factor into their decisions. Councilors selected by the legislature are typically former state deputies belonging to the dominant coalition and have a clearly political, rather than technical, profile. All councilors have protected tenure until a mandatory age of retirement.

The TCEs' institutional structure means that audit decisions are at least partly responsive to political factors, especially in the annual review of accounts, their most visible and legally consequential function. Many courts are dominated by councilors with partisan or family ties to politicians. In Pernambuco, the court had five members with political ties: three were former state deputies, one was the former cabinet chief of the governor, and another was a cousin of the governor who appointed him (Paiva and Sakai 2014). An emerging literature on horizontal accountability in Brazil has documented the implications of these political ties. TCEs tend to punish governments more readily when the councilors are politically diverse (Melo, Pereira, and Figueiredo 2009), whereas councilors tend to treat co-partisan politicians more leniently (Hidalgo, Canello, and Lima-de-Oliveira 2016). The degree to which political factors influence decision-making varies substantially, and Pernambuco's court is considered one of the most professional and least politicized (Melo, Pereira, and Figueiredo 2009). However, decisions by its councilors with political careers do show some evidence of favoritism toward co-partisan mayors (Hidalgo, Canello, and Lima-de-Oliveira 2016).

In addition to sometimes treating co-partisans more leniently, there is clear evidence that the bar for rejection of everyone's accounts is raised or lowered depending on the political sensitivity



**Figure 2** – Electoral Cycles in Accounts Court Decision-making. Top panel shows the proportion of municipalities with rejected accounts by year. Bottom panel shows the normalized median number of irregularities reported by federal auditors in municipalities with accepted accounts, by year.

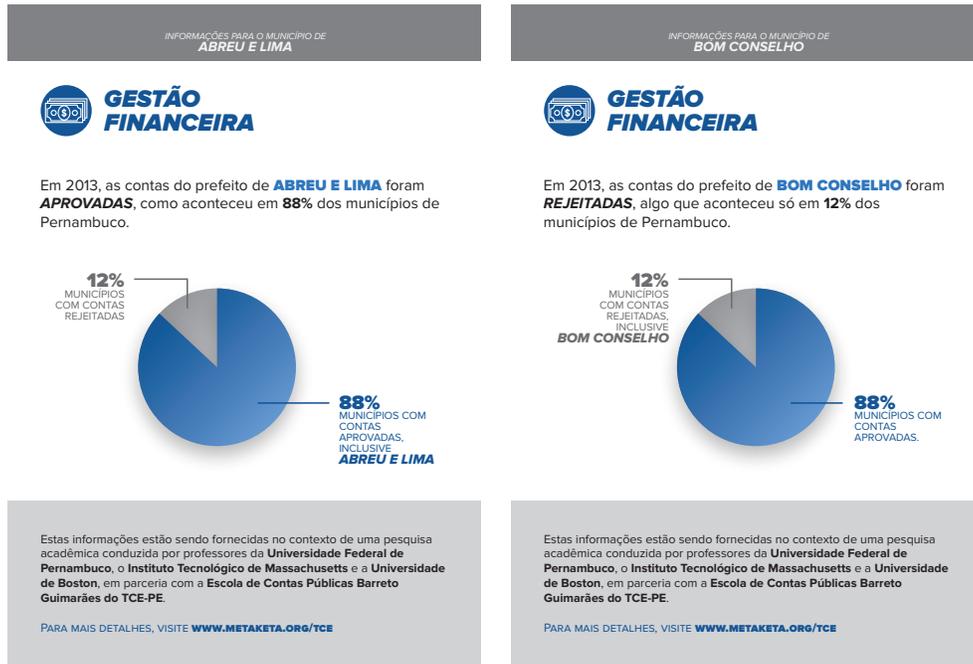
of the period in which they are being judged. The top panel of Figure 2 shows the percentage of municipalities with rejected accounts over time. The first year of each mayoral term—the only one for which accounts are likely to be adjudicated before the next election, and hence, the most politically consequential judgment—always has the lowest rejection rate. Indeed, our contacts within the court confirm that first-year accounts are judged more leniently because of election-year sensitivities.<sup>3</sup>

This electoral accountability cycle has direct implications for the types of municipalities that are approved or rejected in particular years. In the bottom panel of Figure 2, we plot the normalized number of CGU irregularities—as noted above, an independent measure of municipal malfeasance—in the median municipality with approved accounts. In electorally sensitive years, the number of irregularities found by federal auditors is unusually high in municipalities declared to be law-abiding by the TCE-PE.

The TCE-PE’s hesitation when rejecting accounts was likely exacerbated by the 2010 passage of the Clean Slate Law, which, as discussed above, made the court’s decision itself grounds for barring a candidacy. Starting in 2006—the accounts for which would have been judged in 2009, while the Clean Slate Law was being debated—we see a steady decline in rejection rates. Likely due to a combination of these two dynamics, the rejection rate in 2013 was the lowest since our data series began, at only 12% of municipalities in the state. This drop was accompanied by a large increase in the number of audited irregularities among municipalities with approved accounts.

The TCE-PE’s unusually lenient approach to judging mayors’ accounts in 2013 has several

3. An alternative explanation is that mayors are simply too inexperienced in their first year to engage in serious malfeasance. This hypothesis is belied by the fact that we see similar patterns among mayors in their second term.



(a) Example Accounts Accepted Flier

(b) Example Accounts Rejected Flier

**Figure 3** – Example of Fliers Distributed to Voters

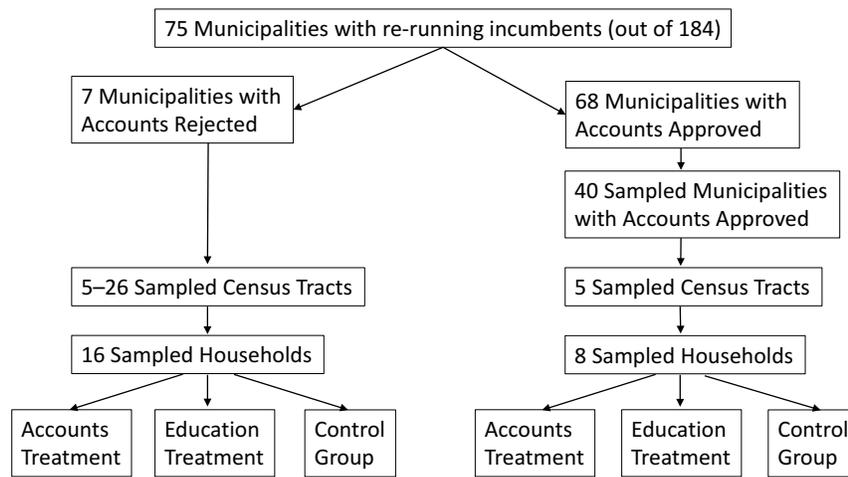
implications for our research design. At the time of the 2016 election, the 2013 accounts were the most recent ones that had been judged for nearly all municipalities in the state and the only ones corresponding to the first year of the incumbent mayor's term. The fact that the rejection rate in 2013 was unusually low, and that only a subset of mayors with rejected accounts ran for reelection, meant that we would have to sample every eligible municipality with rejected accounts, and a larger proportion of voters within these municipalities, in order for equal numbers of respondents to receive "good news" and "bad news" about their mayor's performance in office. Moreover, the set of municipalities with both approved and rejected accounts in 2013 was unusual compared to other years. As the federal auditing data indicates, municipal governments with approved accounts were more likely to be "bad types" than in previous years, possibly rendering court's decisions less informative about the overall quality of municipal governance. Meanwhile, the small set of municipalities that cleared the bar for rejection were likely to be especially egregious violators.

