Don't Know, Don't Care? Experimental Evidence on Barriers to Responsive Public Services in Ethiopia*

Daniel Agness
University of Maryland

Pascaline Dupas
Princeton University

Fiker Negash
Princeton University

Tigabu Getahun
Policy Studies Institute

January 15, 2026

Abstract

In a sample of local public officials working in Ethiopia's decentralized bureaucracy, we document low knowledge of public priorities and substantial mismatch between citizen and government priorities. Using three experiments, we show that bureaucrats fail to update beliefs or adapt policy when provided with information about local conditions and citizen preferences for public services, and bureaucrat demand for this information is low. Mismatch between citizen and local government priorities in this setting cannot be explained by information frictions, lack of autonomy, or bureaucrat turnover. Our results challenge standard assumptions in models of fiscal federalism and highlight the limitations of decentralized bureaucracies in providing locally tailored public services without incentives or intrinsic motivation to meet public demand.

JEL Classifications: C93, H11, H72, D73, D82, D83, O55
Keywords: decentralization, public services, bureaucrat incentives, bureaucrat beliefs, policy mismatch

*We are extremely grateful to Abis Getachew, Deivy Houeix, Eva Lestant, and Yalew Mekonnen for outstanding field research assistance and to seminar participants at the University of Maryland, and Princeton University for their valuable feedback. This study was reviewed and approved by the Institutional Review Board (IRB) of Stanford University and is registered in the AEA RCT Registry as AEARCTR-0002114. We thank the Stanford King Center on Global Development (AUDRI initiative) and the Gates foundation (J-PAL Urban Services Initiative) for financial support. Dupas also gratefully acknowledges support from the Guggenheim Foundation. Agness was generously supported by the National Science Foundation's Graduate Research Fellowship Program.

1 Introduction

The value of decentralization lies in the belief that local public officials, with better access to information on local conditions and preferences, are best positioned to efficiently provide or procure public services (Oates, 1972; Garicano, 2000). Economic theory and canonical models of fiscal federalism suggest the informational advantages and, in many contexts, greater accountability associated with the local control of public expenditures can improve efficiency, with decentralization generally making policy more responsive to public demands (Tiebout, 1956; Wallis and Oates, 1988) and promoting growth (Bardhan, 2002; Hatfield, 2015).[1] This rationale has led to the promotion of policy decentralization across lower-income countries (Ahmad et al., 2003; Gadenne and Singhal, 2014), where stakes are particularly high given limited fiscal capacity and lower quality infrastructure and services.

The realization of decentralization's benefits requires at a minimum that: (1) local officials either know or seek out information about public preferences and local conditions; and (2) these officials have the capacity and the will to shape responsive policy. However, information frictions and incentive misalignment within a bureaucracy could prevent these conditions from holding. While elected officials, the focus of the majority of the literature, have electoral incentives that promote alignment with the public (Besley and Case, 1995), much less is known about the motivation and performance of unelected bureaucrats in decentralized systems (Ujhelyi, 2014).[2] Yet, as these bureaucrats wield substantial control over policy implementation, they play a critical role in the provision of public goods (Besley et al., 2022; Best et al., 2023).

We study responsiveness to public priorities through three experiments with local bureaucrats overseeing administrative units in and around Addis Ababa, Ethiopia. Having experienced dramatic growth over the past decade, Addis Ababa and its surrounding region, Oromia, face the challenges of urban service provision common to rapidly urbanizing

[1] Some models suggest that under certain conditions, decentralization can decrease accountability. See Boffa et al. (2016) for one such example.

[2] In their review article on personnel economics of developing states, Finan et al. (2017) write "As this survey documents, the (experimental) research on trying to understand how bureaucracies work is still in its infancy, so there is plenty to do and a lot to learn."

areas in Africa and throughout the global South. Provision of public goods and services falls predominately to a large, and increasingly decentralized, system of bureaucrats (Kosec and Mogues, 2020). But public services have failed to keep up with demand: in a baseline survey we conducted in 2016 with 3,540 households living in the outskirts of Addis Ababa, households reported spending 27% of days without electricity for at least 1 hour and 43% of days without drinking water. In the same survey, 58%, 83%, and 82% of households report being dissatisfied with the public electric power supply, water and sanitation, and road quality, respectively. Around 32% of households had ever complained to public officials. We probe this mismatch between public needs and public service provision. Is the rapidly changing economic and social landscape creating information frictions that undermine the ability of local officials to be responsive to public needs? Or are there other barriers to local government's responsiveness?

High levels of dissatisfaction with public services are not, in themselves, evidence of inefficient governance or information frictions. They could simply reflect limited resources for public service provision. We thus further motivate our experiments with two novel facts about citizen and bureaucrat priorities in Addis Ababa. First, through a survey of 626 local bureaucrats drawn from the periphery of Addis Ababa, we show that the majority are unaware of the public service utilization or policy priorities of their constituents. Comparing bureaucrats' estimates with a representative household survey, we find only 7% (21%) and 11% (16%) of bureaucrats accurately estimate the frequency of power and water outages within 10% (25%) of residents' reports. Bureaucrat responses also show high variance in estimating service utilization and unemployment rates. When asked to identify the five sectors (out of 14) most important to the public, administrators correctly identified 2.5, on average. They fail to rank the public's top priority in the top five 41% of the time and fail to rank at least two of public's top three priorities 43% of the time. Even education and health, which are top citizen priorities in 67% percent of woredas and represent the largest shares of woreda budgets, are only correctly identified by administrators 34% of the time.[3]

[3]The survey was conducted with woreda chief administrators and an average of 10 sector bureau heads in each woreda in Addis Ababa that had a border touching the city limits, as well as woredas proximate to Addis Ababa in the surrounding Oromia region. More information on the survey and sampling can be

This result is consistent with descriptive work from rural Uganda documented in Azfar et al. (2006).

Second, we show that there is substantial misalignment between the policy priorities of the public and those of the local administrations: while 76% of administrators identify at least one of the public's priority sectors as one of their own, only 36% have two or more of their top three sectors align with citizens' top three. Taken together, these results imply either that information frictions prevent bureaucrats from knowing what the public desires or that organizational incentives render this information unimportant.

In our first experiment, we test whether and how local bureaucrats respond to information about public service utilization and citizen priorities in their jurisdictions. If information frictions are the primary determinant of the policy misalignment we observe, we would expect bureaucrats to internalize new information about public priorities, update their organizational priorities, and adapt policy accordingly. We randomly assign *woredas*, the lowest administrative level in Ethiopia, and *bureaus* managing particular policy sectors within those woredas, into treatments where they receive "report cards" containing information about public priorities and service satisfaction. These report cards aggregate information drawn from contemporaneous, representative surveys with households and firms located in each respective woreda so that the information provided is highly localized and relevant. Within treated woredas, we cross-randomize whether bureaus receive a report card for their policy sector directly, or whether the report card is shared only with the woreda chief administrator to test for differential responses to information within the bureaucratic hierarchy.

We estimate precise null treatment effects from receiving a report card on chief administrator knowledge of citizen priorities and conditions. Consistent with this result, we show that the report cards do not improve the alignment between bureaucrat's priorities and those of citizens: administrators in treated woreda do not update their stated policy priorities, implement new targeted projects, nor use the report cards to request or receive additional funds. Using detailed, line-item budget data, we can reject 1% increases in treated woreda- and bureau-level budgets, respectively. Unsurprisingly, given the lack of administrative re-

found in Section 5.

sponse, report cards do not increase citizen satisfaction with public services in their woreda a year later. There is no treatment effect heterogeneity based on whether report cards are given to bureau heads or woreda chief administrators.

Having shown that woreda officials neither internalize nor respond to citizen information, in our second experiment we test directly for bureaucrat demand for information about their constituents. Eighteen months after the report card experiment, we cross-randomize whether woreda administrators are given the opportunity to "opt-in" to or "opt-out" of receiving new report cards. To elicit their "willingness to pay", we further randomize the length of the form they have to fill in order to opt-in or out. Bureaucrats were invited to request report cards on sectors and topics of their choice, and to request statistics of their choice (since we elicited their interest prior to the second round of data collection).

We find that demand for information is low, with only 14% (2%) of bureaucrats willing to spend five (ten) or more minutes to "order" a report card. We estimate an arc-elasticity of -1.21 with respect to time. There is no differential demand based on treatment status in the first experiment, ruling out that low demand and our null effects are driven by the specific design or content of our original report cards. Furthermore, we show that 5% of woreda administrators are willing to spend at least 5 minutes to opt-out, i.e., to *not* receive a report card. We rationalize this finding in a theoretical framework where high costs of policy implementation or psychological costs associated with negative information could drive costly avoidance behavior.

Report cards have the potential to influence decision-makers due to both the information provided and the "threat" of outside monitoring. Our third experiment is designed to separate these mechanisms and test whether the outside monitoring alone can impact bureaucrat behavior. Using the remaining woredas in Addis Ababa, for which we did not have household or firm surveys, we provide notices to randomly selected woreda chief administrators and bureau heads that we will be conducting research in their woreda. Randomization and the design of these notices mirrored the report card experiment, but did not contain locally relevant statistics. Consistent with the lack of impact of the Report Cards themselves, we find that the monitoring-only treatment had negligible effects on each of our primary outcomes,

confirming that third-party monitoring is insufficient to change local bureaucrat behavior or impact public service delivery when the center's priority is elsewhere.

Our results cannot be explained by insufficient bureaucrat autonomy, where decentralized bureaucrats have policymaking authority according to the law, but *de facto* authority remains centralized (Faguet, 2014). First, we document substantial variation in sectoral spending across woredas showing that woreda administrations hold substantial budgetary autonomy. Second, while only 43% of surveyed administrators report having at least some influence over policy determination in 3 or more sectors, large majorities report having the authority to deliver performance-based rewards or punishments. But administrators who report higher levels of autonomy are no more knowledgeable or aligned with public priority sectors at baseline, and there are no differential treatment effects on knowledge, alignment with public priorities, or budget allocations for high-autonomy sectors. Our results strongly suggest, consistent with our theoretical framework, that while local bureaucrats have autonomy to make policy decisions, they are not incentivized to make policy responsive to local demand.

