

Bureaucratic Quality and Electoral Accountability

Tara Slough*

August 21, 2022

Abstract

In many theories of electoral accountability, voters learn about an incumbent's quality by observing public goods outcomes. But empirical findings are mixed, suggesting that increasing the visibility of these outcomes only sometimes improves accountability. I reconcile these apparently conflicting findings by highlighting bureaucrats' role in the production of public-goods. In a simple model of electoral accountability involving a voter, a politician, and a bureaucrat, I show that accountability relations yield distinct empirical implications at different levels of bureaucratic quality. These predictions rationalize some existing conflicting empirical findings. To illustrate how my model makes sense of these otherwise inconsistent results, I develop a new research design—a theoretically structured meta-study—to synthesize existing findings. Evidence from two such meta-studies suggests that a common model of electoral accountability that allows for variation in bureaucratic quality can reconcile conflicting findings on accountability from multiple contexts.

*Assistant Professor, New York University, tara.slough@nyu.edu. Thanks to Deborah Beim, Natalia Garbiras Díaz, Germán Gieczewski, John Huber, John Marshall, Greg Martin, Lucy Martin, Pablo Querubín, Arturas Rozenas, David Stasavage, Mike Ting, and Stephane Wolton, and seminar audiences at Caltech, LSE, NYU, Washington University in St. Louis, and the 2022 Formal Theory and Comparative Politics Conference for comments. I am very grateful to Galileu Kim for sharing data and to the authors of the replicated studies for providing access to replication data. Jiawei Fu and Carolina Torreblanca provided outstanding research assistance. This project is supported in part by an NSF Graduate Research Fellowship, DGE-11-44155; NSF Doctoral Dissertation Research Improvement Grant 1729334; and grants from J-PAL, Evidence in Governance and Politics (EGAP), the Center for the Study of Development Strategies, and the Columbia Department of Political Science.

Electoral accountability is a normative goal of democracy (Przeworski, Stokes, and Manin, 1999). Yet, recent empirical assessments of the existence of these relationships between voters and politicians present grounds for pessimism. In developing democracies, widespread malfeasance by politicians, underprovision of public goods and services, and low levels of citizen political knowledge motivate questions about whether and when voters can hold these apparently poorly-performing politicians to account. More concerning, efforts to inform citizens about politician behavior or political outcomes provide little systematic evidence that citizens respond by sanctioning underperforming or rewarding high-performing politicians (Dunning et al., 2019).

I present a new explanation for these empirical patterns that focuses on bureaucratic quality as a constraint on politicians' incentives to provide public goods, even when voters are sufficiently informed and rational. The theory suggests that empirical manifestations of accountability across democracies should vary in bureaucratic quality. Specifically, I argue that higher bureaucratic quality increases the efficiency of public goods investments, which in turn gives politicians the incentive to invest in public goods over private rents. But politicians' investment behavior also influences what voters can learn from the observation of public goods outputs, and hence, how voters can use information to retain or select "good" politicians. I identify qualitatively distinct equilibria that emerge at different levels of bureaucratic quality, and how show how the observable implications of these equilibria can reconcile mixed empirical findings about accountability.

My account departs from recent arguments that focus on uninformed or irrational voters as a barrier to electoral accountability. In such accounts, citizens that lack access to an informational signal about a politician's type do not learn about the politician's quality (i.e., Ferraz and Finan, 2008; Humphreys and Weinstein, 2012; Dunning et al., 2019). Alternatively, citizens may receive information but fail to rationally update (Achen and Bartels, 2016; Healy, Malhotra, and Mo, 2010).¹ In either case, failure to access or update on performance-relevant information breaks the link between a politician's actions and her subsequent electoral fortunes, so the politician maximizes her own utility, regardless of whether their actions are congruent with voters' preferences.

¹But see Fowler and Montagnes (2015); Fowler and Hall (2018); Ashworth, Bueno de Mesquita, and Fridenberg (2018).

To understand the implications of my argument versus these existing alternatives, I develop a simple two-period model of electoral accountability with a bureaucrat. As is standard, voters evaluate politicians on the basis of the quality of observed policy outcomes, here levels of public goods provision (Ashworth, 2012). However, this model departs from standard practice by considering co-production of public goods by a politician and an unelected bureaucrat. The politician allocates a budget to public goods or private rents (corruption). In turn, bureaucrats produce public goods with the allocated funds. Bureaucratic quality influences the efficiency with which these public goods are produced. Politicians differ in their competence at “getting things done,” which translates into their ability to monitor bureaucrats. Voters prefer competent types and update their belief about a politician’s type on the basis of public goods outputs. The voter then decides whether or not to retain the incumbent for a second period or to elect a challenger.

When bureaucratic quality is low, it is inefficient for either type of politicians to fund public goods. As a result, both competent and incompetent politicians pool on corrupt behavior, allocating all funds to private rents instead of public goods. Due to this pooling, voters are not able to ascertain politician quality regardless of whether they observe the lack of public goods (the signal). At moderate levels of bureaucratic capacity, public goods provision is efficient for competent types but inefficient for incompetent types. In this case, voters’ observation of public goods allows for updating on politician type which yields higher retention of competent types and may induce incompetent type to make costly investments in public goods in their first term. Finally, at higher levels of bureaucratic quality, a partially-pooling and then pooling equilibria emerge in which both types invest in public goods with positive probability. As in the separating equilibrium, in these equilibria, voter observation of public goods outputs allows for updating which facilitates more frequent retention of competent politician.

These theoretical results have implications for how we interpret the results of empirical studies of electoral accountability. Suppose that researchers conduct an experiment in which they provide voters with information on an incumbent’s corruption (the politician’s behavior). At low levels of bureaucratic quality, information that the incumbent is corrupt should not have an effect on

voters beliefs about the politician or their voting behavior because both types of politicians pool by allocating the budget to rents. The pooling equilibria at high levels of bureaucratic quality further suggests that voters' beliefs should not change in response to revelation of "clean" (non-corrupt) politician behavior since both types of politicians pool on allocating the budget to public goods.²

I illustrate the plausibility of the theory by presenting two theoretically structured meta-studies that synthesize work on electoral accountability. In the first meta-study, I examine studies of accountability of Brazilian mayors. Specifically, I develop and validate a measure of bureaucratic quality across Brazilian municipal bureaucracies. I use this measure to extend seminal studies of corruption by politicians (Ferraz and Finan, 2011; Avis, Ferraz, and Finan, 2018); voter updating on politician performance (Weitz-Shapiro and Winters, 2016; Winters and Weitz-Shapiro, 2016); and the emergence of incumbency disadvantage for local mayors (Klašnja and Titunik, 2017). I find that a separating equilibrium where competent invest in public goods and incompetent politicians invest in rents emerges only in municipalities with comparatively low measures of bureaucratic quality (relative to the sample of Brazilian municipalities). In municipalities with higher quality bureaucracies, politicians abstain from appropriating private rents, politicians do not shirk in their second term, and voters do not update on a signal of a clean politician. This is consistent with the implications of a pooling equilibrium that emerges at high levels of bureaucratic quality.

In the second structured meta-study, I reconsider evidence from eleven field experiments on pre-election voter information campaigns. I first show that these experiments have only been conducted in democracies that are somewhat below average in bureaucratic quality. This is important because when experiments are conducted in settings where competent and incompetent (or good or bad) incumbents pool by making the same budget allocations, information about politician behavior should have no effect on voter beliefs or behavior, even when voters are Bayesian. I show suggestive evidence from these eleven studies that treatment effects of "good" and "bad" signals are indeed attenuated toward zero when bureaucratic quality is low (within the sample).

This paper makes theoretical and empirical contributions. The theory connects accountabil-

²In contrast, if the researchers provided information about public goods outputs rather than politician behavior the voter could update on the politician's type at high levels of bureaucratic quality.

ity models focused on a voter and politician (Fearon, 1999; Ashworth, Bueno de Mesquita, and Fridenberg, 2017) and to models of moral hazard in bureaucracies. Joining Yazaki (2018), Li, Sasso, and Turner (2019), and Foarta (2021), the model posits distinct roles for politicians and bureaucrats in the production of public goods as central to our understanding of accountability. This represents a departure from spatial models of delegation such as Fox and Jordan (2011) and analyses of optimal institutional design (e.g., Maskin and Tirole, 2004; Alesina and Tabellini, 2007) in which bureaucrats and politicians (ultimately) use the same instrument to affect policy. These departures from standard accountability models provide new insights about accountability relations manifest across the world's democracies.

The results engage large empirical literature on information and accountability in developing democracies (Dunning et al., 2019; Chong et al., 2015; Banerjee et al., 2011; Bhandari, Larreguy, and Marshall, 2019; Cruz, Keefer, and Labonne, 2018). The findings of these studies are mixed with respect to the (average) effects of information on voter beliefs and election outcomes (Enríquez et al., 2019). This study rationalizes mixed results in this literature and provides guidance for research design. Joining an important recent contribution by Raffler and Martin (2019), I argue that the co-production of public goods by bureaucrats and politicians conditions voters learning about politicians. In contrast to this work, however, I show that the finding that bureaucratic co-production of public goods hinders voter learning is not general across all levels of bureaucratic quality. When bureaucratic quality is (relatively) high, reliance on bureaucrats to produce public goods *improves* voters' information and thus promotes positive selection of politicians.

In sum, I suggest that the siloing of recent work—by method, context, and critique—obscures important contributions of these works to a more general understanding of democratic accountability. I show that a common theoretical framework that is symmetrically applicable in developing and advanced democracies can help to understand how accountability relations between voters and politicians might manifest in different settings. To show the plausibility of this account, I employ a novel research design—the theoretically structured meta-study—as a way to integrate and cumulate seemingly-unrelated or contradictory empirical findings.

1 Theory

Consider three actors: an incumbent politician, P , a bureaucrat, B , and a voter, V . In each of two periods, the politician and bureaucrat jointly produce public goods that are observed (or unobserved) by the voter. After first term public goods production, there is an election in which the politician re-election office against a challenger.

Politicians are of an incompetent or competent type, $\theta \in \{\underline{\theta}, \bar{\theta}\}$, respectively. The politician's type is private information to the politician and the bureaucrat. The voter holds a prior belief that the politician is a competent type with probability $Pr(\theta = \bar{\theta}) = \pi \in (0, 1)$. I conceive of competence as ability to manage the bureaucracy or “get things done” via oversight. A competent politician monitors the bureaucrat at intensity $\bar{\theta}$ while an incompetent politician monitors the bureaucrat at intensity $\underline{\theta}$, where $0 < \underline{\theta} < \bar{\theta} < 1$.

Public goods are produced as a function of the funding allocated by the politician in period t and the quality and effort of the bureaucracy. Specifically, politicians allocate a budget, normalized to 1 in each period, between public goods (a_t) and private rents ($1 - a_t$). I assume a binary allocation decision, i.e. $a_t \in \{0, 1\}$. I treat the quality of the bureaucracy, $q > 1$ as exogenous. While bureaucratic quality may be an outcome of policies pursued by a politician, the model simply assumes that quality is slow-moving and requires sustained investment to realize changes (Rauch, 1995; Huber and Ting, 2020).³ Recent evidence by Besley et al. (2021) suggests that, globally, bureaucratic quality has been remarkably persistent since at least 1900.

I assume that bureaucrat the bureaucrat exerts costly effort, e in response to some intensity of oversight, given by $\theta \in \{\underline{\theta}, \bar{\theta}\}$. As such, the utility of the bureaucrat, in period t , net of a wage satisfying his participation constraint, can be written:

$$u_t^B(e) = -\theta(1 - e_t) - \frac{e_t^2}{2} \quad (1)$$

³There is less systematic empirical evidence on reforms designed to improve bureaucratic quality. Notably, however, interventions studied in existing literature that intend to improve bureaucratic quality via hiring (selection) are initiated by higher levels of government from *outside* the localities they serve, not by local politicians who may be judged on the quality of their services (Dal Bó, Finan, and Rossi, 2013; Ashraf et al., 2020).

Note that θ is given by the politician's type. The bureaucrat is not forward-looking and chooses a level of effort in each period. This characterization of a bureaucrat is intentionally very simplistic as the focus of this paper is to draw out implications for the voter-politician accountability relations.