## 2 Experimental Design

### 2.1 Treatment

Our common arm intervention informed voters as to whether the mayor's accounts were approved or rejected by the TCE-PE in 2013, the first year of the current mayoral term. Information was delivered to voters in the form of a flier handed out by enumerators during the baseline wave of the survey; examples for each type of municipality are contained in Figures 3a and

3b. Enumerators also summarized the information orally to maximize information retention and facilitate comprehension among illiterate voters. The flier design was refined based on feedback from two rounds of focus groups conducted with voters from three municipalities as well as review by our government partner, the TCE-PE. The front of the flier bore the logos of the TCE-PE and its affiliated academic institution, the Public Accounts School, and it briefly explained the court's auditing responsibilities. The reverse side conveyed municipality-specific details, including a pie chart with comparative metrics.<sup>4</sup>



**Figure 4** – Overall Structure of the Sampling and Treatment Assignment Process

The implementation of the field experiment involved sampling municipalities, census tracts, and individuals, who were then individually randomized to three different treatment conditions. The overall structure of our research design is summarized in Figure 4. Below we describe in detail each stage of the process and their implications for interpretation of our findings. The timing of the pilot, baseline, and endline is shown in Figure 5.

## 2.2 Subjects and Contexts

### 2.2.1 Sampling Municipalities

The primary criterion for sampling municipalities was achieving a balanced sample, such that an equal number of respondents would receive positive and negative information about the

<sup>4</sup> We initially designed the fliers to mimic vivid advertising common in campaigns, but our focus groups indicated that recipients would likely believe that the fliers were distributed by politicians and not by the TCE. As a result, we adopted a more neutral and staid design.

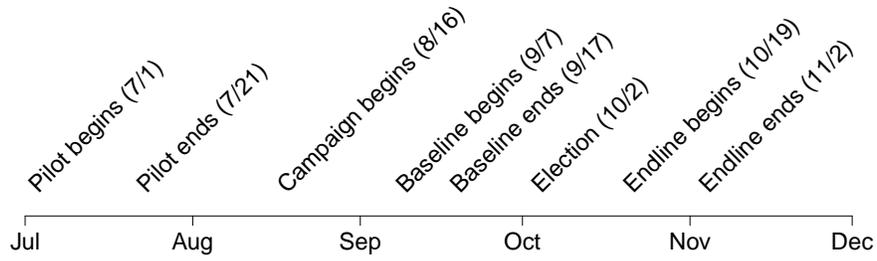


Figure 5 – Project Timeline

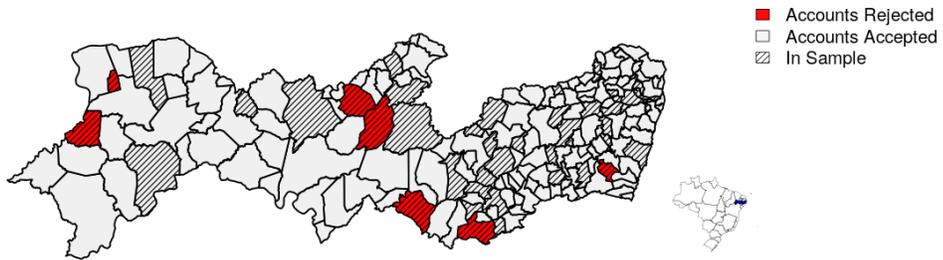


Figure 6 – Sampled Municipalities in Pernambuco, Brazil

incumbent government. Because there were only seven municipalities where incumbents with rejected accounts ran for reelection, we included all of them in the study. To construct the sampling frame of municipalities with approved accounts, we eliminated the smallest municipalities (where it might be difficult to sample the requisite number of voters) as well as the state capital, Recife. We then grouped the remaining 75 municipalities into two strata based on information to be conveyed in our alternative arm—performance on the National Literacy Exam (Avaliação Nacional da Alfabetização, ANA), a standardized test given in elementary schools—and sampled an equal number of municipalities from each stratum with inclusion probabilities proportional to the 2010 population. Our sample of municipalities is depicted in Figure 6, which shows the geographic distribution of sampled communities as well as accounts status. As shown by the map, the accounts rejected municipalities are geographically distributed across the state, though six out of the seven municipalities are located in the poor, semi-arid region known as the *sertão*, whose politics is often characterized as relatively traditional and clientelistic.

**Table 1** – Socio-Demographics and Political Characteristics of Sampled Municipalities. Source: 2010 Census and Superior Electoral Court.

Variable	Brazil	Pernambuco	Sampled Accounts Accepted	Sampled Accounts Rejected
<b>Socio-Demographic Variables</b>				
% in Extreme Poverty	6.6	12.3	15.1	22.9
% with Running Water	92.7	83.7	76.6	65.2
% Students Held Back	38.2	43.6	45.2	52.0
% Working in Agriculture	13.6	18.9	29.3	44.7
Average Income (BRL)	794	525	347	279
Human Development Index	0.73	0.67	0.62	0.57
<b>Electoral Variables</b>				
2012 Winner Vote %	54.4	55.6	53.0	52.2
2012 Turnout (% of registered)	83.5	83.7	83.3	78.2
<b>Baseline Survey Variables</b>				
Muni. Government Evaluation (5 point scale)	-	-	3.0	3.1
Confidence in Muni. Government (7 point scale)	-	-	4.2	4.1
Confidence in Accounts Tribunal (7 point scale)	-	-	4.3	4.5
Vote Buying Offer is Somewhat or Very Probable (%)	-	-	33.0	30.1
Vote Monitoring is Somewhat or Very Probable (%)	-	-	29.3	31.0

On basic socio-demographic variables, our sample is poorer and more rural than Brazil or Pernambuco as a whole. As indicated in Table 1, municipalities with both approved and rejected accounts have higher levels of extreme poverty, lower rates of access to running water, lower average monthly incomes, and worse educational performance. The poorer socioeconomic profile of our sample is not surprising, given the Northeast’s persistent under-development and our exclusion of Recife, the wealthiest city in the state. Within our sample, however, the accounts rejected municipalities are substantially poorer and more rural than their accounts approved counterparts. For example, the average monthly income in accounts rejected municipalities is less than half of the Brazilian average and 20% smaller than the accounts approved average. Similarly, the agriculture sector’s share of the workforce is about 13 percentage points higher in the accounts rejected municipalities.

With respect to electoral participation and competition, our sample is quite representative of Pernambuco and Brazil as a whole. Turnout and average vote share for the incumbent in the last election (2012) are broadly comparable to national and state averages. Elections are quite competitive, with incumbents only garnering around 52–53% of valid votes. Due to compulsory voting, turnout is high, with average rates of 78–83%. On these basic political indicators, accounts

approved and accounts rejected municipalities are broadly comparable.

Attitudes towards the government and perceptions of the electoral process were also quite similar across the two types of municipalities. In the bottom rows of Table 1, we present data from our baseline survey on evaluations of the municipal government and the TCE, as well as perceptions about the prevalence of vote buying and the secrecy of the ballot. Across all five variables, we see quite similar responses in both accounts rejected and accounts approved municipalities.

### 2.2.2 Sampling Respondents

Within each municipality, we used a two-stage sampling procedure that involved choosing census tracts and then respondents. Census tracts were sampled with probability proportional to the number of households in the 2010 census, excluding the least populous and most rural tracts where, based on pretesting, we anticipated logistical problems during fieldwork. In accounts approved municipalities we sampled five census tracts, while in accounts rejected municipalities we sampled between 5 and 26 census tracts, varying with municipality size. Within each tract, enumerators sampled sixteen households in accounts rejected municipalities and eight households in accounts approved municipalities. To ensure that interviews were geographically distributed throughout the tract—thus reducing the risk of spillover among neighbors assigned to different treatment conditions—we calculated a tract-specific number of houses to skip after a successful interview by dividing the total number of households in the tract by twice the number of interviews to be conducted. To avoid large imbalances in basic demographics, we used census data to construct sex-specific age quotas that interviewers were required to meet in each census tract.