Finally, we examine whether *turnover* among chief administrator and bureau heads could have limited information sharing and subsequent treatment effects. Turnover was high during our period of study, though consistent with historical levels of turnover. Consequently, understanding whether bureaucratic performance can be maintained, and knowledge retained, following bureaucrat exit is important. After showing that the treatment had no impact on turnover, we look at heterogeneity in impacts by whether the woreda experienced bureaucrat turnover. First, we show that the treatment *increased* policy misalignment among longer-tenured chief administrators; this rules out the average null effects can be explained by turnover alone.

Our results call into question key theoretical assumptions in models of fiscal federalism; they suggest that absent stronger incentives to align bureaucrats with public priorities, decentralization is likely to fail to promote more responsive policy. In doing so, we contribute to a literature in economics and political science on the merits and drawbacks of decentralization (Wallis and Oates, 1988; Prud'homme, 1995; Gadenne and Singhal, 2014; Boffa et al., 2016) and the principal-agent problems inherent in bureaucratic hierarchies (Bertrand

et al., 2019; Besley et al., 2022). By focusing on unelected bureaucrats, rather than political appointees or politicians, we highlight an understudied layer of the administrative hierarchy that has been shown to play an important role in policy performance (Bhavnani and Lee, 2018; Dal Bó et al., 2018; Best et al., 2023; Aneja and Xu, 2024; Arora et al., 2024).

We also contribute to a growing literature on policymaker responses to information.[4] While this literature has generally studied politicians (Butler et al., 2011; Hjort et al., 2021; Callen et al., 2020), invited workshop delegates (Vivalt and Coville, 2023; Vivalt et al., 2025), or high-level agencies (Toma and Bell, 2024), few have focused on low-level bureaucrats. We add a particular emphasis on misalignment with public priorities (Liaqat, 2023; Bazzi et al., 2025). To our knowledge, we are the first to experimentally estimate demand for information by bureaucrats and to document the potentially negative demoralization effects of information.

The most closely related paper to ours in this literature is Rogger and Somani (2023), conducted concurrently with a sample of bureaucrats throughout the Ethiopian administrative hierarchy. That paper first shows that while woreda-level bureaucrats have more accurate information about local statistics (such as population size, or the number of children enrolled in school) than their federal or regional counterparts, they are still substantially misinformed. Second, it shows that evidence briefings including these statistics marginally improved accuracy among bureaucrats two weeks after their delivery. We build upon this work in a number of ways. First, we focus on bureaucrat knowledge of key service delivery indicators and measures of public preferences instead of administrative statistics collected 2-9 years before the intervention; repeated visits to respondents in our sample, and the inclusion of a baseline, allow us to study changes over time. Our experiments additionally measure knowledge and policy alignment between four months and up to two years after the intervention and allow estimation of an explicit measure of demand for information. Finally, we are able to measure downstream impacts (or lack thereof) on realized policy choices using outcome data from households, firms, and woreda budgets.

[4] A related, but distinct, literature in political science focuses on the biased beliefs of politicians (e.g. Badin et al. (2025); Brookman and Skovron (2018); Pereira (2021); Banuri et al. (2019)). We support this literature by documenting low knowledge of public priorities by local bureaucrats.

The paper proceeds as follows. We describe our context in Section 2. We then build a simple theoretical framework to motivate our experiments in Section 3. Next, we describe the experimental design in Section 4 and the data in Section 5. In Section 6 we present our primary results before finishing with a discussion in Section 7.

2 Context

2.1 Decentralization in Ethiopia

Following the fall of the Marxist Derg regime in 1991, Ethiopia began a process of decentralization in order to promote regional autonomy and localized public service delivery. In a first phase immediately following the regime change, the federal government devolved substantial powers to nine ethnic regions and two chartered cities (Addis Ababa and Dire Dawa). This was followed by a second phase in 2002-2003, giving more control to woredas, the lowest level of municipal government in most of the country (Garcia and Rajkumar, 2008; Bongwa et al., 2011). Today there are approximately 1,050 woredas across the country, with an average population of 110,000 per woreda. Studies leveraging the staggered roll-out of these policies in Ethiopia's rural regions find mixed impacts on public service delivery (Kosec and Mogues, 2019; Faguet et al., 2020; Kosec and Mogues, 2020). Implementation is widely viewed as imperfect, with central government organizations maintaining substantial influence over technically delegated responsibilities (Yimenu, 2023).

This phased decentralization resulted in a three-tier municipal structure below the federal government—regions, zones (sub-cities), and woredas—with special status given to the cities of Addis Ababa and Dire Dawa.[5] Our study takes place in Addis Ababa and the region that surrounds it, the Oromia region. Addis Ababa is governed much like a region, with similar levels of autonomy, while containing 11 zone equivalent sub-cities and 140 woredas.[6]

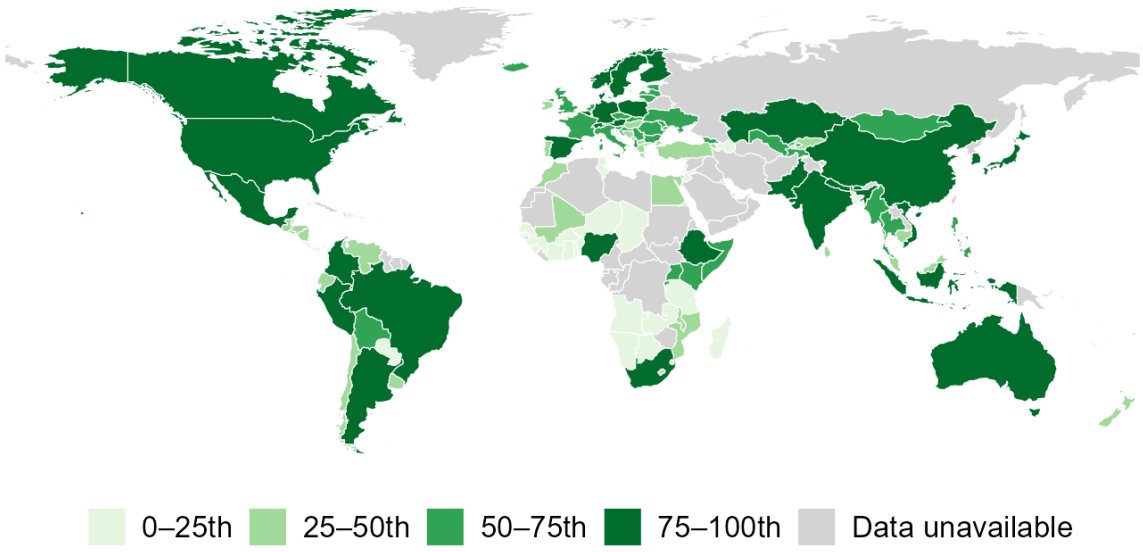
[5]There is a fourth level, the kebele, below woredas that is primarily concerned with record keeping. Kebeles have few governance powers or implementation responsibilities, particularly in our region of study. Within each kebele there are non-administrative divisions, ketenas and gotes, that support population tracking and political organization.

[6]At the time of the study, there were 10 sub-cities and 117 woredas in Addis Ababa.

Amendments to the Addis Ababa city charter in 2003 enumerated the role of woredas to serve as foci for local service delivery in the city, as in the surrounding regions (Bongwa et al., 2011). Woredas were given not only responsibilities, but also real resources in the form of personnel and block grants, such that nearly 80% of city expenditures were managed by woredas in Addis Ababa in 2012; in Oromia, this figure was 50% (Kiringai et al., 2016).

Figure 1 documents how Ethiopia compares to other countries across the globe in terms of sub-national (state or regional) government share of total government expenditures, a proxy for decentralization. Ethiopia's level of decentralization is considerably higher than most others on the African continent, and comparable to those in the United States, India, China, and Indonesia, which have been studied as models of decentralization in prior research. In Appendix Figure A1 we show analogous map looking at local government (second or third tier) share of total expenditures using alternative sources of data in which Ethiopia appears as the most decentralized country in Africa. Focusing on Africa's most decentralized nation allows us to explore the value of decentralization on the continent and investigate its success in a context with low state capacity.

Figure 1: Share of general government expenditures executed at the state or regional level



Notes: The figure shows the share of general government expenditure executed at the state or regional government level, grouped into quartiles across countries. Darker shades represent higher shares of the national budget executed at the state or regional level. The state or regional level includes federated regions in federal and quasi-federal countries, as well as related public entities such as special-purpose bodies and state public institutions. Data are drawn from the OECD–UCLG World Observatory on Subnational Government Finance and Investment (SNG-WOFI) for 2020, including Ethiopia. Appendix A1 shows the share of general government expenditures executed at the local level using a mix of SNG-WOFI and national fiscal data.

2.2 Woreda Responsibilities and Structure

The median woreda in our sample covers a population of 30,098 in Addis Ababa and 96,034 in Oromia.[7] Within each woreda there is an appointed chief administrator, an elected council, and numerous sector bureau managers (e.g. health, education, planning, women's affairs, and sanitation). Woredas have autonomy over which bureaus they maintain, resulting in heterogeneity in bureaus across woredas. Figure A2 displays the share of woredas that allocated funding to each sector bureau. On average, woreda administrations direct funds

[7]The 25th percentile of woreda population is 24,059 (78,637) in Addis Ababa (Oromia). The 75th is 36,467 (112,578).

across 22 bureaus and 560 total administrative employees.[8]

Responsibility for the procurement of public infrastructure (e.g. roads, water, electricity), infrastructure maintenance, and the provision of specific public services (e.g. health, education), varies by sector and is shared across the different administrative levels. For example, in our baseline survey with woreda administrators, described in further detail in the next section, approximately 50% of woreda chief administrators report having at least some influence over health, water, and small-enterprise policy *decisions* in their woreda; and 89% report having at least some influence over policy *implementation* in those sectors. Conversely, only 14% report having any influence over electricity policy in their woreda, a service largely managed at the city or regional level.