Given the allocation of funds by the politician and the effort exerted by a bureaucrat, the public good, $g_t(a_t, e_t)$ is produced according to the production function in (2). The production function assumes that allocation to public goods and bureaucratic quality are complements.

$$g_t(a_t, e_t) = \begin{cases} a_t q & \text{with probability } e_t \\ 0 & \text{with probability } 1 - e_t \end{cases} \quad (2)$$

This production function indicates that if the politician invests in public goods ($a_t > 0$), then the expected quantity of public goods outputs is increasing in bureaucratic quality (q) and effort (e_t). In contrast, if the politician starves public goods funding ($a_t = 0$), they are not produced. The production function in (2) further clarifies the relationship between bureaucratic quality, q , and broader notions of bureaucratic capacity. Bureaucratic capacity consists of both the skill of bureaucrats (Geddes, 1994), their allocation of bureaucrats across a jurisdiction (Acemoglu, García-Jimeno, and Robinson, 2015), and the effort exerted by bureaucrats. I capture the first two features in quality (q) and the third in bureaucratic effort (e_t). Thus, in the present framework both bureaucratic quality and effort increase the efficiency with which a politician's funding allocation is converted to a public goods output.

The politician trades off private rents for public goods when allocating the budget. Both types of politicians value the provision of public goods. The assumption that both types to value public goods provision is not common to all accountability models. Note, however, that many of the qualitative results of the model can be generated by a venal/non-venal politician type space (instead of competence). I note these departures when discussing results. Competent types' superior ability to induce the bureaucrat to work is captured in the realization of g_t in the per-period utility of the

(incumbent) politician in (3).

$$u_t^P(a_t; \theta) = 1 - a_t + g_t \quad (3)$$

The politician receives $u_t^P(a_t; \theta)$ for each period she is in office, and utility normalized to 0 if she is not in office. This normalization creates a re-election incentive for the politician. As such, the politician's utility over two periods is given by:

$$u^P(\mathbf{a}; \theta) = \begin{cases} 2 - a_1 - a_2 + g_1 + g_2 & \text{if re-elected} \\ 1 - a_1 + g_1 & \text{if not re-elected} \end{cases} \quad (4)$$

The voter may observe the realization of first-term public goods provision, and forms a posterior belief about the politician's type, μ . To understand the role of voter information—here, whether or not a voter observes the public goods—in generating results, I assume that the voter observes a signal, $z = g_1$, of first-term public goods with probability $p \in [0, 1]$ and that the voter does not receive a signal, $z = \emptyset$, with complementary probability $(1 - p)$. While it is natural to think that incumbents, challengers, or civil society would publicize performance signals, the information and accountability literature generally assumes some barrier to diffusion of this information, for instance, a lack of local media (Ferraz and Finan, 2008; Larreguy, Marshall, and Snyder Jr., 2020). The assumption of exogenous revelation solely maintains that these diffusion technologies are not manipulated by politicians in the short-run. Further, treating p as exogenous is consistent with the theoretical treatment of information revelation in experiments on information and accountability (Izzo, Dewan, and Wolton, 2019). In this paper, setting $p = 0$ allows for characterization of equilibria where voters are (completely) uninformed (see anecdotes throughout Dunning et al., 2019). In experiments that provide voters with information, we are interested in comparing how citizen beliefs and voting behavior vary in p .

The voter values consumption of the public good. This implies that the voter cares about a politician's competence to the extent that competent politicians produce more public goods (in

expectation). The voter's utility is thus given by expected public goods provision in the second period and a valence shock for the incumbent, parameterized as $\phi \sim U[-b, b]$, where $b > q$. The voter votes, $v \in \{i, c\}$, to re-elect the incumbent (i) or elect the challenger (c). If elected, a challenger acts as a first-period incumbent. For that reason, I index second-period actions by i and c , respectively. The voter's second period expected utility from the a vote for incumbent or a vote for a challenger (c), can be expressed:

$$E[u_2^V(i)] = E[g_2^i|z] + \phi \quad (5)$$

$$E[u_2^V(c)] = E[g_2^c] \quad (6)$$

I assume that if the challenger wins, she acts as a first-term incumbent. Thus, in evaluating $E[g_2^i|z]$ and $E[g_2^c]$, the voter considers differences in expected politician competence and differences in allocation behavior that depend on a politician's term as in Klašnja and Titunik (2017).

1.1 Sequence and Equilibrium Concept

The game proceeds according to the sequence:

1. Nature determines θ , the incumbent's competence. Only the incumbent and bureaucrat observe θ .
2. The incumbent allocates a_1 to the public good.
3. The bureaucrat exerts effort e_1 to produce the first-term public good, g_1 .
4. With probability p , the voter observes $z = g_1$ and forms a posterior belief about the politician's type, μ . The valence shock ϕ is revealed, and the voter chooses whether to re-elect the incumbent or elect the challenger.
5. If the incumbent was re-elected, she allocates a_2^i to the public good. Otherwise, the challenger allocates a_2^c to the public good.

6. If the incumbent was re-elected, the bureaucrat exerts effort e_2^i to produce the public good g_2^i . Otherwise, the bureaucrat exerts effort e_2^c to produce the public good g_2^c .

I characterize the Perfect Bayesian Equilibria (PBE) of the game. The incumbent's allocation decision is the choice $a_1 \in \{0, 1\}$. The bureaucrat's effort allocation is $e_1 \in \mathbb{R}_+$. Public goods production, $g_1 : \{0, 1\} \times \mathbb{R}_+ \rightarrow \{0, q\}$, maps the budget allocation and bureaucratic effort into a public goods output observed by all players. Voters update beliefs on the observation of public goods $\mu : \{0, q\} \rightarrow [0, 1]$ and the voter's voting strategy is a mapping $v : \{0, q\} \times [0, 1] \rightarrow \{i, c\}$. The second period incumbent's allocation strategy is a mapping $a_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \rightarrow \{0, 1\}$. Finally, second period bureaucratic effort and public goods production represents the mapping: $e_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \times \{0, 1\} \rightarrow \mathbb{R}_+$ and public goods provision represents the mapping $g_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \times \{0, 1\} \times \mathbb{R}_+ \rightarrow \{0, q\}$. As in many signaling games, there exist multiple equilibria in some regions of bureaucratic quality. I invoke the intuitive criterion refinement to ensure that the equilibria discussed below are unique (at each level of bureaucratic quality) (Cho and Kreps, 1987).

2 Equilibrium Analysis

First, consider the bureaucrat's equilibrium level of effort. By straightforward inspection of the bureaucrat's objective, it is clear that optimal effort, $e_t^* = \theta$. The bureaucrat's effort depends only on the politician's type in either period. When combined with Equation (2), this optimal effort indicates that politician competence and bureaucratic effort are complements with respect to the production of public goods. This contrasts with the idea that politician type (quality) and bureaucratic quality are substitutes, which is typically motivated by the observation that high-quality bureaucracies tend to insulate outputs from the follies of bad politicians. Instead, the model develops an alternate mechanism for this observed insulation focused on how bureaucratic quality shapes a politician's allocation decision.

Turning to the incumbent's second-term allocation strategy, the politician considers the expectation second-term public goods provision, $E[g_2^i(a_2^i, e_2^i)] = \theta q a_2^i$. Where $E[g_2^i(a_2^i, e_2^i)] \geq 1$, a

politician will invest the budget in public goods, $a_2^i = 1$. In contrast, where $E[g_2^i(a_2^i, e_2^i)] < 1$, a politician will invest nothing, $a_2^i = 0$. The politician's optimal second-period allocation strategy is given by:

$$a_2^{i*} = \begin{cases} 1 & \text{if } q \geq \frac{1}{\theta} \\ 0 & \text{else} \end{cases} \quad (7)$$

Intuitively, if the bureaucracy is of sufficiently low capacity (low q), neither type has an incentive to fund public goods because it is inefficient to do so. This implies that even a competent politician that values public goods outputs will “take the money and run” when the state is incapable of efficiently producing public goods. On the other hand, when q is sufficiently high, both types will fund public goods. The efficiency gains in the provision of public goods from a high-quality bureaucracy induce both types of politician to fund public goods, insulating outputs (to some extent) from incompetent politicians.

Consider the voter's voting decision. The voter votes for the incumbent if $E[u_2^V(i)] > E[u_1^V(c)]$. Given the distribution of the valence shock and denoting equilibrium allocation strategies for each term and politician type as \mathbf{a} , the incumbent's probability of victory is:

$$\begin{aligned} \tau(\mu, \mathbf{a}) &= \frac{1}{2} + \frac{E[g_2^i|z] - E[g_2^c]}{2b} \\ &= \frac{1}{2} + \frac{\mu E[g_2^i(a_2^i, e_2^i|\theta = \bar{\theta})] + (1 - \mu) E[g_2^i(a_2^i, e_2^i|\theta = \underline{\theta})]}{2b} - \\ &\quad \frac{\pi E[g_2^c(a_2^c, e_2^c|\theta = \bar{\theta})] + (1 - \pi) E[g_2^c(a_2^c, e_2^c|\theta = \underline{\theta})]}{2b} \end{aligned} \quad (8)$$

Turning to the voter's beliefs and voting decision, recall that the voter observes $z = g_1$ with probability p . With probability $1 - p$, the voter does not observe public goods outputs ($z = \emptyset$). Consider the latter case first. In this case, $\mu = \pi$, which follows (trivially) from Bayes' rule. If voters do not update, a politician's re-election fate is independent of her first-period allocation decision. If this occurs, the politician maximizes her utility by adopting the same allocation strategy in both periods, always adopting the optimal allocation strategy given by (7). Thus, following (8),

the probability of re-election is $\tau(\pi, \mathbf{a}) = \frac{1}{2}$.

When voters do observe first period public goods, $z \in \{0, q\}$, they are able to update their beliefs on the basis of observed outputs. However, at different levels of bureaucratic quality, the signal offered by the realization of public goods differs in its informativeness. Politicians choose their first period allocation behavior on the basis of the efficiency with which a public good could be produced combined with their anticipated prospects for re-election. Define four cutpoints in q that are relevant for characterizing the resultant equilibria: $q_1 \equiv \frac{1}{\theta}$, $q_2 \equiv \max\{\frac{1}{\theta}, \frac{2b(1-\pi\bar{\theta})}{\theta(2b(1-\pi\bar{\theta})+\bar{\theta}p(1-\pi))}\}$, $q_3 \equiv \frac{2b(\underline{\theta}(\pi-1)-\bar{\theta}\pi)(1+\underline{\theta}(\pi-1)-\bar{\theta}\pi)}{\underline{\theta}(2b(\underline{\theta}(\pi-1)-\bar{\theta}\pi)(1+\underline{\theta}(\pi-1)-\bar{\theta}\pi)+\bar{\theta}(\bar{\theta}-\underline{\theta})p(\pi-1)\pi)}$, and $q_4 \equiv \frac{1}{\bar{\theta}}$. It is straightforward to see that $q_1 \leq q_2 \leq q_3 \leq q_4$ and that $q_1 < q_4$ under the parametric assumptions of the model.

Proposition 1. *Equilibrium* *In the unique perfect Bayesian equilibrium:*

- (i) *If $q < q_1$, both types of politicians allocate $a_1 = a_2 = 0$ to public goods.*
- (ii) *If $q \in [q_1, q_2)$, a competent-type politician allocates $a_1 = a_2 = 1$ while an incompetent-type politician allocates $a_1 = a_2 = 0$ to public goods.*
- (iii) *If $q \in [q_2, q_3)$, competent-type politician allocates $a_1 = a_2 = 1$ while an incompetent-type politician allocates $a_1 = 1$ with probability $k \in (0, 1)$, $a_1 = 0$ with probability $1 - k$, and $a_2 = 0$ to public goods.*
- (iv) *If $q \in [q_3, q_4)$, a competent-type politician allocates $a_1 = a_2 = 1$ while an incompetent-type politician allocates $a_1 = 1$ and $a_2 = 0$ to public goods.*
- (v) *If $q \geq q_4$, both types of politicians allocate $a_1 = a_2 = 1$ to public goods.*

Figure 1 depicts both politician allocations and voter beliefs, in each equilibrium region. Importantly we see that in multiple regions of the parameter space, competent and incompetent politicians pool (or partially pool) by making the same first-period allocation. This allocation behavior by politicians, however, determines how much voters can learn by observing public goods.