## 2.3 Threats to Validity

### 2.3.1 Attrition

**Table 2** – Comparison of Attrited and Re-interviewed Respondents.

Variable	Attritted	Re-interviewed
Age	38.1	41.2
Male (%)	58%	48%
Income < R\$880.00 (%)	44%	51%
Evaluation of Incumbent (5 point scale)	3.0	3.1
2012 Vote for Incumbent (%)	55%	52%
2012 Turnout (%)	84%	85%
Accounts Treatment	36%	33%

Because we measure our outcome using a post-election survey, it was important to minimize attrition, both to preserve statistical power and to reduce the possibility of post-treatment bias. Our survey enumerators returned multiple times to interviewed households and often tracked missing respondents to their workplace or other locations to complete the endline interview. Recontact was more difficult, and attrition was noticeably higher, in more urban municipalities where respondents tended to work further from home. Overall, we achieved a re-contact rate of 81%.

Figure 2 shows how basic demographic and political variables vary by attrition status. As expected, men and higher income individuals were more difficult to reinterview as they were more likely to be employed and away from home during second-round visits. With respect to political variables, the two groups are broadly comparable: 2012 voting behavior as well as evaluation of the incumbent were very similar. Finally, the attrition rate in the treatment group was 3 percentage points higher than in the control group. While this difference is not large, it is statistically significant with a  $p$ -value of 0.03.<sup>5</sup> While the significant difference in attrition is a potential threat to inference, the fact that attrition is weakly correlated with political variables suggest that any bias would likely be small.

### 2.3.2 Measurement Error

A second important threat to the validity of our estimates is the fact that our main dependent variables are self-reported. In Brazil, electoral precincts (*seçôes*) do not correspond to specific, mutually exclusive geographical units in which voters reside, so we did not have the option of randomizing and measuring outcomes at the precinct level, as some other projects did. Self-reported outcomes are, of course, subject to recall, demand, and pro-winner biases. Of these three potential sources of error, demand effects are most likely to induce a correlation between measurement error and the treatment, as voters who received fliers may seek to pander to survey enumerators and falsify their vote choice. Recall effects, even if correlated with treatment, are unlikely to be correlated with the outcome of interest—vote for the incumbent mayor—so they should not affect our estimates. Likewise, the well-known propensity to over-report vote for the winning candidate is also unlikely to induce differential measurement error because our block randomization design ensures that we compare voters *within* municipalities, who would all presumably experience the same pro-winner effect.<sup>6</sup>

To reduce all three forms of bias, we used a secret ballot vote choice question. Respondents were given municipality-specific printed ballots (see Figure 7 for an example) and asked to privately check off their vote choice, fold the ballot, and deposit it into a sealed "ballot box" carried by the enumerator. To facilitate recall, our ballots included the same candidate photographs that are displayed on the confirmation screen of the electronic voting machine. Respondents also had the option of marking a blank or null vote, as is possible with electronic voting.

To assess the extent to which we successfully minimized measurement error, it is informative to benchmark our survey against actual election outcomes. This exercise has limitations, as our sampling frame omitted 16- and 17-year-olds (enfranchised in Brazil) as well as residents of the most rural census tracts. Furthermore, our target population excludes voters who do not reside the municipality where they are registered to vote, a common occurrence in areas, such as rural Pernambuco, that have experienced substantial out-migration. Non-resident registered voters tend to inflate official abstention rates, which would not be captured in our survey. Despite these caveats, benchmarking our sample can be informative, in that substantial deviations from electoral outcomes could be indicative of measurement error.

To construct a benchmark for comparison, we weighted each municipality's electoral results in proportion to its share of respondents in the endline sample. Results are displayed in Figure 8. The largest discrepancy is with respect to turnout, which could be driven by social desirability

---

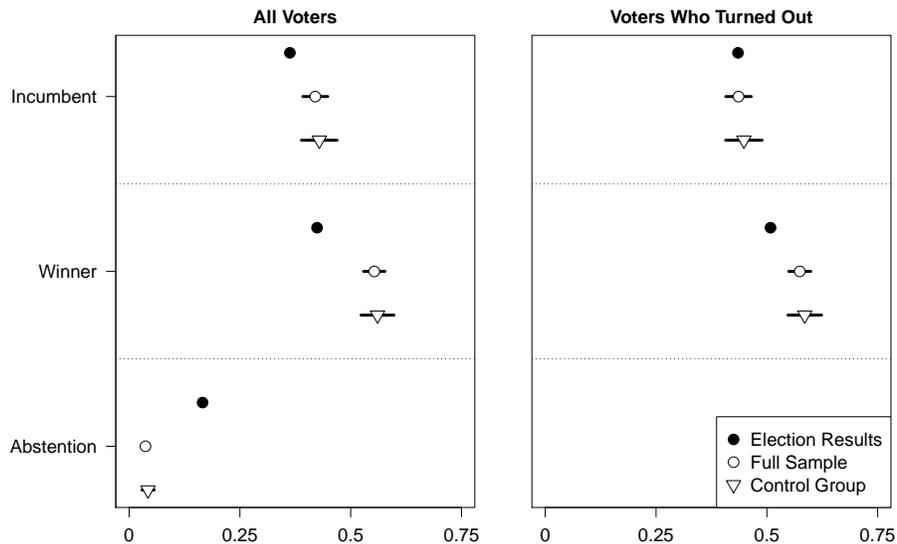
5. This  $p$ -value was estimated using permutation inference, testing the null of no effect on any individual's propensity to attrit. We employed the  $t$ -statistic from the specification listed in Section 3 as the test-statistic and performed 10,000 simulated randomizations.

6. If treatment causes a pro-winner (as opposed to pro-incumbent bias), this source of measurement error could lead to erroneous inferences. It is difficult to conceive of a scenario in which this would be the case.

**PARA PREFEITO DE ABREU E LIMA**

	<u>NOME</u>	<u>NÚMERO</u>	<u>PARTIDO</u>	
	KATIANA GADELHA	12	PDT	<input type="checkbox"/>
	FLAVIO GADELHA	15	PMDB	<input type="checkbox"/>
	PR. MARCOS JOSÉ	40	PSB	<input type="checkbox"/>
	BRANCO / NULO			<input type="checkbox"/>

**Figure 7** – Simulated Ballot Used to Measure Vote Choice



**Figure 8** – Vote for Mayor: Sample vs. Official Election Results

bias but probably also reflects our inability to sample non-residents. As is common in post-election surveys, we also find that voters somewhat over-report vote for the winning candidate. When we exclude abstainers, however, our estimate of the incumbent vote share is statistically indistinguishable from the official electoral results. Furthermore, the difference between the official vote returns and self-reported vote for the winning candidate diminishes—but does not disappear—after conditioning on turnout.

### 2.3.3 Candidate Self-Selection

A final potential concern is that, among mayors with rejected accounts, politically weak incumbents may have expected to lose and chosen not to run again. The remaining incumbents with rejected accounts might have a strong record of achievement or attractive personal qualities, making it less likely that supporters would change their vote when presented with negative information. If so, our sample would exclude those places where effects might be larger, creating a bias in favor of a null finding.

Evidence from our pilot argues against this interpretation. About a month prior to the candidate registration deadline, we conducted a large-scale ( $n = 2000$ ) pilot study in all municipalities where the incumbent was eligible to run for reelection, providing the same treatment information and inquiring about intended vote if the incumbent were to rerun. We obtained similar results to those from the field experiment, suggesting that the findings reported below are not an artifact of self-selection into the sample of candidates.