Woreda administrators report high levels of autonomy over personnel and woreda-level budgeting: 97% report having authority to punish poor performers and 69% claim to have substantial influence over budget allocation across sectors.[9] Consistent with the maintenance of budgetary authority, we see substantial heterogeneity across woredas in budget allocations across bureaus. Figure 2 shows the share of woreda budgets dedicated to three key services: education, health, and small and micro enterprise (SME) development. Woreda expenditure on health varies from less than 2% of the overall budget to over 20%; expenditure on education and SME development vary from less than 5% and 2% to nearly 50% and 10% of the overall budget, respectively. Variation in these budget shares, across woredas, together with our survey evidence indicating that budgets had been previously reallocated based on public feedback, supports woreda administrations having autonomy over policy choices.[10]

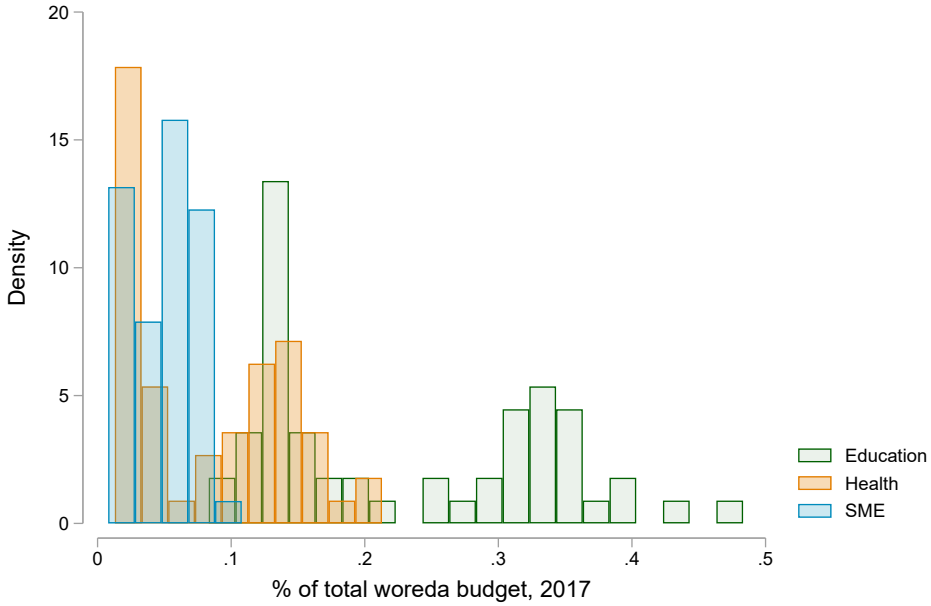
Woreda chief administrators are appointed by city and regional administrations to indefinite terms and are not necessarily replaced following political transitions. Bureau heads are, in turn, appointed by chief administrators. While there has been little research into

[8]This total does not include frontline service providers such as teachers and nurses who are typically paid by zones or the city administration.

[9]Over 80% of woreda finance managers report having substantial authority over budget allocations (Table A5).

[10]Approximately 15% of budget allocation changes in the previous 2 years were directly attributed by WCAs to public requests. The remained were either decided unilaterally by the woreda administration, responding to higher authority priorites, or a due to a combination of factors.

Figure 2: Share of woreda budget across 3 key sectors in FY 2017.



the nature of appointments and promotions in the Ethiopian bureaucracy, qualitative work (Terefe Gemechu et al., 2020) and our surveys with bureaucrats imply that political connections play a central role.

Relative to their elected counterparts, bureaucrats in public administrations are likely to have weaker incentives to prioritize public or respond to public concerns (Kosec and Mogues, 2020). Our surveys present some of the first evidence on how low-level bureaucrats in Sub-Saharan Africa interact with and respond to the needs of their constituents. Woreda administrators report spending substantial portions of their time interacting with citizens and receiving regular feedback from the public: on average administrators claim to spend 15% of their working hours interacting with citizens and 81% of woredas claim they receive at least monthly feedback from the public (Table A1). These results are consistent with the duties and mandates outlined by Ethiopian policy (Bongwa et al., 2011). Using citizen self-reports, however, Table 1 shows that between 5 and 23% of citizens report having ever lodged a complaint about education, water, power or transport, suggesting that administrators may have overstated their degree of contact with citizens.

Table 1: Baseline self-reported contact with woreda among citizens

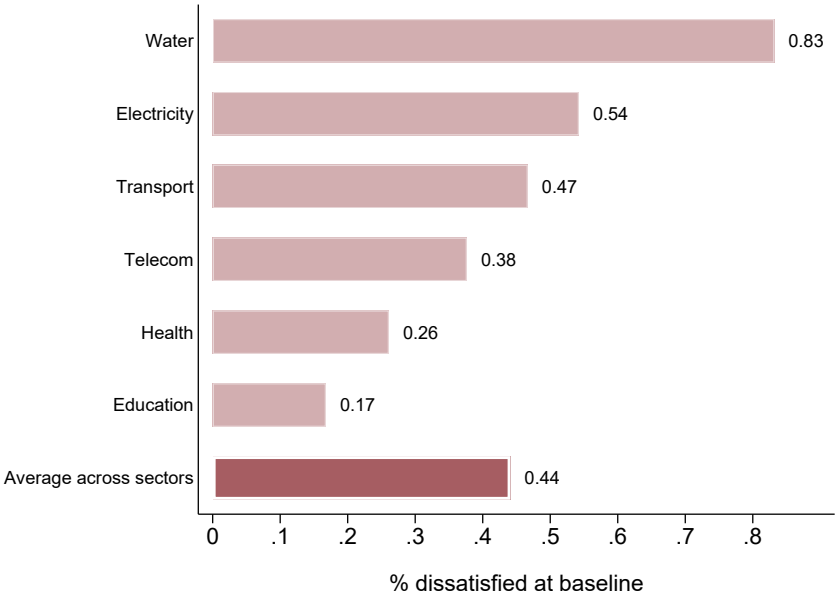
	N	Mean
Ever lodged complaint about:		
Public education	3496	0.10
Water quality	3496	0.18
Electric power supply	3496	0.23
Public transport	3496	0.05
Street lights	3496	0.09

Notes: The unit of observation is an individual. Data from the baseline citizen survey.

2.3 Baseline Service Delivery and Misalignment

Public Satisfaction: Addis Ababa's rapid growth over the past two decades has stretched the city's ability to provide public services and infrastructure (Maaskant, 2023; Agness and Getahun, 2024). We show that at baseline, citizens are highly dissatisfied with public services in their woreda. Across key sectors of service delivery, an average of 44% of Addis Ababa residents report being dissatisfied or highly dissatisfied with service provision or quality (Figure 3). This displeasure is not borne in silence as 40% of citizens report having complained to public officials or knowing someone who has complained about services and infrastructure in their neighborhood.

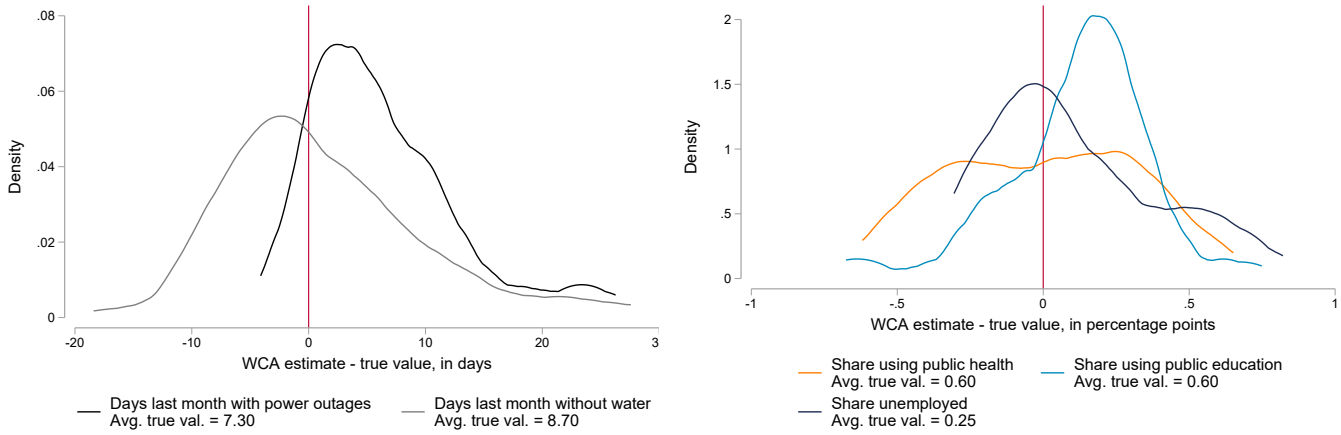
Figure 3: Baseline citizen dissatisfaction across sectors



Notes: The last bar in the figure reports the average share of citizens who reported being dissatisfied with a typical sector within the list shown.

Chief Administrator Knowledge: Public dissatisfaction may simply reflect bureaucratic budget constraints, not a lack of information, capacity, or will to address public concerns. To test whether local bureaucrats are knowledgeable about service quality and utilization in their woreda, we combine our administrator and household surveys. We find that chief administrators are inaccurate about service delivery and utilization. The results are presented in Figure 4. Administrators consistently overestimate the number days citizens are without power and have highly variable estimates of the number of days without water in a given month; only 7% and 11% of administrators are accurate within 10% of the public's reports about electricity and water shortages, respectively. There is similarly high variance in the estimates of the public utilizing public health services, public education, or currently unemployed. Together, these results imply that woreda administrations operate in low information environments, despite reports of regular public feedback. They are not only inaccurate about demographic characteristics of their constituents as shown by (Rogger and Somani, 2023),

Figure 4: Gap between WCA estimates and service quality/use



(a) WCA estimates v. service disruptions

(b) WCA estimates v. service use

Notes: Panel (a) shows a kernel density estimate of the distribution of the gap between admins' estimates and the true values of days of electricity and water service interruptions in a given month. Panel (b) shows the difference between admin's estimates and the true value of public health facility, public education, and unemployment.

they are also inaccurate about what their constituents care about.