When bureaucratic quality is low ($q < q_1$), neither type of politician invests in public goods. If the voter observes the signal that no public goods were produced ($z = 0$), they do not learn anything because both politicians are making the same allocation. This produces the same outcomes—in terms of voter beliefs and behavior—that we would expect if voters never saw the signal ($z = \emptyset$).

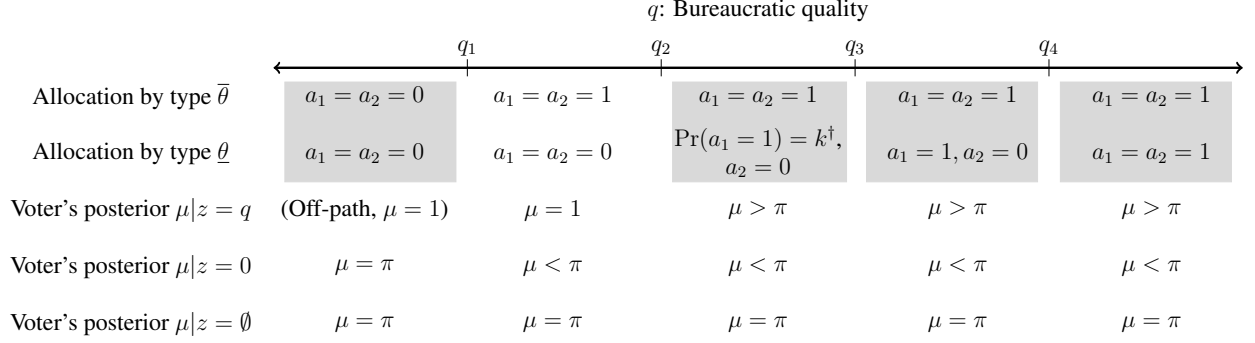


Figure 1: Equilibrium allocation behavior by politicians and the voter's posterior belief as a function of bureaucratic quality, q . The shaded rectangles represent pooling or partially pooling equilibria in which both types of politicians make the same allocation in the first period with positive probability.

[†]In the partially pooling equilibrium, the incompetent type allocates $a_1 = 1$ with probability k and $a_1 = 0$ with probability $1 - k$. The appendix provides an explicit derivation of k .

Indeed, in this pooling equilibrium, voters do not care about the politician's type because neither type will allocate funds to or produce the public good.

In the remaining equilibria, the voter can learn if they observe the public goods signal, as is apparent in Figure 1. As is standard, when the voter sees that public goods were produced ($z = q$), they update positively such that $\mu > \pi$. When the voter sees no public goods produced ($z = 0$), they update negatively such that $\mu < \pi$. Because re-election rates are increasing in the voter's posterior belief that the incumbent is competent, this learning means the voter re-elects competent types at a strictly higher rate than incompetent types (see Lemma A1). This generates positive selection of second-period incumbents. Voters stand to learn more in the separating equilibrium than the partially-pooling or pooling equilibria that emerge at higher levels of bureaucratic quality.⁴ In this separating equilibrium, an observation that public goods were produced ($z = q$) reveals the incumbent to be a competent type. Importantly, however, even in the two pooling equilibria in which both types of politician allocate first-period funds to public goods, voters can update on the politician's type. Here, the complementarity between politician competence and bureaucratic quality allows voters to update on the basis of observed public goods.

⁴The pooling equilibrium that emerges for $q \geq q_4$ would not emerge with a venal/non-venal type space (as opposed to competence). With such a type space, it is possible to generate equilibria that are qualitatively similar to the three equilibria that emerge at any $q < q_4$.

Three attributes of the partially pooling and pooling equilibria that emerge in the region $q \in [q_2, q_4)$ are important. First, note that this equilibrium only emerges when the voter could be informed (when $p > 0$).⁵ If the politician knows the voter will not observe public goods, the incompetent type has no incentive to pool with the competent type in the first period because it will not improve their re-election chances. Second, we should only observe an *increase* in shirking by the politician in the second period in these equilibria. In all other equilibria, each type of politician makes the same public goods allocations in each period. This term effect on shirking is clearly driven by incompetent types who are re-elected. Third, this equilibrium features an incumbency *disadvantage*. Due to this second-period shirking, it is quite costly to re-elect an incompetent type for a second term. Because both types of politicians will allocate public goods in the first period, the voter is more likely to elect the challenger. This explanation for incumbency disadvantage is analogous to the mechanism in Klačnja and Titunik (2017), though my theory introduces new scope conditions—with respect to bureaucratic quality—on where we should observe it.

2.1 Extension: Observable Politician Behavior

In many experiments, researchers provide voters with information about a politician's actions—here, their allocation behavior—rather than information about public goods outputs. It is therefore useful to consider an extension of the model in which voters observe politician allocation behavior, a_1 , instead of public goods outputs, g_1 . Voter observation of politician allocation behavior instead of public goods outcomes removes a source of randomness—whether or not public goods are successfully produced by the bureaucrat. I characterize the Perfect Bayesian equilibria of this game in Proposition A1 and plot equilibrium allocation behavior and voter beliefs in Figure 2. I denote the thresholds in bureaucratic quality in this extension as $\tilde{q}_1 = \frac{1}{\theta}$, $\tilde{q}_2 = \max\{\frac{1}{\theta}, \frac{2b}{\theta 2b + p\theta}\}$, $\tilde{q}_3 = \frac{2b}{\theta 2b + p\theta\pi}$, and $\tilde{q}_4 = \frac{1}{\theta}$.

While the four equilibria are similar to those in Proposition 1 and Figure 1, the differences are informative. First, note that $\tilde{q}_2 \leq q_2$. When voters observe politician actions—not outputs—the separating equilibrium is sustained for a smaller region of the bureaucratic quality parameter

⁵This is apparent from inspection of q_2 , q_3 , and q_4 . When $p = 0$, $q_2 = q_3 = q_4$.

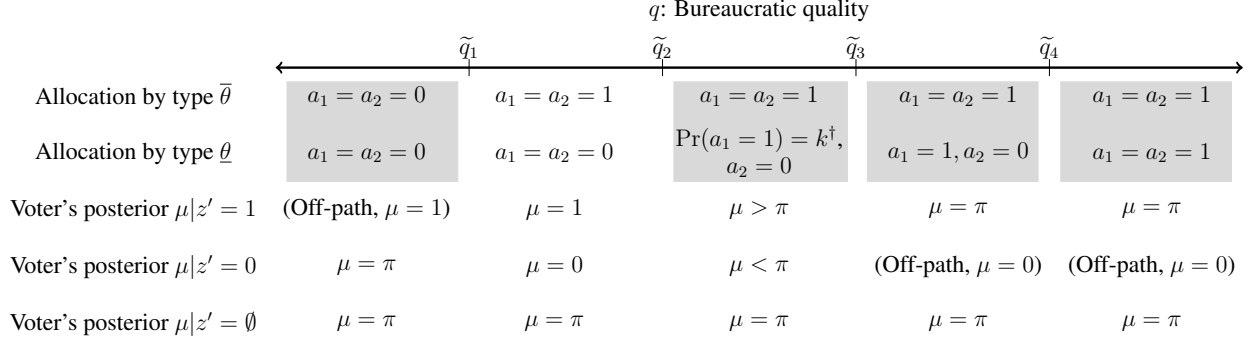


Figure 2: Equilibrium allocation behavior by politicians and the voter's posterior belief as a function of bureaucratic quality, q . The shaded rectangles represent pooling equilibria in which both types of politicians make the same allocation in the first period. When $q < q_L$ or $q \geq q_H$, both types also make the same allocation in the second period.

space. Consistent with the findings of Raffler and Martin (2019), in the separating equilibrium, co-production with the bureaucrat adds noise to the signal. As a consequence, the voter learns more in a separating equilibrium when they observe politician actions, not public goods. But this also makes this equilibrium “rarer” because it incentivizes the incompetent type to pool with the competent type to increase their likelihood of re-election at lower levels of bureaucratic quality.

While this separating equilibrium ($q \in [\tilde{q}_1, \tilde{q}_2)$) is consistent with the results in Raffler and Martin (2019), the pooling equilibria that emerge for $q \geq \tilde{q}_3$ are very different. In these equilibria, both politicians pool and allocate $a_1 = 1$ to public goods. In this case, the voter cannot learn anything from observing the politician's action because both types make identical allocations. Here, we would expect that information that the politician allocates funds to public goods to have no effect. In contrast, in these regions, the voter would be better off observing public goods outputs. Here, the complementarity between politician competence and bureaucratic quality ensures that competent politicians deliver the public goods at higher rates than incompetent politicians. As such, in contrast to the findings in Raffler and Martin (2019), for sufficiently high bureaucratic quality, co-production of public goods between politicians and bureaucrats *facilitates* rather than hinders electoral accountability.

3 Study #1: Accountability of Brazilian Mayors

I present empirical evidence in support of this theory in the form of two theoretically-structured meta-studies. This design seeks to synthesize the findings of multiple studies that measure different implications of a common equilibrium (resp. equilibria). I use the theory to generate predictions about different measures of politician allocation behavior, voter beliefs, and voter behavior. In the first study, I consider the case of municipal governance in Brazil, which is arguably the most researched case in the study of electoral accountability of local politicians. Since the work of Ferraz and Finan (2008), researchers have studied mayoral corruption and voters' responses to revelations of corruption or lack thereof (Ferraz and Finan, 2011; Winters and Weitz-Shapiro, 2016; Weitz-Shapiro and Winters, 2016). These articles use the results—real or hypothetical—of federal or state audits of municipalities to understand politician behavior and voter responses. In separate work, Klašnja and Titunik (2017) find evidence of a strong incumbency disadvantage in Brazilian mayoral elections. My theory helps to connect these sets of findings. To do so, I leverage a new measure of bureaucratic quality to extend these existing analyses.

I measure bureaucratic quality at the municipal level. However I do not manipulate bureaucratic quality, nor do I identify a research design which generates plausibly exogenous variation in this variable. Indeed, historical evidence emphasize stability in bureaucratic quality over the last century (Besley et al., 2021). Theories of civil service adoption similarly point to the rarity of these reforms (Geddes, 1994; Ting et al., 2012; Huber and Ting, 2020). As such, I rely on observational variation in bureaucratic quality in a series of analyses intended to test several implications of the theory.

3.1 Measuring Bureaucratic Quality

As the theory clarifies, bureaucratic quality (q) is distinct from bureaucratic effort (e) and public goods outputs (g). The measure of bureaucratic quality should therefore abstract from effort or public goods outputs. I thus operationalize quality as a measure of human capital of individuals employed in municipal administration. I rely on Brazil's Basic Municipal Information Survey

(MUNIC) to measure characteristics of employment in (direct) municipal administration. This survey, implemented by the Instituto Brasileiro de Geografia e Estatística, requires municipalities to report counts of public employees working in direct municipal administration, disaggregated according to several categories including education and contract type. Given that the raw data consists of counts of public employees, the level of cross-sectional aggregation is the municipality.

I operationalize bureaucratic quality as the average education level of bureaucrats working in municipal administration. The measure of bureaucratic quality capture features of a representative (average) public employee. I abstract from measures of the number of public employees per capita for two reasons. First, per-capita measures of municipal employment do not account for efficiencies of scale: running a fixed set of programs requires more employees per capita in small municipalities. Second, classic descriptions of patronage in Latin America include accounts of low-wage workers filling out the ranks of public employment (Calvo and Murillo, 2004; Grindle, 2012). This is generally believed to reduce bureaucratic quality.

In treating bureaucratic education as a measure of quality, several legal, economic, and political considerations are warranted. Legally, municipal employees in direct administration should be hired with civil service provisions, though empirically adherence varies substantially, with many municipalities relying heavily on contractors. Variation in hiring practices is substantial across municipal governments, which accords with the wide observed variation in bureaucratic quality (e.g., Toral, 2019).