## 2.4 Implementation Challenges

In the vast majority of municipalities, our project encountered no implementation difficulties, but in four municipalities where the mayor’s accounts had been rejected, our survey prompted reactions from local politicians or their allies. While the negative valence of our treatment

information may partially account for these reactions, it is also important to remember that we sampled a much larger fraction of the electorate in these municipalities, making the intervention more noticeable to local political actors. According to enumerators, reactions from local politicians were most often triggered not by the content of the fliers but rather by questions in the baseline survey about the likelihood of vote buying—an activity that is illegal in Brazil and severely punished by electoral authorities.

Political reactions to our survey fell into two categories: inquiries or complaints through official channels, and harassment of enumerators in the field. In two municipalities, public servants sent an email to our official project account or complained to the Brazilian IRB and State Accounts Court. The complaint to the Accounts Court generated some concern, and they asked us to cease the intervention in the corresponding municipality, but fieldwork had already been completed so our results were unaffected. In three municipalities, enumerators were harassed by allies of local politicians, and in some cases, they were followed and observed during fieldwork. As a result, they were unable to finish a handful of interviews (7 during the first round and 48 during the second) in two municipalities. In both cases, incidents were confined to one or two census tracts in peripheral communities where there was little or no police presence. Fieldwork was unaffected in more centrally-located parts of these municipalities, from which the majority of respondents had been sampled.

## 2.5 Ethical Considerations

In concert with the overall objectives of the Metaketa initiative, we sought to ensure that our study adhered to ethical principles. First, one of the Principal Investigators is Brazilian and a resident of the state of Pernambuco, so our study is not an instance of a strictly foreign team of academics intervening in an election abroad. Second, we obtained approval from the Institutional Review Boards of each of our universities.<sup>7</sup> This includes the Comit  de  tica em Pesquisa (Ethics in Research Committee) of the Federal University of Pernambuco, which, like all Brazilian IRBs, generally reviews only medical studies.<sup>8</sup> Thus, we went much further in obtaining approval than is typically done for studies of Brazil, especially those done by Brazilian social scientists.

Third, we partnered with the Brazilian government agency that produces the auditing decisions that we conveyed to voters in the common arm. The TCE-PE gave us formal permission to use its name and logo in the study, and it reviewed, requested modifications to, and ultimately approved the final version of the fliers. While we cannot claim that the intervention would have happened anyway without our participation, the design of the study is entirely consistent with the public education mission of the TCE-PE. We presented our research proposal to the TCE-PE as evaluating a method of direct outreach to individual citizens that they might consider adopting in the future.

Fourth, prior to the experiment we had very little basis for believing that our intervention could change the outcome of the election. In no municipality did our treatments reach more than 1 percent of the electorate. To estimate the number of votes moved in each municipality prior to launching the study, we relied on a full-scale ( $N = 2000$ ) pilot conducted in July 2016, administering the same treatments in many of the same municipalities. Based on the average treatment effects on intended vote for mayor in our pilot study, we estimated that our interventions could shift the votes of 8 percent of the number of treated voters in each municipality.

---

7. Boston University, protocol 4094X; MIT, protocol 1604551604; Universidade Federal de Pernambuco, n mero de parecer 1571592.

8. IRB review in Brazil is similar to the situation in the United States, in that university-based committees review research protocols according to a set of regulations that are defined at the federal level and apply nationwide.

In most places, this amounts to a mere 2 votes; in the most heavily sampled municipality, it constitutes 22 votes. As shown in our pre-analysis plan, our treatment effect would have had to be 3.25 times larger than estimated in the pilot study to have had a chance of changing the outcome of the closest prior election in these 47 municipalities over the past 16 years. In fact, our treatment effect was much smaller than estimated in the pilot, so it is even more unlikely that our intervention made the difference between any candidate winning and losing.

### 3 Results

To examine the overall impact of our treatment, we estimate the average treatment effect using the following estimating equation:

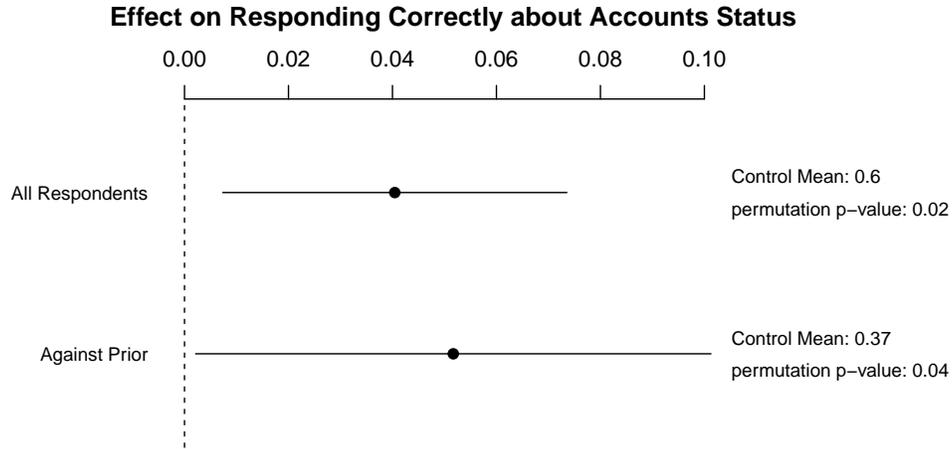
$$Y_{im} = \beta_0 + \beta_1 \mathbf{T}_{im} + \sum_{j=1}^k (\mu_j X_{im}^j + \gamma_j X_{im}^j \cdot \mathbf{T}_{im}) + \epsilon_{im} \quad (1)$$

$Y_{im}$  is the outcome variable for individual  $i$  in municipality  $m$ ,  $\mathbf{T}_{im}$  is the treatment indicator,  $X_{im}^j$  is the  $j$ th pre-treatment covariate (demeaned using the sample average) and  $\epsilon_{im}$  is the disturbance term. In the results presented here,  $X_{im}^j$  only include census tract dummies, which are our blocking variable; in the appendix, we present all relevant specifications pre-specified in the “Meta-PAP,” including those with covariate adjustment. Because we demean the covariates and include their interaction with treatment,  $\beta_1$  is a consistent estimator for the average treatment effect (Lin 2013). For the standard error of our estimates, we employ the “HC2” heteroskedastic consistent estimator. In addition to conventional inference, we also test the sharp null of no treatment effect using permutation inference in each of our specifications. Our test-statistic is the  $t$ -statistic of our experimental estimate. Our pre-specified hypotheses are directional, so reported  $p$ -values are one-sided.

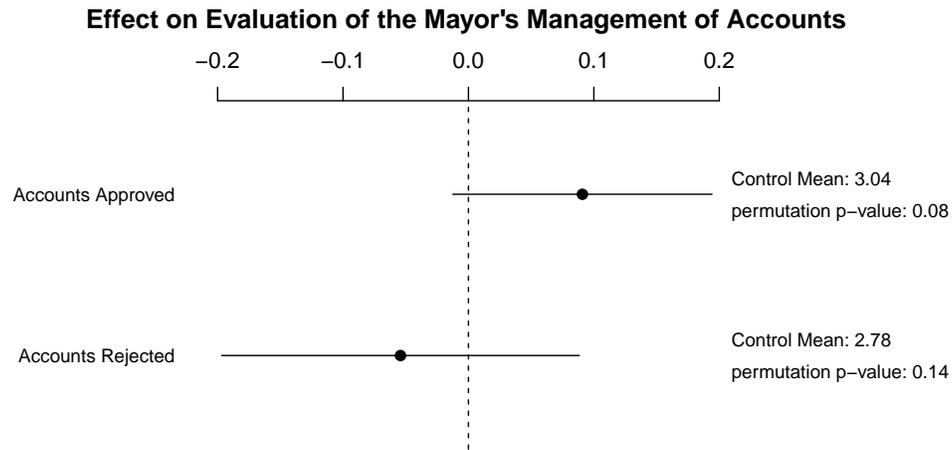
While the effect on vote choice is of primary interest, we first show the effect of our intervention on intermediate knowledge and attitudinal variables that should be affected if the information were to change behavior. Specifically, we examine whether or not respondents learn as a result of the treatment, change their evaluation of the mayor in the specific domain of management of accounts, decrease self-reported uncertainty over their evaluations, and change their overall evaluation of the mayor’s record.