Second-order Beliefs and Misalignment: Even in a low information environment, the priorities of woreda administrations may be well aligned with public demand. Again combining our administrative and household surveys, we compare the policy priorities of woreda administrations and those of the public. Table 2 shows that, on average, only one woreda administration priority is a top three public priority. Further, the average chief administrator is able to predict only 50% of the public's top five priorities. This confirms the presence of important information gaps between bureaucrats and the public.

Chief administrators consistently underestimate public preferences for improvements in water, transit, education, and roads, while overestimating their demand for housing and SME support (Panel A of Table A2). While the public consistently prioritizes health, electricity, jobs, roads, and education, few woreda administrators list these as top woreda priorities (Panel B of Table A2). Overall, for any issue outside of "security", there is alignment between the chief administrator and its constituents for a minority of woredas. Interestingly, misalignment with public priorities, and second-order beliefs about public priorities, do not

Table 2: Baseline misalignment between admin and citizen priorities

	Mean
Panel A: WCA preferences	
Share of the top 3 citizen priorities that are also WCA priorities	0.37
At least 1 is also a WCA priority	0.76
At least 2 are also WCA priorities	0.36
Panel B: WCA second-order beliefs	
Share of the top 5 citizen priorities recognized as such	0.50
At least 2 priorities recognized	0.91
At least 2 recognized (excl. Education & Health)	0.62
At least 3 priorities recognized	0.45
At least 3 recognized (excl. Education & Health)	0.28
N	58

Notes: Panel A reports the share of citizens' top 3 sector priorities that matched admins' own priorities. Panel B shows the share of citizens' top 5 sector priorities that admins correctly identified as priorities at baseline. Citizens' top 3 and top 5 priorities were determined from the citizen survey, where individuals ranked the top 5 sectors they wanted to see improved. These rankings were converted into scores, averaged, and re-ranked at the woreda level.

vary with woreda sectoral autonomy (Table A3). This suggests that misalignment cannot be explained by lack of woreda autonomy alone.

3 Theoretical Framework

We model a local bureaucrat's decision to acquire information about the public's ideal policy and to propose a policy that trades off alignment with the center's priority (e.g., fiscal conservatism, war on crime) and responsiveness to the public.

There are two focal points:

- The *center's* priority is $\theta_c \in \mathbb{R}$, which is known to the local bureaucrat.
- The *public's* ideal policy is $\theta_p \in \mathbb{R}$, which is initially unknown to the local bureaucrat.

The local bureaucrat chooses a proposed policy $p \in \mathbb{R}$. Throughout, we use a quadratic distance metric:

$$D(p, \theta) \equiv (p - \theta)^2. \tag{1}$$

This yields closed-form policy rules and a transparent expression for the value of information. Appendix D.2 discusses how the main intuition extends to other convex distance functions.

3.1 Timing and primitives

1. Nature draws the public ideal $\theta_p \sim \mathcal{N}(0, \sigma^2)$. The local knows the distribution and also knows the center ideal θ_c.

2. The local chooses *binary* information acquisition effort $e \in \{0, 1\}$ at cost ce with $c \geq 0$.

3. If $e = 1$, the local observes θ_p perfectly; if $e = 0$, the local retains the prior.

4. The local proposes a policy p. The implemented policy is assumed to be p, though we consider finite discretion or feasibility constraints in Appendix D.

5. The center observes p and θ_p.

Beliefs. Let $\mathcal{I}$ denote the local's information set at the time of choosing p. Define

$$m \equiv \mathbb{E}[\theta_p \mid \mathcal{I}], \qquad V \equiv \mathrm{Var}(\theta_p \mid \mathcal{I}).$$

Under the information structure:

$$\begin{aligned} e = 0: &\quad m = 0, \quad V = \sigma^2; \\ e = 1: &\quad m = \theta_p, \quad V = 0. \end{aligned}$$

3.2 Preferences

The local chooses (e, p) to minimize the loss function

$$L(p, \theta_c, \theta_p, e) \;=\; \beta\, D(p, \theta_c) \;+\; (\alpha_\ell + \alpha_c)\, D(p, \theta_p) \;+\; ce, \tag{2}$$

where:

- $\beta \geq 0$ represents autonomy and captures the extent to which the local is disciplined by alignment with the center's priority θ_c;[11]
- $\alpha_\ell \geq 0$ captures the local's intrinsic motivation to be responsive to the public;
- $\alpha_c \geq 0$ captures the center's desire for policy to be locally responsive (modeled as an additional weight in the local's objective, i.e., as internalized by the local);
- $c \geq 0$ is the fixed cost of exerting effort.

Define the total weight placed on public responsiveness as

$$\alpha \equiv \alpha_\ell + \alpha_c. \tag{3}$$

3.3 Optimal policy

Fix an information set $\mathcal{I}$ summarized by (m, V). Using (1), the local's expected loss conditional on $\mathcal{I}$ is

$$\begin{aligned} \mathbb{E}[L \mid \mathcal{I}] &= \beta(p - \theta_c)^2 + \alpha\, \mathbb{E}[(p - \theta_p)^2 \mid \mathcal{I}] + ce \\ &= \beta(p - \theta_c)^2 + \alpha((p - m)^2 + V) + ce. \end{aligned} \tag{4}$$

Lemma 1 (Optimal policy rule)**.** *For any beliefs (m, V), the unique policy that minimizes (4) is*

$$p^*(m) = \frac{\beta\theta_c + \alpha m}{\beta + \alpha}. \tag{5}$$

The local sets policy as a weighted average of the center ideal θ_c and the believed public ideal m. Responsiveness to the public is governed by the share $\alpha/(\beta+\alpha)$, while responsiveness to the center is governed by $\beta/(\beta + \alpha)$.

Corollary 1 (Comparative statics for responsiveness)**.** *Policy responds to the belief mean m*

[11] We could alternatively model autonomy as a separate parameter constraining the set of feasible policies available to the local bureaucrat. All intuition from the model goes through.

with slope

$$\frac{\partial p^*}{\partial m} = \frac{\alpha}{\beta + \alpha} \in [0, 1],$$

which is increasing in α (and thus in both α_ℓ and α_c) and decreasing in β.

3.4 Effort choice

Given Lemma 1, the local compares expected losses under $e = 0$ and $e = 1$.

No effort ($e = 0$). Then $(m, V) = (0, \sigma^2)$ and

$$p_0^* \equiv p^*(0) = \frac{\beta}{\beta + \alpha}\, \theta_c. \tag{6}$$

Effort ($e = 1$). Then $m = \theta_p$ and $V = 0$, so

$$p_1^*(\theta_p) \equiv p^*(\theta_p) = \frac{\beta\theta_c + \alpha\theta_p}{\beta + \alpha}. \tag{7}$$

Proposition 1 (Binary effort rule)**.** *Assume $\theta_p \sim \mathcal{N}(0, \sigma^2)$. The local exerts effort if and only if*

$$e^* = \mathbf{1}\{\, c \leq \frac{\alpha^2}{\beta + \alpha}\, \sigma^2 \,\}. \tag{8}$$

Equation (8) implies that information acquisition depends on how much the local cares about public alignment relative to center alignment and on uncertainty about public preferences: the likelihood of effort is increasing in α (hence in α_ℓ, the local's intrinsic motivation to meet the public's demand, and α_c, the extent to which the local internalizes the value the center puts on meeting the public's demand) and in σ^2 (the uncertainty about what citizens demand); and decreasing in c. Similarly, the likelihood of effort is decreasing in β: stronger emphasis on alignment with the center's own priority makes public information less valuable. Note that the effort threshold does not depend on θ_c: the value of learning θ_p is driven by the *uncertainty* in public preferences and the weight placed on public responsiveness. So θ_c affects the level of expected loss identically in both the informed and uninformed states, but not the incremental benefit of acquiring information about θ_p.

The model separates two channels through which the center shapes local behavior:

- **Alignment discipline/Autonomy** (β): pulls policy toward the center's priority θ_c and dampens both responsiveness to public beliefs (Corollary 1) and information acquisition (Proposition 1).
- **Responsiveness mandate** (α_c): increases responsiveness to the public directly and increases the value of learning θ_p.

Observed weak policy adjustment to information can reflect either strong center discipline or low autonomy (high β), weak responsiveness mandates (low α_c) and weak intrinsic motivation (low α_ℓ), or high information costs (high c).

3.5 Mapping the model to our experiments

While we explain our experiments in more detail in the next section, it is helpful to preview how each maps into the model.

In the first experiment we set $c \approx 0$, providing a subset of administrators access to information at essentially zero acquisition cost. In the model, this removes the direct barrier to learning, so any deviation in policy response from the public's ideal must come from downstream constraints — limited autonomy (large β) or weak incentives (low α) —which jointly determine how information translates into policy. The experiment therefore isolates whether low updating reflects information frictions per se or constraints on acting on information. If policy responds, information acquisition costs are binding.

The second experiment randomizes c, generating exogenous variation in the marginal cost of acquiring information. This allows us to estimate the elasticity of information effort with respect to acquisition costs. We also introduce the opportunity for costly information avoidance, as in the model extensions in Appendix D: when receiving a signal may increase expected costs or the contents of the signal may directly impact utility, some administrators may pay to avoid receiving a signal altogether.

The final experiment increases the salience of monitoring by making it more likely that θ_p is observable to the center, such that the local has a stronger incentive to internalize the

center's interest to be responsive to the public (greater α_c, unless the center does not care about the public). The basic model assumes that θ_p is observed by the center ex-post, but in practice monitoring may be imperfect or uncertain. By shifting perceived observability, the experiment tests how administrators adjust information use and policy choices when the threat of evaluation is stronger.