Average bureaucratic education clearly is driven, in part, by local labor market conditions. The scope of heterogeneity across Brazilian labor markets is likewise impressive. I account for regional variation using state fixed effects. I also use flexible covariate specifications to adjust for municipal population, average municipal education (years of education), formality (percent of workers working in the formal sector), and per-capita GDP. All of these features correlate with the measure of bureaucratic quality (see Figure A2). However, they (collectively) account for less than 20% of the variation in bureaucratic quality, indicating that this variation is not simply a function of variation in local labor markets (see Figure A3).

Analysis of persistence of this measure of bureaucratic education within municipalities over five waves of MUNIC (2005, 2008, 2011, 2014, and 2018) accords with some qualitative assumptions of the model. While quality is secularly increasing over time across the sample of (Figure A1), within municipalities bureaucratic quality is sticky. The annualized autocorrelation of bureaucratic quality between waves of the survey is 0.84 (Table A3). I further show that such persistence obtains across the constituent education categories. In the model, q is treated as exogenous and the politician does not alter bureaucratic quality. The data is consistent with this assumption. Table A4 reports the results of first-difference models that regress changes in bureaucratic quality (from consecutive waves of MUNIC) on indicators measuring a change in mayor and change in the mayor's party in an intervening election. I find no consistent evidence that changes in mayor or mayor's party yield differential shifts in average bureaucratic quality: point estimates are near-zero and precisely estimated. Further evidence from visualization of the ECDF of changes in bureaucratic quality shows no evidence of differential changes in variance (Figure A4).

Finally, given the importance of distinguishing the implications of the present theory of bureaucratic quality and accountability from theories premised only on voter information, I examine the association between bureaucratic quality and local media presence in Figure A5 and Table A5. In the Brazilian municipal context, the presence of community radio is argued to be the most important form of media for diffusing local news (Ferraz and Finan, 2008; Boas and Hidalgo, 2011; Varjão, 2019). I show that while raw measures of bureaucratic quality and radio station presence are positively correlated, conditional on the local labor market covariates and state fixed effects, the conditional association between bureaucratic quality and radio presence is estimated to be a precise zero. Further, I use an indicator for radio presence as an additional covariate to allay possible concerns.

3.2 Measuring Allocations to Public Goods

In this analysis, I measure politicians' allocations to rents, $1 - a$, relying on the results of federal audits of municipal governments. Such audits have gained prominence across Latin America in the last two decades and have been used to measure corruption in academic literature (Ferraz and

Finan, 2008; Chong et al., 2015). In Brazil, the audits I use to measure allocations are conducted by the federal Controladoria-Geral da União (CGU), through a municipal auditing program inaugurated in 2003. Because municipalities in Brazil receive the majority of their budgets from the federal government, such audits cover sizable shares of municipal budgets.⁶ Audits consist of visits by a team of federal officials to municipalities to oversee allocation and disbursement of funds and observe outputs. They report their findings in reports which are disseminated by local media (Ferraz and Finan, 2008).

Brazil's audits are oft-studied as a natural experiment because the federal government randomly selects municipalities by lottery. While this random assignment has facilitated many studies of the effects of audits (i.e., Ferraz and Finan, 2008; Avis, Ferraz, and Finan, 2018), from the perspective of this paper, it primarily ensures random sampling of municipalities. This sampling ensures support across all levels of bureaucratic quality. Support across the distribution of bureaucratic quality is essential to making inferences about the theory proposed by this paper, as I describe in more detail below.

A final consideration about the use of audit data to measure allocations considers whether audits measure the actions of politicians, here, mayors, or the municipal administration generally. I follow existing studies of accountability in attributing corrupt or malfeasant spending to mayors. This is precisely the inference that experimental studies of accountability ask voters to make (Boas, Hidalgo, and Melo, 2019; Chong et al., 2015; Arias et al., 2019). In Brazilian municipalities, executives are responsible for proposing a budget and, alongside the city council, monitoring its execution (Gonçalves, 2013). It is not surprising, therefore, that opposition to audits in the form of lawsuits have come from elected politicians, not other municipal officials (Seabra, 2018). Furthermore, studies of the effects of audits find few consequences for bureaucrats, at least in terms of retention (Ferrali and Kim, 2019).

⁶Among the sample of municipalities audited in the first rounds of randomized audits, audits covered 60% of local budgets.

3.3 Measuring Voter Updating and Behavior

Finally, I measure voter updating and voting behavior in response to provision of information about politician allocation behavior. The study of information revelation and voter updating has spawned a large body of recent survey experiments and experiments (for a list of these studies, see Incerti, 2019; Dunning et al., 2019; Bhandari, Larreguy, and Marshall, 2019). I extend a survey experiment fielded in Brazil as reported in Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016) to measure how voter responses to evidence of politician corruption/lack of corruption vary in bureaucratic quality. This allows for the cleanest test of the voter updating mechanism in isolation, separated from the effect of a politician's (strategic) allocation behavior or baseline voter information. It is important to emphasize that the manipulation to voter information means that voters observe politician allocations (more specifically, corrupt spending), not public goods. As such, it tests the extension of the model described in Figure 2 and Proposition A1. By design, the hypothetical mayor did not allocate funds in anticipation of this revelation. As such, this test represents a partial equilibrium test of the model I advance.

To study equilibrium voter behavior, I consider the phenomenon of incumbency disadvantage to test whether voters anticipate the second-term behavior of mayors and condition their votes accordingly. Klačnja and Titunik (2017) find evidence of an incumbency disadvantage for Brazilian mayors' parties. Using an electoral regression discontinuity design, they find that barely elected mayors' parties are less likely to win the subsequent election than barely defeated parties. My model suggests that we should only observe this incumbency disadvantage in the region of bureaucratic quality in which the incumbent shirks in the second term but not the first. This equilibrium emerges only at intermediate levels of bureaucratic quality. I extend the analysis of Klačnja and Titunik (2017) to examine heterogeneity in the local average treatment effect (LATE) of incumbency on (re)-election as a function of bureaucratic quality. This evidence allows me to investigate whether voters re-elect incumbents at rates consistent with theoretical predictions.

3.4 Mapping Theory to Research Design

Empirically, I focus on the beliefs and actions of two actors in the model: the politician and the voter, examining four predictions of the model. In assessing the plausibility of this model, I compare it to two special cases which capture arguments in the existing empirical literature:

1. *No bureaucratic co-production*: Assume that $\underline{\theta} = 0$ and $\bar{\theta} = 1$.⁷ This eliminates both pooling equilibria and ensures that public goods production is independent of bureaucratic quality. If voters observe public goods, g_1 , the incumbent's type is revealed. Per Fearon (1999), this is a model of pure selection.
2. *Completely uninformed voters*: Assume that $p = 0$. The voter does not observe a signal and thus does not update. This breaks the link between first-period allocations and electoral fortunes, implying that any politician that could be induced to allocate first-period funds to public goods in order to improve her electoral prospects if voters were informed would not do so.

The first case is obviously rather narrow. Here, competent (or “high”) types always produce public goods but incompetent (or “low”) types never do. Voters seek to elect competent types since it is their only shot at receiving public goods. This is consistent with the motivation of many studies based in low- and middle-income democracies including the studies from Brazil that I extend. Note further that cases #1 and #2 are not mutually exclusive. Indeed, the arguments in which bureaucrats are absent often coincide with arguments about limited voter information.

To adjudicate the general (“unrestricted”) case of the model from these alternatives, I motivate four reduced-form tests in Table 1. The tests I propose rely on the observable implications of the equilibria that emerge at different levels of bureaucratic quality. This presents a challenge because we do not know the true mapping between q and the empirical measure of bureaucratic quality. For example, it is possible that all municipalities have sufficiently uniform bureaucratic

⁷Note that this model falls outside parametric restrictions imposed on $\underline{\theta}$ and $\bar{\theta}$ in the model, so this is not precisely a “case” of the general model.

quality to ensure that all municipalities fall into one of the equilibria identified in Proposition 1. The inferences from the proposed tests rely upon variation in the underlying equilibria. To this end, I adopt three complementary approaches to interpretation of the data. First, description of the data and reading of the existing literature provides some guidance on what is plausible. Second, I leverage insights *across* the tests described in Table 1. Because these data come from different sources, qualitative consistency across tests is, in principle, more challenging to achieve and also more informative. Finally and most practically, I am careful to model bureaucratic quality flexibly in regression specifications given the possibility of non-linearities suggested by the model.

3.5 Estimation

Table 1 presents four testable implications of the theory. In the first test, I examine the association between politician allocations to rents (corruption) and municipal bureaucratic quality. To examine this relationship, I estimate an OLS regression of the form:

$$Y_{msl} = \beta_0 + \beta_1 Q_m + \gamma_s + \lambda_l + \delta \mathbf{X}_m + \epsilon_{msl} \quad (9)$$

where Y_{msl} is the proportion of audited funds allocated to rents in municipality m in state s , as measured in lottery round l . Q_m is the measure of municipal bureaucratic quality and β_1 is the coefficient of interest. γ_s is a vector of state fixed effects and λ_l is a vector of lottery round fixed effects. \mathbf{X}_m is a matrix of decile indicators for each of four municipal-level covariates: population, average education, formality rate, and GDP per capita, as well as an indicator for community radio presence. The unrestricted model implies that $\beta_1 < 0$: corruption declines in bureaucratic quality. I also estimate this specification with tercile and quartile bins of Q_m given the potential for non-linearity suggested by the model.

Second, I examine whether term differences in the allocation of rents varies in bureaucratic quality. I test this prediction by estimating an OLS regression of the form:

	TESTABLE IMPLICATIONS		EVIDENCE
	“Unrestricted” model $p > 0, 0 < \underline{\theta} < \bar{\theta} < 1$	No bureaucratic co-production $\underline{\theta} = 0, \bar{\theta} = 1$ No voter information $p = 0$	
1	<p><i>Politician allocations to rents (1 – a) weakly decrease in bureaucratic quality (q).</i></p>	<p>Allocations to rents are independent of bureaucratic quality (q).</p>	<p>Original analysis using measure of rents from Avis, Ferraz, and Finan (2018).</p>
2	<p><i>Politicians allocate more or less to rents in their second term ($t = 2$) than their first term ($t = 1$). This difference is attenuated to zero at low and high levels of bureaucratic quality.</i></p>	<p>Politicians allocate less to rents in their second term than in their first term. This difference is independent of bureaucratic quality.</p>	<p>Extension of (Ferraz and Finan, 2011).</p>
3	<p><i>At high levels of bureaucratic quality, a voter’s posterior belief (μ) is equivalent to her prior (π) upon receiving a signal that a politician allocated no funds to rents ($a = 1$).</i></p>	<p>Voters update positively ($\mu > \pi$) in response to a signal that a politician allocated no funds to rents at any level of bureaucratic quality.</p>	<p>Extension of survey experiment reported in Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016).</p>
4	<p><i>Incumbency disadvantage does not emerge at low or high levels of bureaucratic quality (q).</i></p>	<p>Incumbency disadvantage does not emerge.</p>	<p>Extension of Klášnja and Titunik (2017).</p>

Table 1: Theoretical implications of the model (the “unrestricted” model) relative to the two special cases. Note that “–” indicates that the case does not make predictions that differ from the unrestricted model. *For the “no voter information” model, the interpretation in the text is that voters do not observe the signal; the prediction in the cell represents a slightly different interpretation that yields the same equilibrium. Formal motivation of these predictions is included in the Appendix.

$$Y_{msl} = \beta_0 + \beta_1 Q_m + \beta_2 \text{Second term}_m + \beta_3 Q_m \text{Second term}_m + \gamma_s + \lambda_l + \delta \mathbf{X}_m + \epsilon_{msl} \quad (10)$$

where Y_{msl} is the share of audited funds allocated to rents in municipality m in state s , as measured in lottery round l . Q_m is the measure of municipal bureaucratic quality and Second term_m captures whether the politician is in her second term in the 2001-2004 term (when the audits occurred). Bureaucratic quality, Q_m , is modeled linearly and in quantile bins in different specifications. γ_s is a vector of state fixed effects and λ_l is a vector of lottery round fixed effects. \mathbf{X}_m includes the same set of flexible controls for local labor markets. The quantity of interest is $\hat{\beta}_1 + \hat{\beta}_3 Q_m$, the marginal effect of being a second term mayor, at a given level of bureaucratic quality. I also use a regression discontinuity-like design to decompose selection from shirking effects, enumerated at greater length in Appendix A4.