First, we show that our intervention increased citizens’ knowledge about whether their mayor’s accounts had indeed been rejected or approved in 2013. In Figure 9, we present the estimated effect of treatment on respondents’ knowledge of the TCE’s decision on their mayor’s accounts, both for the full sample and for as those who had been incorrect about the mayor’s accounts at baseline. About 65% and 60% of respondents provided the correct answer in the treatment and control groups, respectively; the difference is statistically significant. While statistically detectable, the intervention did not dramatically increase the number of respondents giving correct answers. This suggests that many respondents either forgot the information or did not believe it. Among the group of respondents who answered *incorrectly* at baseline, about 37% of control group respondents provided the correct answer at endline, possibly indicating that some learning occurred over the course of the campaign.

Upon learning the information, did respondents change their assessment of the incumbent’s handling of the municipality’s accounts? As evident in Figure 10, we find some evidence that respondents changed their views of the incumbent’s performance on this dimension in the expected directions, as measured on a 5 point scale. In both groups, however, the effect is

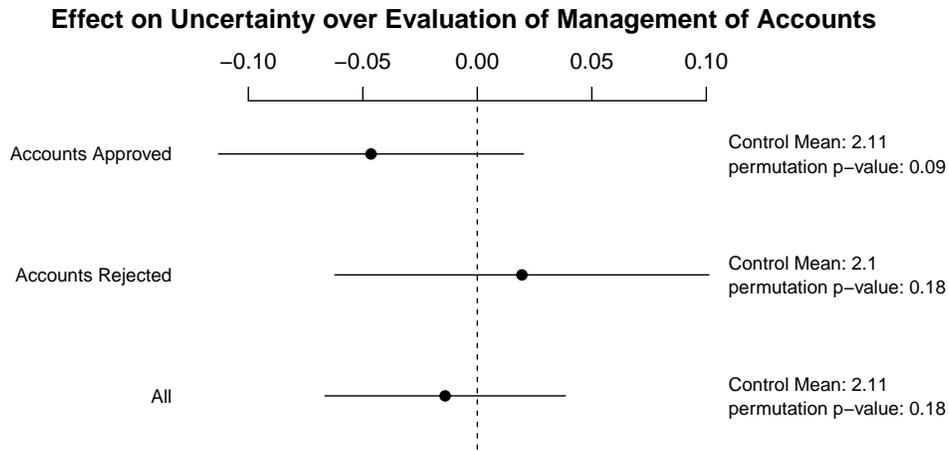


**Figure 9** – Effect of Treatment on Learning. Dependent variable is a variable measuring if the respondent answered correctly about whether municipality accounts had been rejected by the TCE. The "Against Prior" sample consists of respondents who were incorrect about accounts status at baseline. Lines are 95% confidence intervals. P-values in the right margin are from one tailed tests computed using permutation inference.



**Figure 10** – Effect of Treatment on Evaluation of Mayor's Management of Accounts. Dependent variable is the respondent's assessment of the mayor's management of the municipality's accounts on a 5 point scale (higher values indicating a more positive evaluation). Lines are 95% confidence intervals. P-values in the right margin are from one tailed tests computed using permutation inference.

imprecisely estimated. If we combine both groups and change the polarity of dependent variable to match the valence of the information (not shown), then the effect estimate is statistically significant with a point estimate of about .07. This latter specification is not pre-specified, however.



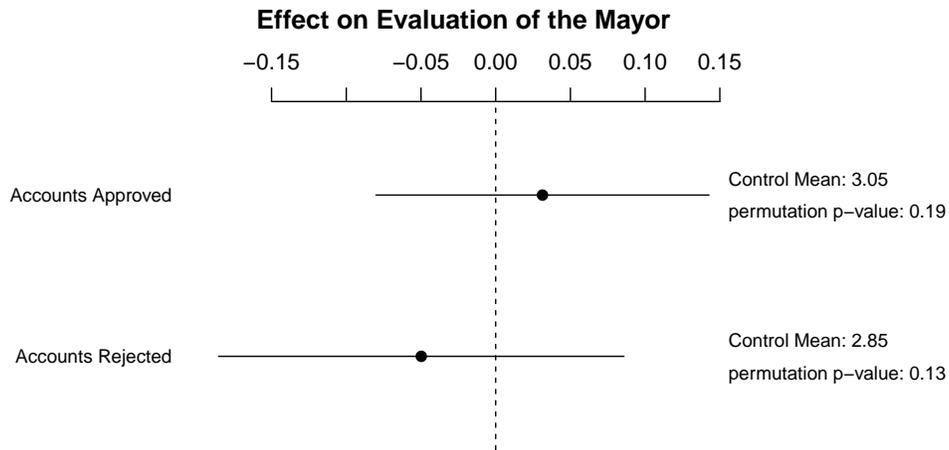
**Figure 11** – Effect of Treatment on Uncertainty Over Evaluation of Mayor’s Management of Accounts. Dependent variable is the respondent’s uncertainty over their own assessment of the mayor’s management of the municipality’s accounts on a 5 point scale (higher values indicating more uncertainty). Lines are 95% confidence intervals. P-values in the right margin are from one tailed tests computed using permutation inference.

We find inconsistent evidence that respondents’ uncertainty over their assessment of the mayor’s management of the municipal accounts diminishes as a result of treatment. Among respondents in accounts approved municipalities, the average effect of the treatment is to diminish uncertainty by .05 on a five point scale. In accounts rejected municipalities, however, we find a small positive (and insignificant) effect. Overall, the effect is negligible.

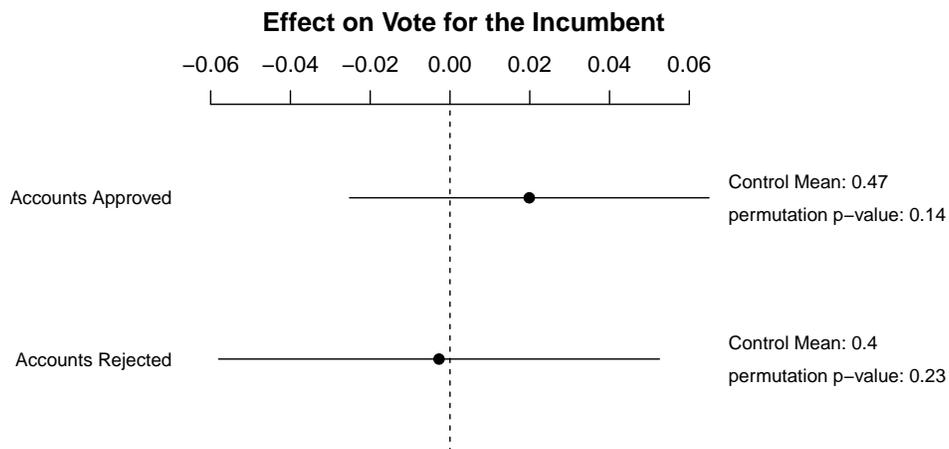
Finally, we find null effects with respect to our main dependent variable, vote for the incumbent, as well as evaluation of the incumbent’s record in office. As shown in Figures 12 and 13 the point estimates are in the expected direction but small and statistically insignificant. For vote choice, the point estimate for respondents in both types of municipality is no greater than .02. Accounting for sampling uncertainty, we can rule out effect sizes greater than about .06 with 95% confidence. Overall, these results indicate that the respondents’ vote choices are not sensitive to the information we distributed. We turn to possible explanations of this result in the next section.