4 Experimental Design

To understand the nature of, and potential remedies to, gaps in administrative knowledge and alignment with public priorities, we conduct three experiments with woreda officials. In the first, we provide information via standardized report cards that report statistics on citizen and firm experiences and priorities for public goods in the woreda. This allows us to test for bureaucrats' responsiveness to information. In the second experiment, we randomize the time needed to opt in, or opt out of, receiving future report cards. We use this experiment to estimate bureaucrat demand for information. In the final experiment, conducted with an entirely new sample of woredas, we inform a random subset of woreda officials of research activities being conducted in their woreda, but do not provide any local statistics. We use this experiment to estimate whether the threat of monitoring provokes bureaucrat response. The three experiments are explained in detail below.

4.1 Experiment 1: Report Cards

In our first experiment, we implemented an intervention providing localized data on the community's satisfaction and experience with public services to woreda officials in the form of structured report cards (RC). The sample for this experiment consists of 58 woredas in Addis Ababa and Oromia where we collected representative household and firm survey data. The experimental design, which involved two levels of randomization, is shown in Figure A3.

In a first step, we used stratified sampling to randomly allocate the 58 woredas into RC-treatment and RC-control states. In a second step, within each RC-treatment woreda, we randomized which bureau managers were selected to receive report cards (hereafter bureau

report cards), drawn from the set of health, education, and SME bureaus. Heads for two of these three bureaus were randomly selected to receive a report card in RC-treatment woredas: ten woredas received report cards in the education and SME bureaus (set ES), another ten receive report cards in the education and health bureaus (set EH), and the remaining nine receive report cards in the health and SME bureaus (set HS).

In addition to the bureau report cards, chief administrators in all RC-treatment woredas received four report cards (hereafter administrator report cards). Three of these are fixed across all treatment woredas—water, electricity, and jobs—while the fourth varied according to the woreda's bureau-level randomization. Specifically, woreda administrators in the ES treatment condition received a health report card, those in EH received an SME report card, and those in HS received an education report card.

Each RC-treatment woreda consequently received the same set of six RCs (health, education, SME, water, electricity, jobs) though the RC recipient varied. We refer to the six bureaus covered by the RCs, a subset of all bureaus in the woreda, as our *focus bureaus*.

Report cards were delivered to focus bureau managers first, in Q4 2017. Then, in early 2018, enumerators delivered additional report cards to the chief woreda administrator.[12] Figure A4 presents the full study timeline. We provide details on survey administration in Section 5.

This innovative experimental design introduces variation both in which bureaus are directly targeted at the manager level and in the additional sectors reported to chief administrators. In doing so, we are able to test not only the overall impact of receiving report cards but also differential effects based on their introduction at multiple levels of the administrative hierarchy.

Report Card Construction The report cards were constructed using surveys of households and firms collected in 2016 and 2017. The surveys were designed to be representative of individuals and firms within a particular woreda, such that the information presented was locally relevant to woreda bureaucrats. On the front side, the report card explained the re-

[12]In 10 woredas, we were unable to deliver the report cards to the administrator themselves and instead interacted with a deputy, typically the vice-administrator.

search collaboration between Stanford University and the Ethiopian Development Research Institute (EDRI). EDRI, which has subsequently been reorganized as the Policy Studies Institute, was a government-affiliated research group charged with providing policy recommendations to the prime minister and federal policymakers. EDRI was well known around Addis Ababa, lending credibility to our reports.

Each RC displayed key statistics regarding the public's perceptions, prioritization, and utilization of services in a particular sector. RCs were printed in color on heavy card stock to imitate high-quality reports frequently distributed by NGOs and government agencies. They clearly stated where the data came from, including the sample size for the statistics shared. Appendix B provides additional details on construction and showcases sample report cards. Figure A5 shows the distribution of sample sizes for statistics provided in report cards.

Bureau Report Cards: The education report cards focused on schooling choice and satisfaction with education services by households with children 4-17 years old. Statistics were reported separately for users of private and public education services. The health report card presented summary statistics on wellness and medical care use, again splitting results for public versus private users. The report card additionally reported on firm health inspections and health extension visits. The SME report card presented responses separately for samples of self-employed individuals and local SMEs. These included statistics on sectoral employment and firm characteristics, financing, and hindrances to business growth.

Administrator Report Cards: The water report card reported on water supply, access, and satisfaction separately for firms and households. It highlighted the share of respondents concerned about water safety and availability, as well as firm costs associated with water shortages. The electricity report card reported statistics, again separated by household and firms, on electricity supply, satisfaction, and complaint resolution. As in the water report card, firms also reported estimated costs associated with electricity outages. Finally, the jobs report card presented broad labor market statistics for individuals and firm-level statistics that compared firms in the woreda with others firms in Addis Ababa or Oromia. Firm statistics included future planning, barriers to business growth, and determinants of

location.[13]

After they received report cards, chief administrators were given the opportunity to provide feedback on the information and/or format and content of the report cards. Feedback could be provided via mail, fax, email, or a phone call. Only 2 out of 58 administrators provided feedback.

4.2 Experiment 2: Demand and Information Avoidance

During the Wave 2 survey with chief administrators and bureau heads, in Fall 2019, we embedded an additional experiment allowing the respondents to Opt-In (hereafter O-I) or Opt-Out (hereafter O-O) of receiving report cards in the future. This was done with both RC-treatment and RC-control woredas. It was explained to RC-treatment woredas that these future report cards would contain updated statistics using the second round of household and firm surveys in the woreda, compared to the first round.

Opting in required that respondents fill out a form, of varying length, in order to receive an updated report card. Conversely, opting out required the bureaucrat to fill out a similar form in order to *not* receive a report card. The respondents were told that our team would return in two days to collect the filled out forms. If the form was filled by the time the enumerator returned, respondents in the O-I conditions were ensured that the required report cards would be produced, whereas those in the O-O conditions were ensured that no future report cards would be delivered. Administrators were offered to receive report cards on water, electricity and jobs, while bureau managers were offered the opportunity to receive a report card on their particular bureau.

Rather than estimate a monetary willingness to pay for information, we randomized the form length and consequently the estimated time required to fill out the form. The design

[13]While chief administrator surveys were intended to be completed prior to any bureau report card delivery, in 15 woredas administrators were surveyed after 1 or 2 bureau report cards had been delivered. Conducting these activities out of order may matter if the bureau managers shared the report cards with the administrator prior to the administrator survey, influencing administrator responses. However, the average gap between manager and administrator surveys in these cases was less than one week, minimizing opportunities to share the report cards. Further, questions with administrator during the midline survey suggest that no administrators had heard of the report cards at the time of the survey.

allows us to estimate demand as a function of bureaucrat time.[14] All woreda administrators and bureau managers from the education, health, and SME sectors were randomized into one of five treatment conditions—O-I 5 minutes, O-I 10 minutes, O-I 25 minutes, O-O 5 minutes, O-O 10 minutes—stratifying by region and RC-treatment status. The randomized treatment conditions varied the form length and consequently the estimated time required to fill out the form. Figure A6 displays the experimental design. Enumerator scripts and forms can be found in Appendix C.

4.3 Experiment 3: Monitoring-only Experiment

Receiving an RC is a bundled treatment: it provides woreda administrators with information about citizen preferences and service gaps while simultaneously informing them that data is being collected from citizen and firms, and analyzed by a research team with permission from the government. To disentangle the respective importance of these two mechanisms, we separately conducted a "monitoring only" experiment in which woredas were told that research was being conducted, but no RC was produced.

This experiment was conducted with a separate set of 89 woredas—all woredas in Addis Ababa not included in the RC experiment sample.[15] After stratifying by sub-city and total woreda budget, a proxy for population, we followed a randomization procedure identical to the report card experiment. Treated woreda chief administrators and two bureau heads received monitoring announcements that indicated that our team would be conducting research in their woreda. The monitoring announcements contained a description of the project and details about the partnership between the Ethiopian Development Research Institute and Stanford University. Chief administrators and bureau heads were left with the same contact information and project description as was included on the front side of the full report cards.

The two bureaus were again randomly selected from the set of education, water, and SME bureaus, generating comparable ES, EH, and HS treatment conditions. The monitoring announcements were designed to look similar to the report cards, while not containing

[14] We chose not to estimate WTP in cash because we did not want to take funding out of the woredas budgets or pressure salary constrained government officials for payments.

[15] These 89 woredas were those that did not share a border with the city boundary.

any specific statistics. Refer to Appendix B for detailed enumerator scripts and a sample monitoring announcement.

5 Data

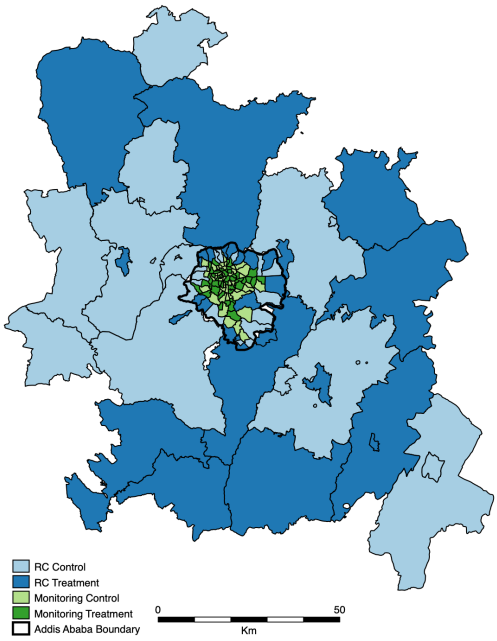
We combine survey and administrative data in our analysis. The survey data—covering households, firms, and public officials—were collected through Stanford's African Urbanization and Development Research Initiative (Abebe et al., 2018). We supplement these surveys with woreda-level administrative budget data. Below, we describe the sampling strategy and the contents of each data source in more detail. Figure A4 summarizes the data collection and experimental timeline.

5.1 Woreda Sampling

Addis Ababa is the capital city of Ethiopia, and sits as an independent entity within the surrounding Oromia region. The majority of Ethiopia's urban growth in recent decades has been concentrated in Addis Ababa's urban periphery: the portion of the Oromia region immediately surrounding Addis Ababa was reclassified as the Oromia Special Zone (OSZ) in 2008 reflecting the area's importance as a hub for economic growth and urbanization. Our primary sample consists of the full set of Addis Ababa woredas sharing a border with the Addis Ababa city limits as well as all of the urbanized woredas in the OSZ. The sample of woredas includes all substantial towns within 20km of Addis Ababa and nearly all towns within 50km.