Third, using survey experimental data, I estimate how the magnitude of voter updating in response to revelation of a clean mayor varies in bureaucratic quality. Given relevant arms of the experimental design of Winters and Weitz-Shapiro (2016), I estimate conditional average treatment effects (CATEs) of the clean and corrupt signals on updating, using OLS regressions of the form:

$$Y_{ims} = \beta_0 + \beta_1 + \beta_2 \text{Clean signal}_i + \beta_3 \text{Clean signal}_i Q_m + \beta_4 \text{Corrupt signal}_i + \beta_5 \text{Corrupt signal}_i Q_m + \gamma_s + \theta \mathbf{X}_m + \epsilon_{ims} \quad (11)$$

where Y_{ims} is the survey response of individual i in municipality θ in state s . The signal indicators measure the two treatment arms in the survey experiment. I cluster standard errors at the level at which Q_m is measured: the municipality. The estimator of the relevant CATE is $\hat{\beta}_2 + \hat{\beta}_3 Q_m$, which measures updating on the clean signal at a given level of bureaucratic quality. As in all models, I also consider binned quantile indicators of the bureaucratic quality measure.

Finally, I examine rates of incumbency disadvantage to see whether voters vote in a manner

consistent with theoretical predictions. Following Klašnja and Titunik (2017), I adopt a close elections regression discontinuity (RD) design. I extend their analysis by estimating the local average treatment effect of incumbency at different levels of bureaucratic quality. Specifically, I estimate the conditional LATE given by:

$$\tau_q = \lim_{x \downarrow c} E[Y_m | Q_m = q] - \lim_{x \uparrow c} E[Y_m | Q_m = q] \quad (12)$$

at different measured levels of bureaucratic quality, q . I use the robust, bias-corrected Calonico, Cattaneo, and Titiunik (2014) regression discontinuity estimator, fit separately on each (quantile) bin of bureaucratic quality. I maintain the optimal bandwidth from the pooled sample for each subset. The “unconditional” regression discontinuity estimand presented in Klašnja and Titunik (2017) can be decomposed into a candidate or party’s decision on whether to contest election $t + 1$ and citizens’ votes given the choices on the ballot (Slough, 2022). To this end, I present unconditional LATE estimates in addition to the LATE on party election contestation and a post-treatment estimand that examines vote choice conditional on the incumbent running. While the latter estimand is *not* a causal estimand, it enables us to better isolate voter (as opposed to candidate) behavior. I formalize the decomposition of the unconditional LATE and motivate these estimands in Appendix A6.1.

4 Study I: Results

I present four findings that parallel the implications of the model elaborated in Table 1 to assess the plausibility of the theoretical model of accountability that I advance.

4.1 Politician Allocations to Rents Decrease in Bureaucratic Quality

In a first observational test of the theory, I examine the relationship between bureaucratic quality and funds diverted to rents, as measured by municipal audits in 2003. Table 2, Column 1, presents the bivariate regression of the share of funds allocated to rents ($1 - a$) to rents on the bureaucratic quality measure. A one standard deviation increase in bureaucratic quality (context) reduces rents

	Share of corrupt spending			Log(Share of corrupt spending + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
A. LINEAR BUREAUCRATIC QUALITY MEASURE (Z-SCORE)						
Bureaucratic quality	-0.014**	-0.014**	-0.017**	-0.012**	-0.012**	-0.014**
	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)	(0.006)
B. BUREAUCRATIC QUALITY MEASURE TERCILES (RELATIVE TO FIRST TERCILE)						
Bureaucratic Quality, Tercile 2	-0.009	-0.009	-0.009	-0.007	-0.007	-0.007
	(0.012)	(0.012)	(0.012)	(0.010)	(0.010)	(0.010)
Bureaucratic Quality, Tercile 3	-0.027**	-0.026*	-0.036**	-0.023**	-0.022*	-0.029**
	(0.012)	(0.014)	(0.018)	(0.010)	(0.011)	(0.014)
C. BUREAUCRATIC QUALITY MEASURE QUARTILE (RELATIVE TO FIRST QUARTILE)						
Bureaucratic Quality, Quartile 2	-0.009	-0.003	-0.002	-0.006	-0.001	0.000
	(0.015)	(0.015)	(0.015)	(0.012)	(0.013)	(0.012)
Bureaucratic Quality, Quartile 3	-0.019	-0.021	-0.029*	-0.015	-0.018	-0.024*
	(0.015)	(0.014)	(0.015)	(0.012)	(0.012)	(0.013)
Bureaucratic Quality, Quartile 4	-0.029**	-0.030*	-0.042**	-0.025**	-0.025*	-0.034**
	(0.014)	(0.016)	(0.021)	(0.012)	(0.013)	(0.017)
State FE		✓	✓		✓	✓
Lottery FE		✓	✓		✓	✓
Demographic controls (decile bins)			✓			✓
Community radio indicator			✓			✓
Outcome Range	[0,0.794]	[0,0.794]	[0,0.794]	[0,0.584]	[0,0.584]	[0,0.584]
Outcome Mean	0.062	0.062	0.062	0.056	0.056	0.056
Outcome Std. Dev.	0.10	0.10	0.10	0.085	0.085	0.085
Num. obs.	448	448	448	448	448	448

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2: Association between bureaucratic quality, q , and allocations to public goods, a . Funds diverted from public goods are measured as the share of corrupt spending, as defined by Avis, Ferraz, and Finan (2018). Heteroskedasticity-robust standard errors in parentheses.

by 1.4 percentage points or 22.5% of the sample mean. The conditional association of bureaucratic quality and mayors' allocations to rents remains substantively similar when including covariates. Given the right-skewed distribution of the share of corrupt spending, I also look at the logged outcome, with substantively similar findings in Columns 4-6. Consistent with the model, more flexible specifications of bureaucratic quality in Panels B and C reveal no evidence of non-monotonicity in this relationship. Corrupt spending is concentrated in the lowest quantiles of bureaucratic quality. On average, 7.3% of federal funds in the lowest tercile of municipalities and 7.5% of such funds in the lowest quartile of municipalities are spent in a corrupt manner. These rates drop in higher quantiles of bureaucratic quality.

Further, Figure 3 plots the density of the corruption (rents) outcome which suggests that the

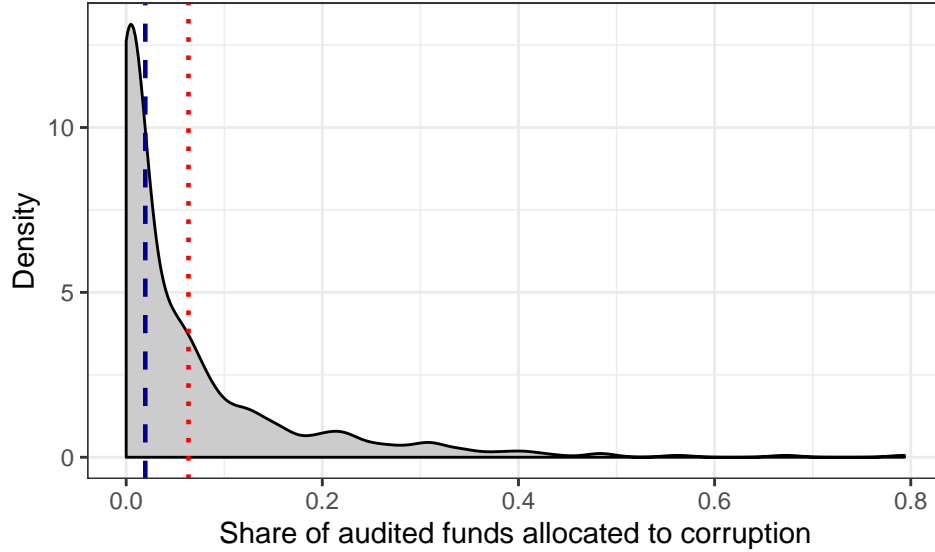


Figure 3: Distribution of the measure of rents: corrupt spending as a share of total audited federal funds. The navy dashed line gives the median and the red dotted line gives the mean.

modal politician did not divert any of the audited federal funds. Indeed, the median politician only allocates 1.9% of these funds to rents. This is consistent with a setting with circumscribed corruption. Per the model, circumscribed corruption occurs when bureaucratic quality is high enough to induce at least one type of politician to invest in public goods (the separating equilibria or the pooling equilibria where both types invest). I return to this observation in the subsequent analyses.

The finding that rent extraction declines in bureaucratic quality is consistent with the unrestricted model advanced here, regardless of whether voters observe signals of an incumbent's performance. It is inconsistent with the special case of the model in which bureaucratic co-production of public goods is absent.

4.2 Second-Term Shirking Disappears as Bureaucratic Quality Increases

Recall that in the model, the partially-pooling and pooling equilibria in which politicians allocate more to public goods in the first period but not the second emerge under two conditions. First, intermediate bureaucratic quality makes public goods production by the incompetent type of politician inefficient while public goods production is efficient for the competent type. Second, voters

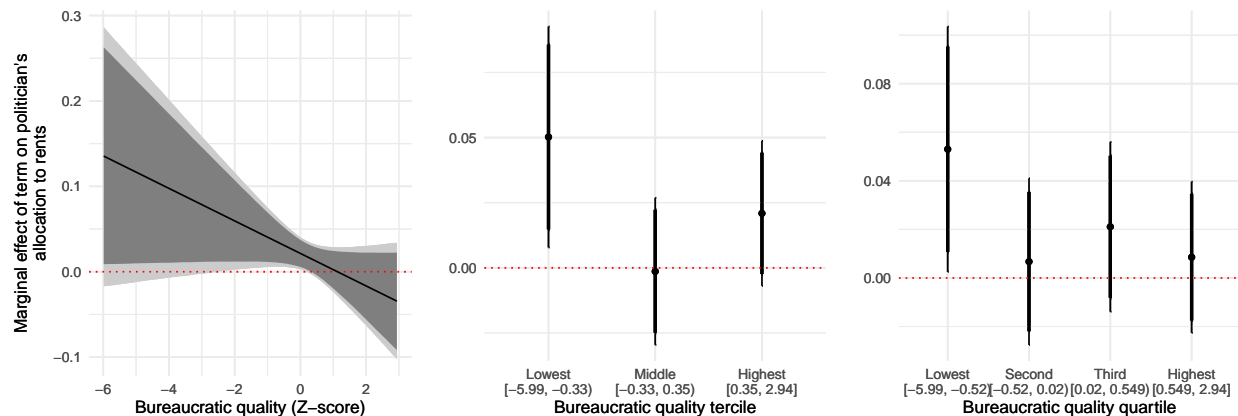


Figure 4: The marginal effect of a second-term politician versus a first-term politician on proportion of audited funds allocated to rents. 90% and 95% confidence intervals constructed on heteroskedasticity-robust standard errors.

are likely enough to observe the signal that costly diversion of funds to public goods in the first term can improve re-election prospects. Ferraz and Finan (2011) establish that corruption is, on average, higher among second-term mayors than first-term mayors. This finding contradicts the predictions of the model without bureaucratic co-production, in which differences in allocations by term are driven only by positive selection in the re-election of mayors. I further examine how this difference between (term-limited) second term and (non-term limited) first term mayors varies in bureaucratic quality. Figure 4 shows that second-term shirking—or diversion of funds from public goods—manifests strongly in municipalities with low bureaucratic quality. However, this difference by term disappears as bureaucratic quality increases.

This finding is consistent with the suggestion that the lowest observed levels of bureaucratic quality correspond to the separating equilibrium and higher levels of bureaucratic quality correspond to the pooling equilibrium with public goods allocations from both types. Most importantly, under the model advanced here, the finding of any difference between first- and second-term mayors suggests that politicians anticipate that voters could learn about their actions (or outcomes). This provides evidence against the case of “no accountability” ($p = 0$). The result in Ferraz and Finan (2011) is sufficient for this conclusion. The finding that this difference varies predictably in bureaucratic quality provides evidence to bolster the plausibility of this model of accountability.