## 4 Explanation of Effects

Why might our intervention have failed to exert any significant effects on the decision to vote for or against mayors running for reelection? One obvious answer might be that Brazilians fail to condemn corruption or malfeasance by elected officials. Yet numerous empirical studies and



**Figure 12** – Effect of Treatment on Evaluation of Mayor's Record. Dependent variable is the respondent's assessment of the mayor's record on a 5 point scale (higher values indicating a more positive evaluation). Lines are 95% confidence intervals. P-values in the right margin are from one tailed tests computed using permutation inference.



**Figure 13** – Effect of Treatment on Vote for the Incumbent. Dependent variable is the respondent's self-reported vote. Lines are 95% confidence intervals. P-values in the right margin are from one tailed tests computed using permutation inference.

public opinion surveys suggest that, to the contrary, Brazilians are some of the most intolerant of official corruption in the world. In the most recent wave of the World Values Survey, Brazil ranked 8th out of 60 countries in the percentage who said that accepting a bribe is "never justifiable." In the AmericasBarometer surveys from 2004–2014, Brazilians have the highest sustained levels of popular concern with corruption in the region, judged by an open-ended question about the most serious problem facing the country (Boas, Hidalgo, and Melo 2017).

In the context of Brazilians' staunch opposition to corruption, survey experiments presenting voters with hypothetical vignettes about a corrupt mayor running for reelection have found large and statistically significant electoral punishment effects, of much greater magnitude than similar studies in Colombia, Moldova, Peru, and Sweden (Avenburg 2016; Botero et al. 2015; Klačnja and Tucker 2013; Vera Rojas 2017; Weitz-Shapiro and Winters 2017; Winters and Weitz-Shapiro 2013, 2016). In our own survey, we were able to replicate these large negative effects when presenting our treatment information about the rejection of accounts in the context of a hypothetical vignette (Boas, Hidalgo, and Melo 2017). The strong negative response in the vignette experiment shows that voters consider a rejection of accounts to be worthy of punishment; hence, the null effects of our field experiment should not be attributable to miscomprehension of the delivered information.

Our survey also underscores that Brazilians strongly support the horizontal accountability mission of the TCE-PE, especially under the original sanctions regime of the Clean Slate Law. In the second wave of our survey, we asked respondents whether mayors who had had their accounts rejected by the TCE should have the right to run for reelection. In the full sample of respondents, 91% answered "no." Even among respondents who reported voting for the incumbent mayor and had been informed of the rejection of his or her accounts, 84% said that such mayors should not have the right to run again—effectively claiming that the candidate they supported should not have been on the ballot.

Another explanation for the null finding might be pervasive cynicism about incumbents and challengers. If voters do not perceive the other candidates in the race as better alternatives with respect to corruption or malfeasance, negative information might not affect their voting behavior. The problem with this explanation is the lack of any effect of *positive* information. If voters' ex-ante opinions about all candidates were highly negative, the provision of positive information should induce some citizens to vote for incumbents with approved accounts since this new information would counter their priors. The absence of both positive and negative effects suggests that low expectations is an unlikely reason for our null findings.

Rather than suggesting that Brazilians do not care about corruption or malfeasance or that their expectations are already too low, we argue that they fail to act upon a strong anti-corruption norm due to a variety of factors that constrain voting behavior (Boas, Hidalgo, and Melo 2017). Here, we highlight two particularly important constraints: the greater salience of more tangible aspects of incumbent performance, such as job creation and the quality of health services, as well as voter loyalty to political dynasties, which serves as a functional equivalent to strong party identification in many municipalities. In addition to our survey, this section draws upon several sources of qualitative data: background reports on 14 municipalities prepared by Brazilian research assistants and post-electoral focus groups in 3 municipalities.

Our research makes it clear that corruption in general, and the judgment of a mayor's accounts by the TCE-PE in particular, are relatively low-salience concerns for voters in Pernambuco.<sup>9</sup> In the baseline survey we asked respondents to name the biggest problem in their

---

9. One possible factor behind the divergence between our findings and those of Ferraz and Finan (2008), which finds large effects of releasing audit information, is that actors such as the media are essential for making audit results salient, as well as facilitating coordination among citizens who wish to act on the information.

municipality, and in the endline survey we asked what issue candidates had most discussed during the campaign. At the top of both lists are health (mentioned by 33% as the biggest problem, and 46% as the biggest campaign issue), crime (15% and 8%), employment (15% and 10%), and dealing with a severe drought affecting much of the state (15% and 7%). Only 2% mentioned corruption or municipal accounts as the biggest campaign issue, and fewer than 1% considered it the biggest problem.

Evidence from the focus groups underscored these findings from the survey. Asked about problems in their municipality, participants most often mentioned poor employment prospects, an issue exacerbated by the region's severe drought, which has made it difficult to earn a living in agriculture. Issues related to corruption and municipal accounts never arose spontaneously. When asked about the quality of the municipal government's "financial management," a term used in the survey to refer to the status of the mayor's accounts, participants talked instead about whether the municipal government paid public servants on time.

A second factor that likely constrains voters' responses to information about malfeasance concerns their loyalty or opposition to traditional political dynasties. Mass partisanship is relatively weak in Brazil, so it is unlikely to play the role that it is often thought to play in advanced democracies—limiting the effect on voting behavior of information gleaned during the campaign. Yet dynastic politics is likely to serve as a functional equivalent in many small towns. In the majority of our fourteen case study municipalities, one or more of the principal candidates for mayor in 2016 was a close relative—parent, child, grandchild, sibling, niece/nephew, or current or former spouse—of a former mayor in that municipality. In some instances, candidates' families had dominated municipal politics for decades. Their campaign strategies often made these family ties explicit, such as featuring photos and names of ex-mayor relatives in their advertising materials.

Evidence from the focus groups underscores that loyalty to political dynasties may serve a similar function as traditional partisan attachments in more established democracies. While members of local political dynasties often switch formal party affiliations from one election to the next, participants often used the term "party" to refer to these groups. According to one participant in Flores, "all my life it's been two parties, either one of them has 5000 votes guaranteed, and there are 2–3000 votes left for them to dispute... the candidate can be Joe Nobody, he enters and gets 5000 votes." In Tabira, another participant said that "whoever votes for that party never ceases to be [loyal]... it's a real tradition. They are people that put on the shirt of their team and never take it off."

In sum, while Brazilian voters strongly condemn corruption and malfeasance in the abstract, it simply ranks too low on their priority lists to have much chance of influencing voting behavior in real elections. Some residents of small towns may support the candidate of a local political dynasty out of longstanding loyalty to that particular clan. For voters such as these, any aspect of incumbent performance may have little influence over their decisions. Where local dynasties are weaker, residents may be more inclined to reward good performers and punish bad ones. Yet those voting based on performance criteria may be largely swayed by a mayor's record on tangible and highly salient issues such as job creation and local health services, leaving little room for additional information about the judgment of an auditing agency to influence their decisions.

## 5 Results from the Pilot and Second Arm

Consistent with the idea that issue salience matters for the effectiveness of informational interventions, results from our pilot study and second arm show that certain types of information about incumbent performance can change voting behavior, but only for those citizens with a personal stake in the issue. In our large-scale pilot study, we provided information on municipal efforts at combating mosquito-borne illnesses such as Zika, dengue, and chikungunya, and in our second arm of the field experiment, we provided information on standardized test scores in municipality-run elementary schools. Mosquito-borne illnesses were a potentially salient issue at the time of our fieldwork because Pernambuco was the epicenter of an outbreak of congenital Zika syndrome, the series of severe birth defects, including microcephaly, associated with the Zika virus. For its part, school performance is potentially salient for parents of children enrolled in local schools, thus providing meaningful heterogeneity in the degree to which respondents had a personal stake in the policy.