The primary sample covers 6 sub-cities and 28 woredas within the Addis Ababa city boundary; and 12 zones in the OSZ which encompass 26 woredas. Figure 5 shows the location of our sample. We survey bureaucrats and collect budget data for all woredas in Addis Ababa: 58 woredas in the primary sample and 89 woredas used in the monitoring-only experiment.

Figure 5: Study area map



5.2 Surveys with Administrators and Bureau Managers

We conducted surveys of both chief administrators in the woreda, and sector-level bureaus within each woreda (e.g. finance, health, education).[16]

The first round of bureaucrat surveys was a baseline in 2017. The survey measured knowledge of key facts about bureau performance, first-order and second-order beliefs on priorities for the woreda and bureaus, as well as bureaucrats' characteristics. Table A4 provides baseline summary statistics on woreda chief administrators and bureau heads. Notably, these officials are young, with an average age of 35 for administrators and 32 for bureau heads; nearly all are male and hold a college degree. While most of the bureaucrats have worked in the public sector for over a decade, reorganization and turnover in the administration means that both chief administrators and bureau heads had only served in their position for a year at the time of the survey, on average. We will return to the implications of turnover

[16]The full list of bureaus surveyed in each woreda was: civil service and human resources, finance, land management and housing, health, education, trade, roads, revenue, micro and small enterprises, and water and sanitation. We also surveyed the municipal infrastructure bureau in Oromia woredas, when applicable. While this was not the exhaustive set in each woreda, it represents the set of bureaus that were consistent across woredas.

in Section 6.[17]

Woreda chief administrators were interviewed for a midline survey in early 2018, shortly after the bureau RCs had been delivered, and just before the additional RCs were delivered directly to the administrator. The midline survey asked questions similar to the baseline.

Finally, both woreda chief administrators and bureau heads were surveyed at endline, between late 2019 and early 2020.

5.3 Household and Firm Surveys

Households surveys were collected in each woreda across two waves: Wave 1 (W1) took place between late 2016 and early 2017 while Wave 2 (W2) took place between late 2018 and early 2019. To sample a representative set of households, we first conducted a household listing, proportional to the estimated population of each kebele within sampled woredas.[18] The listing targeted between 5% of households in less populated kebeles and 0.5% of households in the most populated kebeles, grouping kebeles using empirical discontinuities in the population distribution. Households were listed using a snowball sampling methodology within each kebele, bypassing immediate neighbors at a pre-determined rate to ensure geographic coverage. Through this protocol, we ultimately listed 11,698 households across all study woredas.

Following the listing exercise, we randomly sampled individuals for in-depth surveys based on kebele population, stratified by gender. In general, we sampled at most one individual per household, though in a few cases multiple household members were interviewed in low-population kebeles. Our final sample for in-depth interviews consists of 3,618 individuals, approximately 35% of each kebele-level strata. Compared to a city-wide representative survey conducted by the Ethiopian Central Statistical Service (CSA), households in our sample are smaller, more mobile, better educated, but earn substantially lower monthly

[17] Table E2 compares woreda official characteristics to the labor force in the greater Addis Ababa area. Despite being more educated and reporting working longer hours, woreda officials earn 23% less than the typical worker in this setting. This wage gap is only partially explained by non-pecuniary benefits received by bureaucrats. Nevertheless, bureaucrats report substantially higher levels of job satisfaction.

[18] Since Ethiopia has not completed a census since 2007, population estimates were obtained by combining estimates from the Central Statistical Service and woreda administrators.

wages (Table A6).

Firm surveys were similarly conducted across two waves in 2017 and 2019. As no comprehensive list of firms existed in this context, we combined data sources separately from each region to build a sampling frame. In Oromia, we built near exhaustive lists of firms from administrative data and sampled firms from each woreda proportional to the total number of firms located in the woreda. In Addis Ababa, we constructed a representative sampling frame by combining data from a partial firm census conducted by the Addis Ababa Labor and Social Affairs Bureau and lists from the Addis Ababa Investment Office, Addis Ababa Industry Office, Ethiopian Development Research Institute, and Federal Micro and Small Enterprise Agency. Firms were sampled proportional to the number in each woreda, as in Oromia, and supplemented with additional firms operating in the same sector found via snowball sampling. Further details on the firm sampling can be found in Abebe et al. (2018) and Abebe et al. (2025).

As shown in Abebe et al. (2025), surveyed firms were slightly smaller and less likely to be foreign-owned, relative to representative World Bank Enterprise Surveys. These differences are primarily due to the woredas in our sample being located away from the city center, with fewer large and foreign-owned firms.

5.4 Budget Data

We obtained all woreda-level budget and expense data for fiscal years 2014 to 2020. We use these to study the impact of the report cards on total budget size, since woredas have the discretion to lobby for funding and apply for supplementary allocations from the central government, and on budget allocation across bureaus. The budget data include line-item budgets for the overall administration and for each bureau within the administration. We have this data for all woredas in Addis Ababa and each sampled woreda in Oromia. Tables E3 and E4 display summary statistics for these data for Addis Ababa and Oromia, respectively. Recall that there is substantial heterogeneity in bureau-level budget shares across woredas as displayed in Figure 2 above.

6 Results

This section first presents the results of the main report card experiment and the demand experiment. We then conduct a series of tests for mechanisms and heterogeneity, where we report the results of the monitoring-only experiment.

6.1 Report Card Experiment

For the report card experiment, we focus our results on four sets of pre-specified outcomes: bureaucrat knowledge of public priorities and service utilization, alignment with public priorities, citizen satisfaction with services, and woreda (bureau) budget allocations. Each of these outcomes is defined at the bureau level, providing variation within each woreda, supporting our primary estimating equation:

$$y_{w(s)t} = \beta_0 + \beta_1 \text{RC-Bureau}_{w(s)} + \mathbf{X}_w \kappa + \gamma_s + y_{w(s)0} + \varepsilon_{w(s)t} \tag{9}$$

for some outcome y for sector bureau s located in woreda w at time t. When possible, we estimate ANCOVA models, controlling for the baseline outcome measure, to improve power McKenzie (2012). In Equation 9, "RC-Bureau" is an indicator variable equal to one when a bureau received a report card and β_1 estimates the average effect of the bureau-level treatment on the outcome. The vector $\mathbf{X}_w$ includes woreda and administrator controls or woreda fixed effects; γ_s are sector fixed effects (e.g. education, health, SME).

When outcomes are observed for bureaus other than our six focus bureaus, we estimate Equation 10, allowing for differential treatment effects for non-focus bureaus in treated woredas (δ_1) and treated bureaus (δ_2).[19] "RC-Woreda" is an indicator variable equal to one for all bureaus in RC-treated woredas, while "RC-Bureau" is defined as in Equation 9, again

[19] Recall that our experiment only created and delivered report cards to a subset of bureaus in RC treated woredas. Non-focus bureaus—roads, finance, trade, land management and housing, revenue—never received RCs even in RC treated woredas. However, we still surveyed bureau heads and observe key outcomes for these bureaus. We observe budget allocations for the full set of bureaus in every woreda.

varying at the bureau level.

$$y_{w(s)t} = \delta_0 + \delta_1 \text{RC-Woreda}_w + \delta_2 \text{RC-Bureau}_{w(s)} + \mathbf{X}_w\kappa + \gamma_s + y_{w(s)0} + \varepsilon_{w(s)t} \tag{10}$$

Administrator Knowledge: To begin, we test whether RCs increase chief administrator knowledge of public priorities or utilization of public services. These outcome measures correspond to key statistics displayed directly in the RCs and vary by sector. To construct outcomes, we calculate the absolute difference between the share of citizens in a woreda considering a particular sector as a top 5 priority or using a particular public service net of the administrators estimates of these statistics.

Results in Table 3 show that administrators failed to update their knowledge of public priorities and service utilization in response to the report cards. We estimate precise null effects for each outcome measure separately 4 months and 24 months after the intervention. Columns (1), (2), (5), and (6) estimate effects on knowledge of public priorities at midline and endline, respectively. Similarly, columns (3), (4), (7), and (8) estimate the same for knowledge of service utilization. The null result cannot be explained by generally improved knowledge, as we cannot reject equality between baseline and followup outcome measures in the control group.

There are also null effects on sector-specific (education, health, and SME) knowledge of report card statistics (Table A7). Together, the results are consistent with chief administrators failing to internalize the information provided in the report cards.

Second-order Beliefs and Misalignment: Table 4 reports the treatment effects on the misalignment in chief administrator second-order beliefs about public priorities and, separately, between the administration's and the public's priorities. To measure administrator second-order beliefs about public priorities, we asked them to rank the top 5 services or improvements that they expected the public to demand. The administration's own top 3 priorities were elicited directly. We use the absolute difference between the administrator's rankings and the public's as our measures of misalignment.