Serving as a first- versus second-term mayor is clearly not randomly assigned. One may be tempted to rely on a regression discontinuity (RD) design to identify the conditional (local) ATE of term. Under the present model, however, a RD cannot identify the effect of term alone since second term mayors are also more likely to be of high competence than first-term mayors in the posited parameter space (consistent with the critique in Marshall, 2019). Yet, under the valence assumptions in the voting model, an RD-like estimator that narrows the bandwidth to close races provides a way to decompose term effects from politician type. By reducing the sample to close races (setting a small bandwidth on the RD), the share of incompetent types in the pool of second term incumbents should be higher and thus we would expect more shirking on average (a larger difference between the first and second terms). By varying the bandwidth in Figure A7, I show that this prediction is indeed borne out in the analysis. Consistent with the model’s predictions, in samples where incompetent types theoretically represent a larger share of second-term mayors (close elections), second period shirking is more pronounced.

This test and these results provide evidence against an alternative (unmodeled) account in which bureaucrats exert substantial power over politicians as in Raffer (2022). In this alternate account, high-quality bureaucrats prevent politicians from engaging in corruption whereas weaker, lower-quality bureaucrats do not. If bureaucrats were indeed constraining the politician’s allocations in this way, we should not observe second-term shirking. This finding therefore provides evidence that is inconsistent with accounts in which bureaucrats are unilaterally driving both allocation decisions and policy outcomes.

4.3 Voters Do Not Update in Response to Revelation of “Clean” Politicians when Bureaucratic Quality is High

I now extend a survey experiment to study voter updating in response to politician performance information at different levels of bureaucratic quality. In recent years, scholars have questioned the correspondence between voter (respondent) responses to information in survey experiments as opposed to field settings (Boas, Hidalgo, and Melo, 2019; Incerti, 2019). I contend that the two measure two measure beliefs and actions, respectively. Per the model, effects on beliefs and

actions should be different. I use the survey experiment to test whether voters update beliefs in a manner consistent with the theory. Replication and extension of the survey experiment fielded in 2013 and elaborated by Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016) thus affords a test of the updating mechanism in isolation.

The survey experiment provides respondents with a common vignette about a first-term mayor of a “different city” seeking re-election. A control condition provides no information about municipal audit outcomes. Additional text conveying a “clean” treatment condition indicates that the mayor was found to not have awarded bribes for city contracts. The “corrupt” treatment condition conveys that the mayor engaged in such bribery. Importantly, these three arms allow to test how citizens update in response to a signal of no allocation to rents (the “clean” condition) versus a signal of allocation to rents (the “corrupt” condition).⁸ I measure the respondents’ prior beliefs about the vignette via the control condition. I measure beliefs on a seven-point Likert scale.⁹

Importantly, the survey experiment was conducted with a nationally-representative survey in 140 municipalities. Because random selection of municipalities was weighted by population, the sample skews slightly toward higher levels of bureaucratic quality, though it maintains support over the distribution of quality, as shown in Figure A8.¹⁰ I estimate conditional average treatment effects (CATEs), by bureaucratic quality, as a measure of updating. The primary prediction of the model that I test holds that at high levels of bureaucratic quality, where competent and incompetent pool and allocate funds to public goods in a first term, voters should not update on the basis of a “clean” signal in control. The model does not offer predictions for the “corrupt” signal across the parameter space. At high levels of bureaucratic quality, diversion of funds from public goods is off the equilibrium path. The corrupt treatment condition thus allows for validation of the off-path belief assumptions that support the equilibria in Proposition A1.

⁸Boas, Hidalgo, and Melo (2019) replicate this survey experiment using nearly identical prompts in the state of Pernambuco during the 2016 municipal elections. However, the survey experimental portion of their study does not include a “clean” treatment condition, so I cannot test this implication in the context of their data.

⁹All findings are substantively similar when using Weitz-Shapiro and Winter’s 2016 preferred “vote intention” outcome, as reported in Figure A10. For the purposes of this analysis of updating, I seek a more direct measure of beliefs.

¹⁰Because the survey was fielded in 2013, I use the 2011 levels of bureaucratic quality as the moderator of interest.

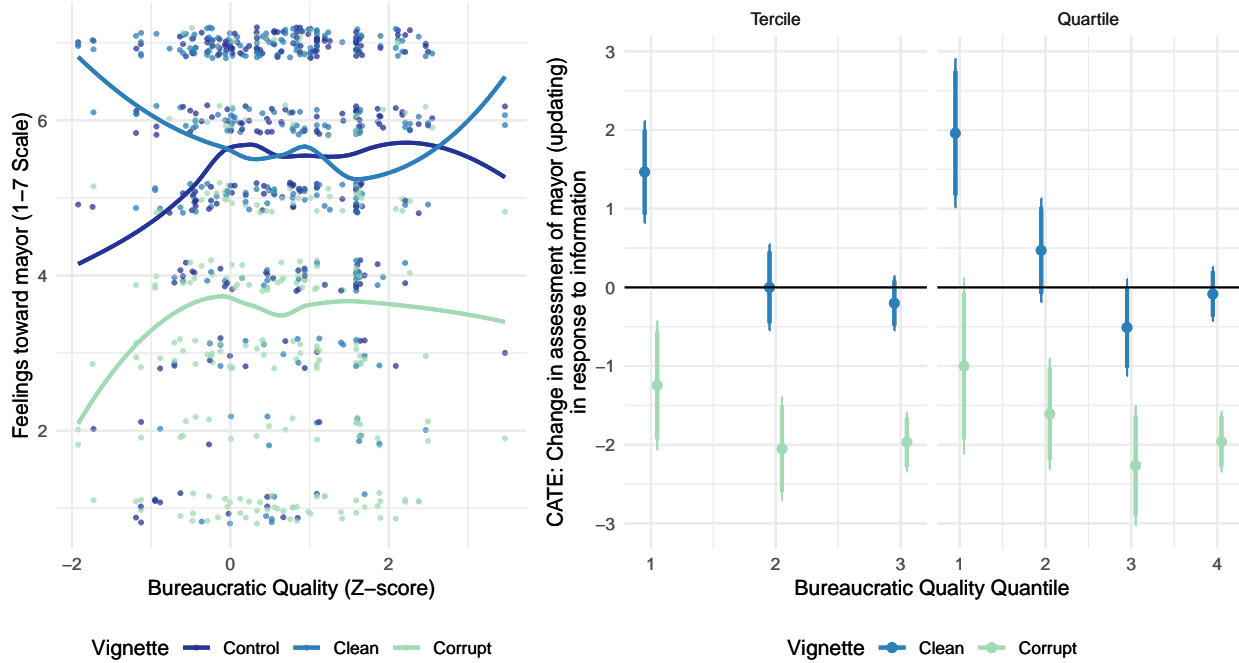


Figure 5: The left panel plots the raw data fitted by Loess for each treatment condition. The right panel estimates the CATEs of both “clean” and “corrupt” signals at different quantiles of bureaucratic quality. All CATEs are estimated from OLS regression models with the labor market covariates. Thick and thin lines represent 90% and 95% confidence intervals. Standard errors are clustered at the municipality, the level of aggregation at which bureaucratic quality moderator is measured.

Figure 5 supports the model’s prediction that voters update on the basis of a clean signal only at moderate levels of bureaucratic quality. In higher quantiles of bureaucratic quality, respondents’ “average” prior (control) and posterior beliefs upon revelation of a clean mayor are not distinguishable.¹¹ Inspection of the data casts doubt on some alternative explanations. First, while the prior belief increases in bureaucratic quality, there is no evidence of a ceiling effect on the CATE, even high levels of bureaucratic quality, per the left panel of Figure 5. Across quantiles, there is “room” to observe positive updating on the basis of a clean signal. Second, while demand effects may be of concern with respect to voter updating about corruption, it is harder to explain why demand effects for the clean treatment would diminish with municipal bureaucratic quality. Finally, in this analysis, confounding is a concern with regard to the difference in CATEs at different

¹¹The finding of substantial updating on politician in low quantiles is distinct from the null ATE across the sample that is reported in Winters and Weitz-Shapiro (2016).

levels of bureaucratic quality. However, this is not the relevant test of the theory; the CATEs are causally identified by the experimental research design (under standard assumptions). Thus, the relevant concern is that another feature that drives bureaucratic quality drives similar patterns of differential updating.

To this end, I do not detect variation in these patterns of updating by individual respondent characteristics. In particular, these patterns do not vary detectably in citizen education or political knowledge on factual questions, as defined by Winters and Weitz-Shapiro (2016) (Figure A11). This helps to allay concerns that differential patterns of updating require sufficient political sophistication or are driven by variation in the educational composition of the electorate that covaries with bureaucratic quality.

The strong treatment effects of the corruption revelation treatment in lower quantiles of the distribution of bureaucratic quality are consistent with predictions of a separating equilibrium. At higher level of bureaucratic quality, the signal provided to voters is “off path”—such voters would not observe a signal about the politician’s action in equilibrium. The negative updating observed in Figure 5 is consistent with the assumption that such a signal is generated by the incompetent type, though note that this tests an assumption rather than a result of the model.

In sum, extension of the survey experiment provides evidence that citizens (voters) respond to information by updating their beliefs in a manner consistent with the theory. Specifically, “good news” of a clean politician record only leads to updating when competent and incompetent types are predicted to take different actions. This suggests that voters internalize expectations about politician behavior and respond to information in light of these expectations. Per the model, differences in expectations are driven by bureaucratic quality.

4.4 Incumbency Disadvantage and Equilibrium Voter Behavior

I consider incumbency disadvantage as an equilibrium implication of the model. Incumbency disadvantage emerges when incompetent-type incumbents allocate funds to public goods in the first period but shirk in the second period. For this reason, voters prefer a first-term incumbent of either type to a second-term incompetent type and vote to re-elect at lower rates. Recall that this

effect emerges only if voters observe first period public goods provision (with positive probability). If this did not occur, a first-period incompetent incumbent could not be induced to pool with the competent type by allocating the budget to public goods.

I report the results on incumbency disadvantage at different levels of bureaucratic quality in Figure 6. The left-most panels in each row of plots present estimates of the *unconditional* incumbency disadvantage reported in Klašnja and Titunik (2017). Two observations are of note. Inconsistent with the “no voter information” and “no bureaucratic co-production models” cases of the model, we observe incumbency disadvantage in all subgroups, as evidenced by the negative and statistically significant LATE estimates. The findings in the previous tests suggests that we should expect incumbency disadvantage to emerge only at low levels of bureaucratic quality in the present sample. While the point estimate is largest in magnitude in this sample, the differences are only marginally statistically significant.¹² Perhaps more concerning, larger incumbency disadvantage emerges at low and high levels of bureaucratic quality (in sample).

The decomposition of the unconditional LATE into the LATE on parties contesting election $t + 1$ and the post-treatment estimands help to consider this tension with the model predictions. The plots in the second column of both panels examine the LATE of incumbency on a party’s decision to contest the mayorship in election $t + 1$. For low levels of bureaucratic quality (in sample), we observe no difference in rates of contesting office in $t + 1$ as a function of incumbency status. These LATEs are both substantially attenuated and statistically distinguishable from the respective unconditional LATEs. In contrast, however for high levels of bureaucratic quality, incumbents ran at lower rates than marginal challengers. This effect is substantively large—12.5 percentage points for the tercile measure and 15.1 percentage points for the quartile measure—both of which are statistically indistinguishable from the relevant unconditional LATE estimate in Column 1. This selection into candidacy is outside the scope of my model, but useful for interpreting incumbency disadvantage estimates.

Finally, the third and fourth two columns report estimates of the post-treatment estimand

¹²For the tercile measure, $p < 0.07$ for the tercile 1 and tercile 2 difference; for the quartile measure, $p < 0.18$ for the quartile 1 and quartile 2 difference and $p < 0.05$ for the quartile 1 and quartile 3 difference.