With respect to the information on combatting mosquito-borne illnesses, our pilot study provided voters in the treatment condition with information about the municipality's hiring of Anti-Endemic Disease Agents (*Agentes de Combate às Endemias*, ACE), specialized public health workers who visit homes to combat mosquitos and to teach residents about disease prevention. Municipalities are in charge of hiring these agents, using both federal and municipal funds; to benchmark their efforts, we use the number of agents funded by the Federal government. Thus, our performance indicator is the ratio of ACE agents per municipality to the maximum number funded by the federal government, which varies substantially across Pernambuco. Treatment information was delivered in a manner similar to that of the field experiment, but vote intention was recorded immediately after information delivery. Because the outcome variable was measured in the same survey—rather than several weeks later, as in our panel study—one might expect larger treatment effects.

Despite the intense media coverage of the Zika outbreak, we find that providing information about the municipality's efforts to combat mosquito-borne illnesses has no detectable effect on intended vote for the incumbent mayor, regardless of whether the municipality was a good or poor performer on our metric (Boas and Hidalgo 2017). These findings also apply to parents of young children or those planning to conceive in the next several years, a population that might be particularly concerned about the effects of the Zika virus on fetal development. However, we find strong reactions to negative information among respondents who know someone with a child affected by microcephaly or the Zika virus. For this group, the treatment lowers support for the incumbent by 37.7 percentage points in poor performing municipalities. Hence, only among respondents with a personal connection to the negative consequences of Zika—for whom information about mosquito control should be particularly salient—do we observe any electoral sanctioning effect.

We find a similar result when examining voters' response to information about municipal school performance (Boas, Hidalgo, and Toral 2017). The second arm of the field experiment informed about changes in scores on the National Literacy Evaluation (*Avaliação Nacional de Alfabetização* or ANA) during the mayor's first term. Among all respondents, we find an unexpected result: voters tend to punish good municipal performance on the ANA and are generally indifferent to poor performance. However, this result masks considerable heterogeneity by whether the respondent has a child enrolled in a municipal school. Among parents of enrolled children, for whom the issue should be most salient, we find the expected relationship: voters punish poor performance and reward (or are indifferent to) good performance. As with anti-disease efforts, a personal connection to the policy in question appears to be a prerequisite for

information about incumbent performance to change voting behavior.

## **6 Conclusion**

Vertical accountability might seem to offer a recourse for institutions of horizontal accountability whose efforts to sanction officeholders are blocked by legal obstacles or political maneuverings. While executives may be able to thwart the constraints imposed by other government entities, in a democracy they are more vulnerable to the punishment imposed by voters. If institutions like Brazil's State Accounts Courts can communicate their decisions directly to the public, they can potentially induce citizens to act directly, voting against and possibly defeating corrupt or malfeasant incumbents. The strength of Brazil's anticorruption norm—confirmed in multiple surveys, including our own—suggests that voters might indeed have the will to do so.

The results of our Metaketa project underscore the limits of this form of "roundabout horizontal accountability." While Brazilian voters condemn corruption and malfeasance in the abstract, they fail to take action in a real election when presented with the same sort of information about their own mayor. We argue that a variety of other factors serve to constrain voting behavior, including attitudes toward local political dynasties and the greater salience of more tangible aspects of incumbent performance, such as job creation. Hence, our findings underscore that robust, direct horizontal accountability is the most promising way to combat corruption and malfeasance in Brazil.

## References

- Avenburg, Alejandro. 2016. "Corruption and Electoral Accountability in Brazil." Ph.D. Dissertation, Boston University.
- Boas, Taylor C., and F. Daniel Hidalgo. 2017. *Electoral Intentives to Combat Mosquito-Borne Illnesses: Experimental Evidence from Brazil*, Working Paper.
- Boas, Taylor C., F. Daniel Hidalgo, and Marcus A. Melo. 2017. *Norms vs Actions: Voting against Malfeasance in Brazil*, Working Paper.
- Boas, Taylor C., F. Daniel Hidalgo, and Guillermo Toral. 2017. *Evaluating Students and Politicians: Test Scores and Electoral Accountability in Brazil*, Working Paper.
- Botero, Sandra, Rodrigo Castro Cornejo, Laura Gamboa, Nara Pavao, and David W. Nickerson. 2015. "Says Who? An Experiment on Allegations of Corruption and Credibility of Sources." *Political Research Quarterly* 68 (3): 493–504.
- Falleti, Tulia G. 2010. *Decentralization and subnational politics in Latin America*. Cambridge University Press.
- Ferraz, Claudio, and Frederico Finan. 2008. "Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes." *The Quarterly Journal of Economics* 123 (2): 703–745.
- Hidalgo, F. Daniel, Júlio Canello, and Renato Lima-de-Oliveira. 2016. "Can Politicians Police Themselves? Natural Experimental Evidence From Brazil's Audit Courts." *Comparative Political Studies* 49 (13): 1739–1773.
- Klašnja, Marko, and Joshua A. Tucker. 2013. "The economy, corruption, and the vote: Evidence from experiments in Sweden and Moldova." *Electoral Studies* 32 (3): 536–543.
- Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7 (1): 295–318.
- Melo, Marcus André, Carlos Pereira, and Carlos Mauricio Figueiredo. 2009. "Political and Institutional Checks on Corruption." *Comparative Political Studies* 42 (9): 1217–1244.
- Novaes, Lucas M. 2017. "Disloyal Brokers and Weak Parties." *American Journal of Political Science*.
- Paiva, Natalia, and Juliana Sakai. 2014. *Quem São os Conselheiros dos Tribunais de Contas*.
- Speck, Bruno W. 2011. "Auditing Institutions." In *Corruption and Democracy in Brazil: The Struggle for Accountability*, edited by Timothy J. Power and Matthew M. Taylor, 127–161. Notre Dame, IN: University of Notre Dame Press.
- Vera Rojas, Sofía Beatriz. 2017. "The Heterogeneous Effects of Corruption: Experimental Evidence from Peru." Manuscript, University of Pittsburgh.
- Weitz-Shapiro, Rebecca, and Matthew S. Winters. 2017. "Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil." *Journal of Politics* 79 (1): 60–74.
- Winters, Matthew S., and Rebecca Weitz-Shapiro. 2013. "Lacking Information or Condoning Corruption: When do Voters Support Corrupt Politicians?" *Comparative Politics* 45 (4): 418–436.

Winters, Matthew S., and Rebecca Weitz-Shapiro. 2016. "Who's in Charge Here? Direct and Indirect Accusations and Voter Punishment of Corruption." *Political Research Quarterly* 69 (2): 207–219.

# Appendices

## A Meta-PAP Pre-Specified Results

This section contains all applicable Meta-PAP pre-specified results. In addition to estimates and heteroskedastic-consistent standard errors, we present permutation  $p$ -values for the sharp null hypothesis of no effect. To calculate  $p$ -values, we drew 10,000 draws from the randomization distribution under the null. The test-statistic is the regression coefficient divided by the standard error.

For specifications with covariate adjustment, the following variables are included in the estimating equation:

- Age
- Gender
- Likelihood that incumbent will engage in vote buying
- Likelihood that the vote is not secret
- Likelihood that vote count is accurate
- Relative well-being
- Respondent most interested in receiving information about corruption
- Respondent most trusts information coming from a person conducting a survey
- Identifies with the party of the mayor
- Same race as mayor
- Turnout in previous mayoral election
- Vote for the incumbent in previous mayoral election
- Years of education

For variables that were asked using a Likert scale format, we transform the ordinal-scale variable into a numeric variable with equal intervals between each possible response.