Table 3: Null effect on woreda chief administrator knowledge of RC-targeted indicators

	Gap in WCA knowledge of:							
	4 months after baseline				24 months after baseline			
	Share of citizens prioritizing a sector		Share of citizens using public services		Share of citizens prioritizing a sector		Share of citizens using public services	
					as of Wave 1	as of Wave 2	as of Wave 1	as of Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RC-Bureau	-0.01	0.00	-0.03	-0.02	-0.02	0.03	0.01	0.02
	[0.06]	[0.05]	[0.03]	[0.04]	[0.06]	[0.07]	[0.03]	[0.02]
Control mean	0.45	0.45	0.26	0.26	0.44	0.42	0.24	0.21
Midline controls	No	No	No	No	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Woreda FEs	No	Yes	No	Yes	No	No	No	No
N	112	112	170	168	112	112	170	170

Notes: The unit of observation is the woreda-sector. All outcomes are absolute differences between admin estimates and data from the citizens' survey. Columns 1-2 and 5-6 show treatment effects on the difference between admins' estimates of the share of citizens that prioritize a sector and the true share; this is available for only Education and Health. For example, the control mean of 0.45 in (1) indicates that, on average, admins in control woredas mis-estimate the share of citizens considering the sector as a top 5 priority by 45 ppt. The rest of the columns show effects on the difference between admins' estimates of the share of citizens using public services and the true services coverage; this is available for Education, Health, and SME. All true values were shared in the report card, calculated using the baseline citizen survey. Columns 1–4 compare admins' midline estimates to baseline citizens' data. Columns 5 and 7 compare admins' endline responses to baseline citizens' data while columns 6 and 8 compare to endline citizens' data. All models control for region. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Administrators are not completely uninformed: relative to a benchmark of random guessing, administrators second-order beliefs and priority alignment with the public are more accurate. However, misalignment in both outcomes is still substantial.

Columns (1) and (3) of Table 4 show that report cards had no impact on the accuracy of administrator second-order beliefs, on average, either 4 or 24 months after baseline. Column (4) shows that, if anything, administrator second-order beliefs about the public's prioritization of treated sectors (as of the second wave of household surveys) were slightly less accurate. As only the public's Wave 1 rankings were reflected in RCs, we interpret this result as RCs failing to improve overall administrator knowledge of public priorities in the medium-run.

The results for the administration's priorities, relative to the public's, are reported in Columns (2), (5), and (6). If anything, misalignment marginally *increased* as a result of the report cards.[20] Finally, administrators do not differentially expect treated sectors to be public priorities, and we can rule out systematic updating by administrators who over- or

[20]We consider alternative measures of misalignment in Table E5 and Table E6. The results are consistent with a null effect: chief administrators are no more likely to correctly identify public priority sectors, or to have priorities that align with the public's following treatment.

Table 4: Report cards do not improve WCA misalignment with citizen priorities

	Misalignment between:					
	4 months after baseline		24 months after baseline			
	WCA perceptions & citizen priorities	WCA's own & citizen priorities	WCA perceptions & citizen priorities		WCA's own & citizen priorities	
			as of Wave 1	as of Wave 2	as of Wave 1	as of Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
RC-Woreda	0.43**	0.40**	-0.23	0.02	0.08	0.00
	[0.19]	[0.15]	[0.20]	[0.22]	[0.20]	[0.19]
RC-Bureau	0.17	0.41	0.79**	0.59*	0.13	0.44
	[0.49]	[0.36]	[0.36]	[0.33]	[0.36]	[0.35]
RC-Woreda + RC-Bureau = 0	0.246	0.050	0.129	0.045	0.482	0.124
Control mean	3.49	3.28	3.78	3.35	3.78	3.45
Random-guess benchmark	5.14	5.49	5.13	4.95	5.48	5.30
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	812	812	812	812	812	812

Notes: The unit of observation is the woreda-sector. All outcomes are absolute differences between admin estimates and data from the citizens' survey. Columns 1, 3, and 4 show treatment effects on the difference between admins' beliefs about citizens' priorities and the true priorities. Admins' beliefs were measured by the rank they assigned top 5 sectors when asked to identify citizens' priorities. The rest of the columns show effects on the difference between their own (top 3) priorities and citizens priorities. Citizen priorities were determined based on the rank individuals gave 5 sectors in terms of priority. These rankings averaged, and re-ranked from 1 (lowest priority) to 14 (highest priority) at the woreda level. All sectors that were unranked by admins were given a value of 6 for the second-order belief outcome and 4 for the woreda priority outcome. For example, the control mean of 3.49 in (1) means that, on average, admins in control woredas misjudge citizens' ranking of a top priority sector by 3.49 positions. All models control for region and baseline misalignment. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

under-estimated a sector's true public ranking (Table A8).

Public Satisfaction: Using our second round of surveys with households and firms, we test for treatment effects on measures of satisfaction with public services. Results are presented in Table 5. We consider two outcomes: the share of citizens reporting to be satisfied or very satisfied with a sector and the share of citizens who believed a sector improved over the prior two years. Unsurprisingly given the lack of impact on bureaucrats' beliefs and priorities, we find null effects on all outcomes. This is consistent with the fact that treated administrators were, by their own account, no more likely to have taken steps to improve water or electricity provision (Figure A7a); there were similarly null effects on the likelihood that public complaints were addressed, as reported by citizens (Figure A7b).

Table 5: No improvement in public satisfaction or opinion

	% citizens satisfied with sector	% believed sector improved
	(1)	(2)
RC-Woreda	-0.00	-0.05
	[0.03]	[0.04]
RC-Bureau	0.01	-0.04
	[0.04]	[0.03]
RC-Woreda + RC-Bureau = 0	0.714	0.047
Control mean	0.50	0.47
Baseline controls	Yes	No
Sector FEs	Yes	Yes
N	348	522

Notes: The unit of observation is woreda-sector. The outcome data from the endline citizen survey. Column 1 shows treatment effects on the share of individuals who report being satisfied or very satisfied with service provision in a specific sector. Column 2 shows effects on the share of people who report that a sector has improved over the past two years. No baseline data is available for the outcome in column 2. All models control for region. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.1.1 Budget

In our final set of outcomes for the report card experiment, we consider woreda budget allocations. This data is detailed, including line-item budgets and expenditures for the years 2015 - 2019 for all woredas and all bureaus. There are two budget outcomes, approved and adjusted. The former reflects the initial budget allocation approved by the sub-city or region while the latter reflects the budget after accounting for supplementary allotments or deductions.

The Report Card treatment had no effect on approved or adjusted total woreda budgets (Table A9). The next question is whether the report cards led to budget reallocation across sectors within a woreda. To leverage the full data, we adjust Equation 10 to allow for multiple post-treatment periods and estimate

$$y_{w(s)t} = \delta_0 + \delta_1 \text{RC-Woreda}_w \times \text{Post}_t + \delta_2 \text{RC-Bureau}_{w(s)} \times \text{Post}_t + \mathbf{X}_w \kappa + \gamma_{st} + y_{w(s)0} + \varepsilon_{w(s)t} \tag{11}$$

where "Post" is a dummy for the 2018 and 2019 budget years such that we estimate average effects across post-treatment years; γ_{st} are sector-by-year fixed effects. As in Equation 10

we control for the pre-treatment outcome measure and add time fixed effects.

Results for Equation 11 are presented in Table 6. We consider two outcomes types – total bureau budget and the bureau's share of the woreda's budget – for both originally approved and subsequently adjusted budgets. We estimate precise null treatment effects of receiving a report card in all specifications indicating that budgets were not reallocated within woredas in response to treatment.

Table 6: Impacts on woreda-sector budget

	(Log) appr. bureau budget		Appr. bureau budget share		(Log) adj. bureau budget		Adj. bureau budget share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RC-Woreda × Post	-0.003	0.016	0.001	0.001	0.013	0.026	-0.002	-0.002
	[0.055]	[0.050]	[0.002]	[0.002]	[0.057]	[0.052]	[0.002]	[0.002]
RC-Bureau × Post	-0.033	-0.032	-0.006	-0.006	-0.005	-0.004	0.002	0.002
	[0.049]	[0.050]	[0.004]	[0.004]	[0.053]	[0.054]	[0.004]	[0.004]
RC-Woreda + RC-Bureau = 0	0.451	0.746	0.086	0.082	0.821	0.517	0.847	0.863
Control mean	14.11	14.11	0.07	0.07	14.20	14.20	0.07	0.07
Woreda FEs	No	Yes	No	Yes	No	Yes	No	Yes
Sector × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,895	3,895	3,932	3,932	3,931	3,931	3,932	3,932

Notes: The unit of observation is the woreda-sector-year. Funds earmarked for *General woreda administration* are excluded. The outcome data span 2015–2019. *Post* is defined as 2018 and onwards. All models control for total approved woreda budget and region. Standard errors in brackets, clustered by woreda in Panel A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For the budget data, we additionally estimate event-studies using the imputation estimator from Borusyak et al. (2024). Specifically, we estimate the following equation:

$$y_{w(s)t} = \sum_{\substack{j=\{-2,2\} \\ j\neq 0}} \gamma_j \text{RC-Bureau}_{w(s)t-j} + \alpha_{w(s)} + \alpha_t + \mathbf{X}_{w(s)t}\gamma + \varepsilon_{w(s)t} \tag{12}$$

where $\text{RC-Bureau}_{w(s)t-j}$ indicates the period in which a bureau received a report card. We also consider extended specifications that allow for heterogeneous treatment effects for RCs delivered to bureau managers versus chief administrators.

The event study results confirm that RC-treated woredas experience no change in overall budgets (Figure A8). They also confirm that the RC-treatment had no effect on how the woreda budget was allocated across bureaus, whether as initially approved or adjusted, and no heterogeneity based on whether report cards were delivered to chief administrators or bureau heads (Figure A9 and Figure A10).

6.2 Demand Experiment

Bureaucrats may not have responded to the report cards if they were deemed irrelevant, whether due to the sectors covered or the specific statistics included. To investigate this, we test bureaucrat demand for receive additional information if allowed to choose the sectors and types of data for their report cards. Conversely, we also test whether they would opt *out* of receiving additional cards for certain sectors if given the option.

As described in Section 4, our experiment cross-randomized the opt-in and opt-out conditions with time (cost) it took to fulfill the conditions of the experiment. Doing so allows us to estimate demand curves for additional information in terms of bureaucrat time.

Shares of bureaucrats either opting-in or out in each treatment condition are show in Table 7. In the Opt-in condition, participants in the 5-minute arm show the highest demand, with 37% completing the opt-in form. Demand drops sharply to 4% in the 10-minute arm and stays low or declines further to 2% in the 25-minute arm. However, overall demand is very low at 15%. This suggests that the content of the report cards cannot fully explain the null effects in the report card experiment. Even when given the option to receive information in sectors of their choosing (5-minute arm) or to select specific data points (10- and 25-minute arms), both chief administrators and bureau heads show little interest.[21]

[21] There are no statistically different differences in demand between chief administrators and bureau heads, as show in Appendix Table E7.