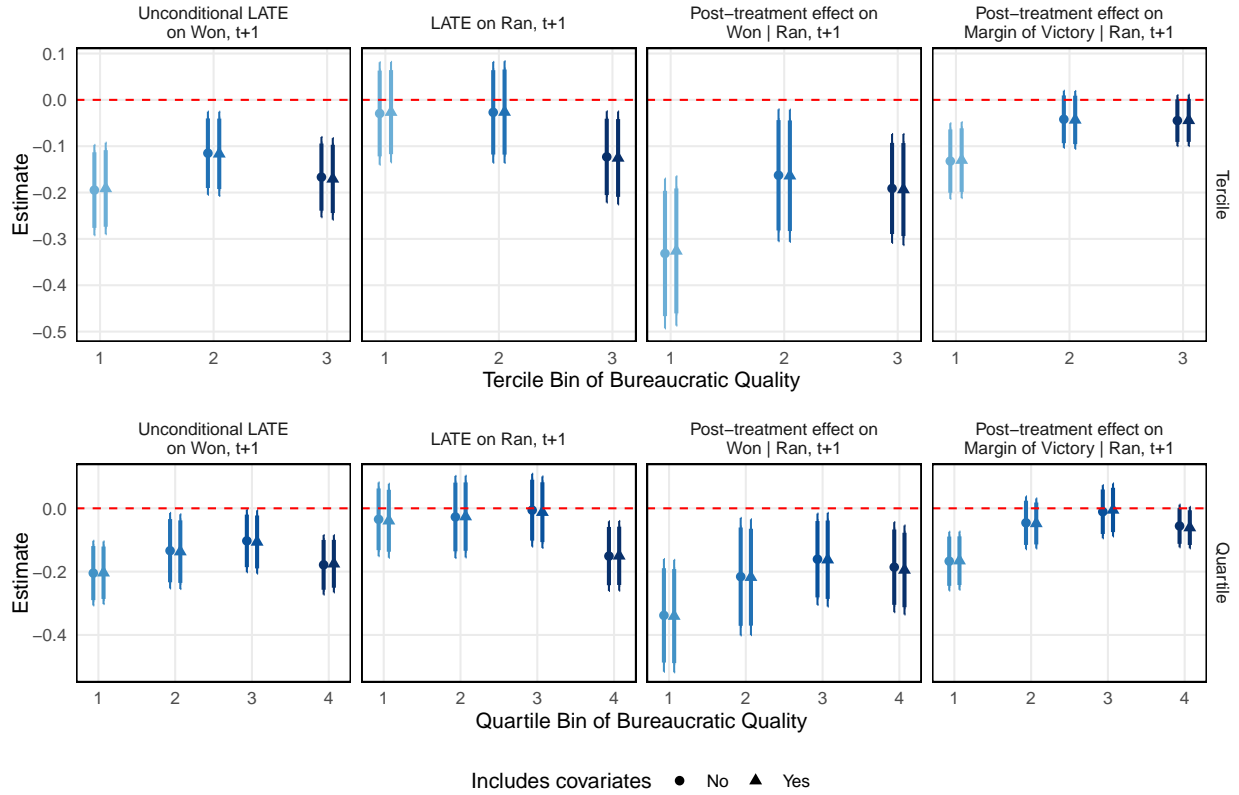


Figure 6: Incumbency disadvantage of Brazilian mayors at different levels of bureaucratic quality.. All estimates are estimated using the Calonico, Cattaneo, and Titiunik (2014) bias-corrected regression discontinuity estimator, with 90% (thin) and 95% (thick) confidence intervals constructed on bias-corrected standard errors. The margin of victory outcome in the fourth column of plots is rescaled to a $[-1, 1]$ range.

(sometimes called the conditional LATE on incumbency disadvantage). Here, the results are more consistent with model predictions. Voters support incumbent candidates at lower rates at low levels of bureaucratic quality (within the sample). These differences are statistically significant at conventional thresholds and evident when examining both vote margins and a binary indicator for winning the election at time $t + 1$. Thus, conditional on voters being given the choice to vote for an incumbent or challenger, voters support the incumbent's bid for re-election at lower rates in municipalities with lower bureaucratic quality.

The evidence on incumbency advantage is somewhat more ambiguous than the other tests. Specifically, incumbency advantage arises across all levels of bureaucratic quality. Nevertheless, the conditional incumbency advantage results suggest that when the option is available to voters, voters vote against incumbents at higher rates in regions of bureaucratic quality consistent with second-term politician shirking. This increases the magnitude of the incumbency disadvantage (in both the unconditional and conditional estimands). These findings provide qualified evidence that, in equilibrium, voter behavior aligns with the predictions of the unrestricted model.

5 Study #2: Information Experiments

The results in the previous section provide evidence consistent with the model advanced in this paper for the case of Brazil. To what extent can the model explain empirical patterns beyond this context? This is an open empirical question. In general, careful operationalization and measurement of variation in bureaucratic quality is necessary to extend the empirical examination to other contexts. While renewed efforts to measure bureaucratic quality at the national level are underway, the study of Brazil emphasizes the need to study variation in the quality of bureaucracies overseen by the politicians under study. Even without such data, however, the model offers some insights for how we consider evidence from other contexts.

Prior to imposing the model in more heterogeneous contexts, it is important to clarify the scope conditions of the theory. First, while the Brazilian meta-study focuses on mayors, the model should apply to elected politicians in national or subnational offices so long as they: (i) rely on bureaucrats

to implement or produce public goods; (ii) have the authority to monitor or oversee those bureaucrats. Second, while term limits are arguably an important feature of Brazilian politics, they are not necessary for the general dynamics with respect to politician allocations and voter updating to emerge. Finally, the theory intentionally considers a continuum of levels of bureaucratic quality and should be applicable in low, middle, and high capacity settings if these scope conditions are satisfied.

Because the empirical manifestations of accountability vary in bureaucratic quality, researcher choices of where to study accountability pose underappreciated limitations for how we understand accountability and its failures. Drawing from the combined efforts of Enríquez et al. (2019) and Incerti (2019), I identify 16 experiments or natural experiments on voter information and accountability conducted in eight countries, enumerated in Table A13. Situating these sites on (national level) macro indicators of bureaucratic quality, corruption, and public goods provision, Figure 7 suggests that such studies have, to date, been confined to democracies with low-to-middling levels of bureaucratic quality. Further, the motivation of some works suggests selection on features of the equilibrium, i.e. poor public goods provision or high corruption.

Examining eleven of these studies across all eight countries that: (i) are experimental and (ii) provide estimates on vote choice for the incumbent subsequent to the revelation of both good and bad news, Figure 8 depicts the relationship between national bureaucratic quality and the effects of information provision. On average, “good news” modestly increases incumbent vote share and “bad news” modestly reduces incumbent vote share, but only as bureaucratic quality (within sample) increases. Because all experiments are conducted in different constituencies the implication here is that the separating equilibrium appears to be more common (across constituencies) at higher levels of (national) bureaucratic quality. This finding is only suggestive and future research could strengthen this analysis in three ways. First, future experiments should avoid building the sample of constituencies by selecting on equilibrium outcomes (corruption or public goods provision). Second, measurement of the quality of bureaucrats actually managed by the politicians in question would allow up to these dynamics much more precisely within cases. Finally, research designs

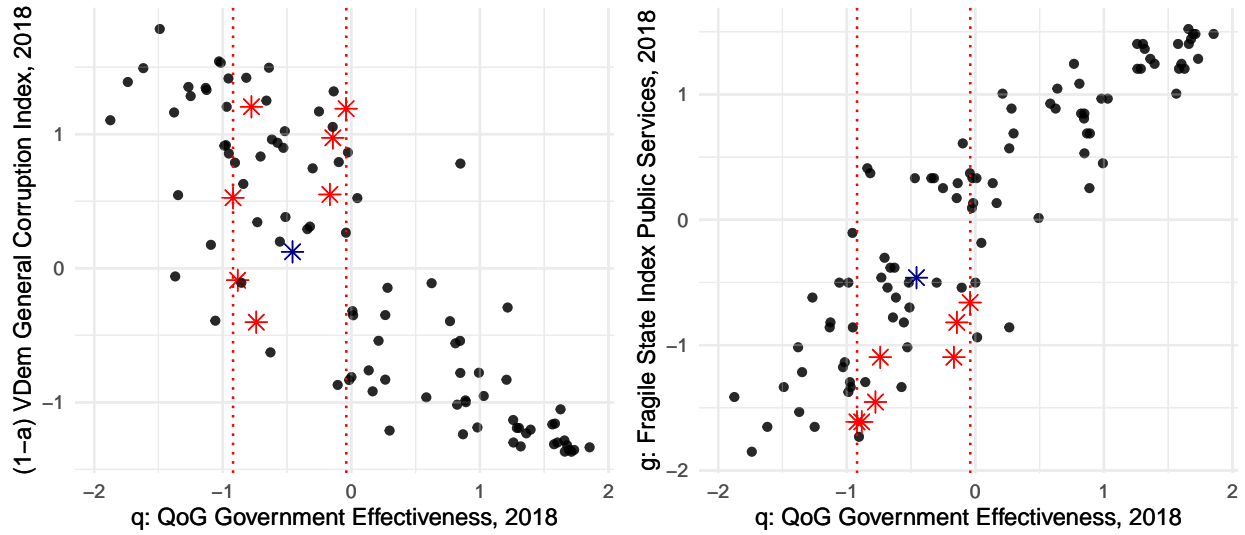


Figure 7: Correlations between bureaucratic quality (x -axis) and corruption (y -axis, left panel) and public service provision (y -axis, right panel). All measures are standardized to the set of democracies (defined by the Quality of Government dataset), such that all measures are z -scores. The navy star indicates Brazil and the red stars indicate other countries in which I identify accountability experiments/natural experiments.

that study accountability in a parallel fashion across places that with more substantial variation in bureaucratic quality (see Figure 7) would allow for more comprehensive tests of the theory than are afforded by the existing estimates.¹³

Considering this evidence from relatively similar studies across multiple sites yields two central takeaways. First, the observational equivalencies generated by the model provide new scope conditions on what inferences we can draw about accountability from partial equilibrium tests of information and accountability. Absent characterization of the underlying equilibrium, zero or null results on the effects of information provision provide less evidence that voters are uninformed or unable to update than is currently implied. Second, under the model I advance, estimates of a common distribution of treatment effects will be attenuated toward zero when sites fall into either pooling equilibrium. This is the central finding of two influential meta-analyses (Dunning et al., 2019; Incerti, 2019). Figure 8 provides suggestive evidence as to why such an approach may not be able to capture the effect of information on accountability. More generally, this reasoning mo-

¹³Figure 8 ultimately includes only eight country-level estimates of bureaucratic quality, which approximate “clusters” for the purpose of this analysis.