Because all included covariates are demeaned, main effects—even when interactions are included—should be interpreted as the estimate of the average treatment effect.

### A.1 Family 1

	Good News	Good News (cov. adj)	Bad News	Bad News (cov. adj)
Control Mean	0.462 (0.015)	0.471 (0.014)	0.400 (0.019)	0.401 (0.019)
ATE Estimate	0.020 (0.023)	0.015 (0.020)	-0.003 (0.028)	-0.002 (0.027)
Num. obs.	847	847	818	818
ATE Permutation P-Value	0.288	0.326	0.456	0.473

**Table A1** – Family 1 Results. Outcome is self-reported vote for the incumbent. Estimates from specifications listed as part of the "Primary Family" in the Meta-PAP.

Results for the hypotheses in the "Primary" family are presented in Table A1. Specifications in the first and third columns do not include covariates.

## A.2 Family 2

	Good News	Good News (cov. adj)	Bad News	Bad News (cov. adj)
Control Mean	0.960 (0.007)	0.956 (0.006)	0.948 (0.010)	0.945 (0.009)
ATE Estimate	0.015 (0.009)	0.020 (0.008)	0.009 (0.012)	0.015 (0.012)
Num. obs.	847	847	818	818
ATE Permutation P-Value	0.162	0.071	0.247	0.137

**Table A2** – Family 2 Results. Outcome is self-reported turnout. Estimates from specifications listed as part of the "Secondary Family" in the Meta-PAP.

Results for the hypotheses in the "Secondary" family are presented in Table A2. Note that turnout is self-reported. Specifications in the first and third columns do not include covariates.

## A.3 Family 3

	Effort (Good News)	Effort (Bad News)	Integrity (Good News)	Integrity (Bad News)
Control Mean	3.248 (0.041)	2.832 (0.046)	2.567 (0.043)	2.257 (0.049)
ATE Estimate	-0.058 (0.057)	0.059 (0.064)	0.018 (0.057)	-0.028 (0.068)
Num. obs.	810	760	825	776
ATE Permutation P-Value	0.292	0.209	0.403	0.360

**Table A3** – Family 3 Results. Outcomes are perceptions of incumbent effort and integrity. Estimates from specifications listed as part of the "Mediators" family in the Meta-PAP. All specifications include pre-specified covariates.

Results for the hypotheses in the "Mediators" family are presented in Table A3.

## A.4 Family 4

<i>Moderator:</i>	<i>Same Race as Mayor</i>		<i>Same Party as Mayor</i>		<i>Clientelism</i>	
	Good News	Bad News	Good News	Bad News	Good News	Bad News
<b>Treatment (ATE)</b>	0.015	-0.012	0.015	-0.012	0.01	-0.01
SE	0.020	0.025	0.020	0.025	0.02	0.02
Perm. P-Value	0.343	0.357	0.331	0.375	0.34	0.36
<b>Treatment x Moderator</b>	0.087	-0.058	0.018	0.038	-0.03	0.01
SE	0.063	0.063	0.015	0.014	0.03	0.02
Perm. P-Value	0.142	0.196	0.176	0.011	0.21	0.36
Num. obs.	847	818	847	818	847	818

**Table A4** – Family 4 Results. Outcome is self-reported vote for the incumbent. Estimate from specifications listed as part of "Substitution" family in the Meta-PAP. Moderator has been demeaned so main effect is the estimated ATE. All specifications include pre-specified covariates.

Results for the hypotheses in the "Substitution" family are presented in Table A4. The moderator "Same Race as Mayor" is whether the respondent believes the incumbent is the same race.

The “Clientelism” moderator is based on the question asking respondents to rate the probability that the incumbent will engage in vote buying.

## A.5 Family 5

<i>Moderator:</i>	<i>Uncertainty</i>		<i>Competition</i>		<i>Free and Fair</i>	
	Good News	Bad News	Good News	Bad News	Good News	Bad News
<b>Treatment (ATE)</b>	0.018	-0.009	0.015	-0.012	0.01	-0.01
SE	0.020	0.025	0.020	0.025	0.02	0.02
Perm. P-Value	0.331	0.398	0.332	0.359	0.35	0.37
<b>Treatment x Moderator</b>	0.031	0.012	1.114	-10.732	-0.02	0.02
SE	0.023	0.028	1.877	2.637	0.03	0.03
Perm. P-Value	0.148	0.371	0.376	0.006	0.26	0.28
Num. obs.	847	818	847	818	847	818

**Table A5** – Family 5 Results. Outcome is self-reported vote for the incumbent. Estimate from specifications listed as part of “Context” family in the Meta-PAP. Moderator has been demeaned so main effect is the estimated ATE. All specifications include pre-specified covariates.

Results for the hypotheses in the “Context” family are presented in Table A5. The moderator “Uncertainty” measures the degree of uncertainty that the respondent has over his or her prior. The moderator “Competition” is simply the vote margin between the top mayoral candidates in the 2012 election. “Free and Fair” is the sum of the questions asking voters to rate the probability that the vote is counted fairly and that their vote is secret.

## A.6 Family 6

<i>Moderator</i>	<i>None</i>	<i>Welfare Relevant</i>		<i>Source Credibility</i>	
	Full Sample	Good News	Bad News	Good News	Bad News
<b>Treatment (ATE)</b>	0.001	0.015	-0.012	0.015	-0.012
SE	0.016	0.020	0.025	0.020	0.025
Perm. P-Value	0.490	0.345	0.353	0.331	0.343
<b>Treatment x Moderator</b>		0.124	0.184	-0.094	-0.015
SE		0.075	0.075	0.106	0.112
Perm. P-Value		0.103	0.017	0.252	0.456
Num. obs.	1665	847	818	847	818

**Table A6** – Family 6 Results. Outcome is self-reported vote for the incumbent. Estimate from specifications listed as part of “Design” family in the Meta-PAP. Moderator has been demeaned so main effect is the estimated ATE. All specifications include pre-specified covariates.

Results for the hypotheses in the “Design” family are presented in Table A6. The results presented in the first column show the effect of the treatment on vote for the incumbent without distinguishing between good news and bad news. For the “Welfare Relevant” moderator, we use whether or not the respondent would most like to obtain information about politicians involved in corruption. For the “Source Credibility” moderator, we use whether or not the respondent would most trust information given by a person conducting a survey.

Note that we do not present results for the hypotheses labeled “Public Channels” and “Hawthorne” because these treatments were not part of our design.

## B Balance Tests

Variable	Estimate	SE	Perm. P-Value
2012 Incumbent Vote	0.011	0.018	0.356
2012 Turnout	-0.044	0.014	0.015
Age	0.328	0.670	0.366
Education	-0.052	0.180	0.445
Female	-0.050	0.020	0.054
Finds Info from Surveyors Credible	0.010	0.010	0.258
Free and Fair Scale	-0.046	0.052	0.289
Prob Vote Can't Be Monitored	-0.017	0.041	0.406
Prob. Vote Buying	-0.085	0.046	0.106
Prob. Vote Count Accurate	-0.030	0.033	0.276
Relative Wellbeing	0.030	0.081	0.389
Same Party as Mayor	-0.024	0.018	0.202
Same Race as Mayor	0.046	0.043	0.218
Wants Info about Corruption	0.020	0.016	0.169

**Table A7** – Covariate Balance Tests. Estimates are based on full sample.

Covariate balance tests on pre-specified covariates are presented in Table A7. These estimates are from a model with block fixed effects run on the full sample.