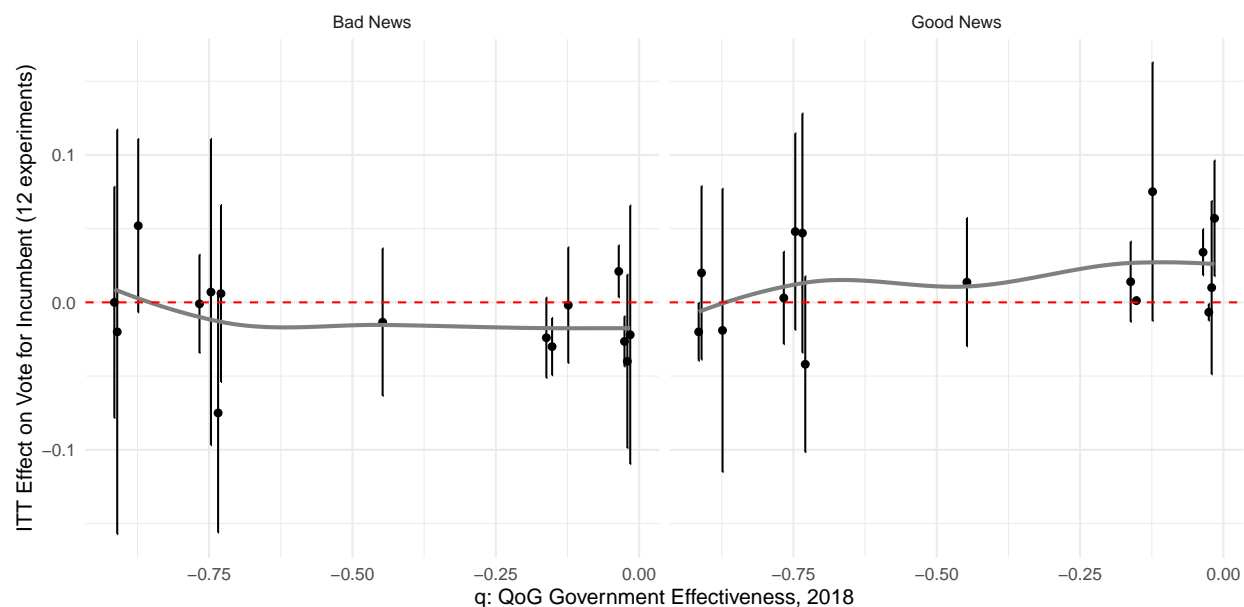


Figure 8: ITTs on vote choice for the incumbent or incumbent's party in 12 experiments in 8 countries using survey and/or administrative data. The dependent variable can be interpreted as the change in incumbent (party) vote share as a function of the revelation of incumbent performance information, by the type of signal “bad news” (left) or “good news” (right). Points are jittered on the x -axis for visibility. Estimates and standard errors come directly from estimated by these 12 studies.

tivates a need to define external validity relative to a mechanism, as opposed to a point estimate. This broader conception of external validity opens new avenues for cumulation of evidence and research design (Slough and Tyson, 2022).

6 Conclusion

Recent literature on accountability remains mixed on whether a better informed electorate can improve political selection, drawing a (relatively) pessimistic conclusion that empowering voters with information is, at best, circumscribed in its ability facilitate better governance. The accumulated evidence tends to describe such outcomes in terms of the deleterious effects of uninformed voters and the venality of politicians. I contend that this attribution of bad outcomes—corruption and limited public goods provision—to a form of “bad politics” is particularly widespread in the study of developing democracies.

The theory advanced in this paper takes a different approach to explaining similar patterns of outcomes. Indeed, the model of electoral accountability advanced in this paper assumes that politicians uniformly value the provision of public goods. Voters are informed and rational (Bayesian). Bureaucrats shirk, but are responsive to oversight and are not otherwise corrupt. Yet, the theory predicts the confluence of corruption, underprovision of public goods, and voter behavior often taken to motivate claims of circumscribed electoral accountability can emerge with informed, rational voters when bureaucratic capacity is low. Furthermore, the dissemination of politician performance metrics conveys no information in the pooling equilibria at very low and, in the case of information about politician performance, high levels of bureaucratic quality. Extension of multiple existing results from Brazil emphasizes the plausibility of this model, providing grounds for extension and further testing in other contexts. Examination of treatment effects from eleven information dissemination experiments provides suggestive evidence consistent with the theory in a wider variety of contexts.

This paper advances a broader appeal for theories that treat political actors symmetrically across contexts in order to generate comparative insight. To extent that the policy implications

of empirical findings depend on the underlying causal process generating outcomes, a bias toward “bad politics” may limit the insight that we can contribute in efforts to inform policies to advance welfare.

References

- Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105 (8): 2364–2409.
- Achen, Christopher H., and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton, NJ: Princeton University Press.
- Alesina, Alberto, and Guido Tabellini. 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97 (1): 169–179.
- Arias, Eric, Horacio Larreguy, John Marshall, and Pablo Querubin. 2019. "Priors Rule: When do Malfeasance Revelations Help or Hurt Incumbent Parties?" Available at https://scholar.harvard.edu/files/jmarshall/files/mexico_accountability_experiment_v13.pdf.
- Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 2020. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review* 110 (5): 1355–1394.
- Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15: 183–201.
- Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Fridenberg. 2017. "Accountability and Information in Elections." *American Economic Journal: Microeconomics* 9 (2): 95–128.
- Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Fridenberg. 2018. "Learning about Voter Rationality." *American Journal of Political Science* 62 (1): 37–54.
- Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126 (5): 1912–1964.
- Banerjee, Abhijit, Selvan Kumar, Rohini Pande, and Felix Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence From India." Available at https://scholar.harvard.edu/files/rpande/files/do_informed_voters_make_better_choices.pdf.
- Besley, Tomithy, Robin Burgess, Adnan Khan, and Guo Xu. 2021. "Bureaucracy and Development." *Annual Review of Economics* Forthcoming.
- Bhandari, Abhit, Horacio Larreguy, and John Marshall. 2019. "Able and Mostly Willing: An Empirical Anatomy of Information's Effect on Voter-Driven Accountability in Senegal." Available at https://scholar.harvard.edu/files/jmarshall/files/accountability_senegal_paper_v5.pdf.
- Boas, Taylor C., and F. Daniel Hidalgo. 2011. "Controlling the Airwaves: Incumbency Advantage and Community Radio in Brazil." *American Journal of Political Science* 55 (4): 869–885.

- Boas, Taylor C., F. Daniel Hidalgo, and Marcus André Melo. 2019. “Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil.” *American Journal of Political Science* 63 (2): 385–400.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs.” *Econometrica* 82 (6): 2285–2326.
- Calvo, Ernesto, and Maria Victoria Murillo. 2004. “Who Delivers? Partisan Clients in the Argentine Electoral Market.” *American Journal of Political Science* 48 (4): 742–757.
- Cho, In-Koo, and David M. Kreps. 1987. “Signaling Games and Stable Equilibria.” *The Quarterly Journal of Economics* 102 (2): 179–221.
- Chong, Alberto, Ana de la O, Dean Karlan, and Leonard Wantchekon. 2015. “Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification.” *Journal of Politics* 77 (1): 51–77.
- Cruz, Cesi, Philip Keefer, and Julien Labonne. 2018. “Buying Informed Voters: New Effects of Information on Voters and Candidates.” Available at https://static1.squarespace.com/static/58c979fad1758e09d030809c/t/5c048e82898583120b1f73cc/1543802523246/buying_informed_voters_web.pdf.
- Dal Bó, Ernesto, Federico Finan, and Martín A. Rossi. 2013. “Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service.” *Quarterly Journal of Economics* 128 (3): 1169–1218.
- Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh, and Gareth Nellis, eds. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*. New York: Cambridge University Press.
- Enríquez, José Ramón, Horacio Larreguy, John Marshall, and Alberto Simpser. 2019. “Information saturation and electoral accountability: Experimental evidence from Facebook in Mexico.” Working paper.
- Fearon, James. 1999. *Democracy, Accountability, and Representation*. Cambridge University Press chapter Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance.
- Ferrali, Romain, and Galileu Kim. 2019. “Are Anti-Corruption Audits Effective? Evidence From Brazilian Municipal Bureaucracies.” Working paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3499117.
- Ferraz, Claudio, and Federico Finan. 2008. “Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes.” *Quarterly Journal of Economics* 123 (2): 703–745.
- Ferraz, Claudio, and Federico Finan. 2011. “Electoral Accountability and Corruption: Evidence from the Audits of Local Governments.” *American Economic Review* 101 (4): 1274–1311.

- Foarta, Dana. 2021. "How Organizational Capacity Can Improve Electoral Accountability." Working paper available at https://gsb-faculty.stanford.edu/dana-foarta/files/2021/01/OrgCap_Nov2021.pdf.
- Fowler, Anthony, and Andrew B. Hall. 2018. "Do Shark Attacks Influence Presidential Elections? Reassessing a Prominent Finding on Voter Competence." *Journal of Politics* 80 (4): 1423–1437.
- Fowler, Anthony, and B. Pablo Montagnes. 2015. "College Football, Elections, and False-Positive Results in Observational Research." *Proceedings of the National Academy of Sciences (PNAS)* 112 (45): 13800–13804.
- Fox, Justin, and Stuart V. Jordan. 2011. "Delegation and Accountability." *Journal of Politics* 73 (3): 831–844.
- Geddes, Barbara. 1994. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley, CA: University of California Press.
- Gonçalves, Sónia. 2013. "The Effects of Participatory Budgeting on Municipal Expenditures and Infant Mortality in Brazil." *World Development* 53: 94–110.
- Grindle, Merilee S. 2012. *Jobs for the Boys: Patronage and the State in Comparative Perspective*. Cambridge, MA: Harvard University Press.
- Healy, Andrew J., Neil Malhotra, and Cecilia Hyunjung Mo. 2010. "Irrelevant Events Affect Voters' Evaluations of Government Performance." *Proceedings of the National Academy of Sciences (PNAS)* 107 (29): 12804–12809.
- Huber, John D., and Michael M. Ting. 2020. "Civil Service and Patronage in Bureaucracies." *Journal of Politics* Forthcoming.
- Humphreys, Macartan, and Jeremy M. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda - Preliminary Analysis." IGC Working Paper S-5021-UGA-1.
- Incerti, Trevor. 2019. "Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design." Working paper, available at https://tincerti.github.io/files/corruption_voting_meta.pdf.
- Izzo, Federica, Torun Dewan, and Stephane Wolton. 2019. "Cumulative Knowledge in the Social Sciences: The Case of Improving Voters' Information." Working paper.
- Klašnja, Marko, and Rocío Titunik. 2017. "The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability." *American Political Science Review* 111 (1): 129–148.
- Larreguy, Horacio, John Marshall, and James M. Snyder Jr. 2020. "Publicizing Malfeasance: When the Local Media Structure Facilitates Electoral Accountability in Mexico." *Economic Journal* Forthcoming.
- Li, Christopher M., Greg Sasso, and Ian Turner. 2019. "Accountability in Governing Hierarchies." Working paper.

- Marshall, John. 2019. “When Can Close Election RDDs Identify the Effects of Winning Politician Characteristics?” Working paper, Columbia University.
- Maskin, Eric, and Jean Tirole. 2004. “The Politician and the Judge: Accountability in Government.” *American Economic Review* 94 (4): 1034–1054.
- Przeworski, Adam, Susan Stokes, and Bernard Manin. 1999. *Democracy, Accountability, and Representation*. Cambridge University Press chapter Introductoin, pp. 1–26.
- Raffer, Pia. 2022. “Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from a Dominant Party Regime.” *American Political Science Review* Forthcoming.
- Raffler, Pia, and Lucy Martin. 2019. “Fault Lines: The Effects of Bureaucratic Power on Electoral Accountability.” *American Journal of Political Science* Forthcoming.
- Rauch, James E. 1995. “Bureaucracy, Infrastructure, and Economic Growth: Evidence from U.S. Cities During the Progressive Era.” *American Economic Review* 85 (4): 968–979.
- Seabra, Sérgio N. 2018. “How Auditing Can be Directed Against Corruption: The Case of Brazil’s Municipality-Facing Auditing Program.” Working paper, University of Oxford.
- Slough, Tara. 2022. “Phantom Counterfactuals.” *American Journal of Political Science* Forthcoming.
- Slough, Tara, and Scott A. Tyson. 2022. “External Validity and Meta-Analysis.” *American Journal of Political Science* Forthcoming.
- Ting, Michael M, James M Snyder Jr., Shigeo Hirano, and Olle Folke. 2012. “Elections and Reform: The Adoption of Civil Service Systems in the U.S. States.” *Journal of Theoretical Politics* 25 (3): 363–387.
- Toral, Guillermo. 2019. “The benefits of patronage: How the political appointment of bureaucrats can enhance their accountability and effectiveness.” Working paper, MIT.
- Varjão, Carlos. 2019. “The Role of Local Media in Selecting and Disciplining Politicians.” Working paper, Stanford University. Available at <https://drive.google.com/file/d/1W14ILSg3tK3AwwfP0hRNG3dNZk5aVxk8/view>.
- Weitz-Shapiro, Rebecca, and Matthew S. Winters. 2016. “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. 2016. “Who’s in Charge Here? Direct and Indirect Accusations and Voter Punishment of Corruption.” *Political Research Quarterly* 69 (2): 207–219.
- Yazaki, Yukihiro. 2018. “The Effects of Bureaucracy on Political Accountability and Electoral Selection.” *European Journal of Political Economy* 51: 57–68.