Table 7: Demand experiment results

	Completed form			
	Opt-in		Opt-out	
	Mean	N	Mean	N
5 min	0.37	46	0.02	47
10 min	0.04	46	0.07	46
25 min	0.02	44		
Total	0.15	136	0.04	93

Notes: The sample includes both woreda chief administrators (WCAs) and bureau managers. The outcome variable equals 1 if a complete form was turned in. We group those who submitted incomplete forms with those who didn't submit at all. In the 5-minute opt-in condition, participants listed sectors they wanted a report card in and could indicate if a higher-level authority should receive it. In the 10-minute opt-in condition, they also specified up to 8 statistics or outcomes they wanted information on. In the 25-minute condition, they listed 16 statistics and noted any neighboring woredas they wanted to learn about, or, if uninterested, explained why. In the 5-minute opt-out condition, they listed sectors they didn't want a report card in. In the 10-minute opt-out condition, participants were asked to also explain their reasons.

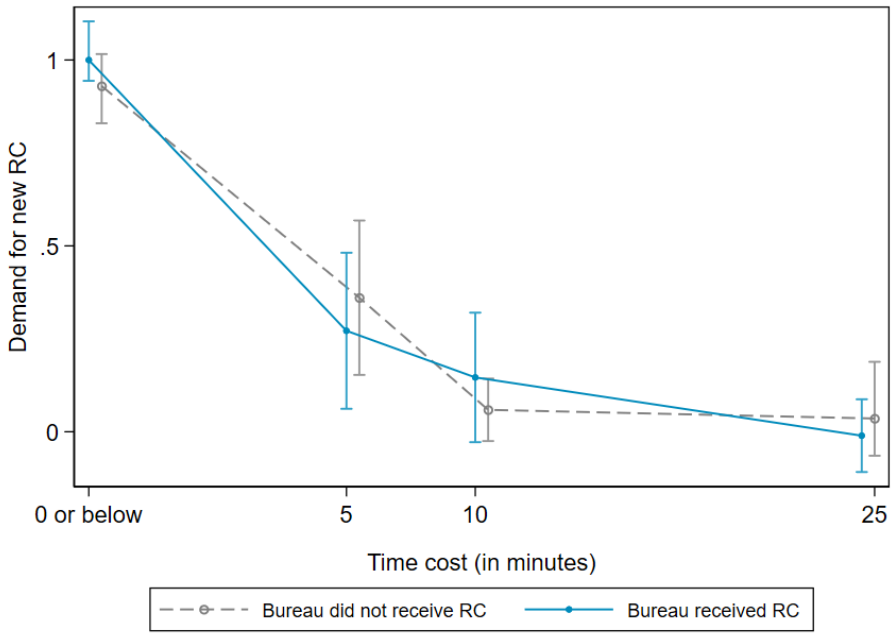
Interestingly, a small but non-zero share of bureaucrats are willing to "pay" at least 5 minutes of time to *not* receive future report cards. We hypothesize that the possibility of receiving documentation indicative of poor performance leads some bureaucrats to demand less information. This result highlights the interaction between information frictions and bureaucrat incentives: even in an environment where bureaucrat *could* be better informed, some may choose not to be if they expect to receive bad news.

Next, we test for RC treatment effects in the demand for information. Bureaucrat demand for report cards as a function of the time cost needed to obtain them are presented in Figure 6, separately for woreda-sectors that did or did not receive RCs in W1. The arc-elasticity in the RC treatment group is indistinguishable from that in the control group, and is estimated at -1.21, suggesting either implausibly high bureaucrat time valuations or, more plausibly, highly elastic demand for information; the receipt of an RC did not change demand for information.

Given bureaucrat turnover between our survey rounds, we explore heterogeneity by bureaucrat tenure in Table E8 within the sub-sample of bureaucrats in RC-treated woredas. We

show that demand is largely comparable between new bureaucrats, who may not have seen the first round of report cards if information is not effectively transmitted across individuals within a given post, and incumbents. This further supports the conclusion that report card content cannot explain the null results. Demand for information simply appears to be low.

Figure 6: WTP for new report cards by **woreda-sector** treatment status



Notes: The figure plots predicted values of local leaders' (WCAs and bureau heads) demand for a new report card at different time costs (in minutes) to receive information. Price of "0 or below" represent the opt-out conditions (5- and 10-minute forms), where leaders received information by default (zero cost) but could submit a form to opt-out (imposing a "negative" cost). Estimates are adjusted for woreda and sector fixed effects, with standard errors clustered at the woreda level. "Demand" is defined as submitting a completed opt-in form or failing to return a complete opt-out form. Treatment status is defined at the woreda-sector level. The solid line represents leaders who had previously received a report card, and the dashed line represents leaders from control woredas or treated woredas who did not receive the bureau report cards.

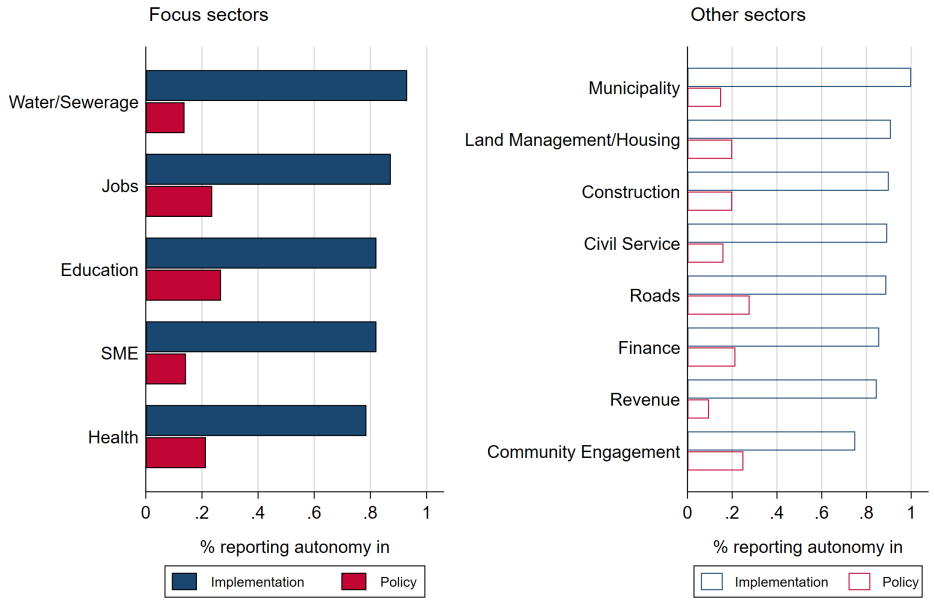
7 Barriers to Responsiveness

Having shown that bureaucrats neither internalize nor respond to information contained in the report cards, and that demand for information is low, we investigate the factors that may explain these results.

7.1 Lack of Autonomy

One potential explanation for the negligible impacts of the report cards and low demand for information could be bureaucrats' lack of autonomy. Even though Ethiopian policy has formally decentralized decision making authority to the woredas, *defacto* authority may still reside with more central bodies. If this were the case, woredas may not have the ability to respond to information about public preferences and demand for this information would be low. To our knowledge, the realities of Ethiopian decentralization have not been rigorously studied, particularly in our peri-urban context. In what follows, we first provide descriptive evidence on markers of autonomy in the data we collected, and then test for heterogeneity in treatment effects by autonomy level.

Figure 7: Reported autonomy of local bureaucrats over policy and implementation decisions



Notes: The unit of observation is the bureau manager. The sectors in solid color are focus bureaus for which report cards were issued in the treatment group.

To start, we note that Ethiopian law differentially delegates responsibility to woredas for specific sectors. This results in variation in the influence that woreda officials have over sector-level policy and implementation. While the share of WCAs reporting having autonomy

to determine sectoral policy varies between 10 and 20% across sectors, the share reporting having autonomy over policy implementation ranges between 45 and 85% (Figure 7). Variation in the shares of woreda budgets, documented in Figure 2 above, is consistent with woredas having budgetary autonomy in key sectors.

Second, as one of the explicit goals of Ethiopian decentralization policies was to improve government responsiveness to local conditions, there exist channels for woreda action even in sectors over which they have relatively minimal formal control. Woreda officials can lobby for project funding or sector-specific budget supplements. We designed our survey to ask about these actions as we expected them to be margins along which woreda officials might respond to the treatment.

We begin by revisiting the main RC experiment, estimating Equation 9 and Equation 10, and testing for treatment effect heterogeneity based on sectoral autonomy. If low autonomy in some sectors is driving our null results, we would expect treatment effects to be concentrated in high autonomy sectors.

Table 8 shows no difference in treatment effects by autonomy level, across a battery of outcomes. Woreda officials are no more likely to be knowledgeable about or aligned with treated high autonomy sectors. Similarly, there are no differential effects on citizen satisfaction or sectoral budget shares. This suggests that little meaningful action was taken in treatment sectors, even in sectors in which the woreda has the authority to implement policy.

7.2 Accountability and Effort

Next, we consider bureaucrat effort and their incentives to be accountable to public demands. If the returns to effort are low or bureaucrat incentives fail to promote alignment with the public, this could explain our experimental results. Bureaucrats report high levels of effort and job satisfaction: respondents claim to work over 50 hours per week, spend approximately 10% of these hours interacting with the public, and over 60% are satisfied with their job (Table 9). Notably, these are self-reported measures and may be affected by desirability

Table 8: Heterogeneity by baseline autonomy

	Gap in WCA knowledge of:		Misalignment between:			
	% citizens prioritizing a sector as of Wave 2	% citizens using public services as of Wave 2	WCA's perceptions and citizen priorities as of Wave 2	WCA's own and citizens' priorities as of Wave 2	% citizens satisfied with sector	Budget share
	(1)	(2)	(3)	(4)	(5)	(6)
RC-Bureau	-0.018	0.014				
	[0.083]	[0.029]				
RC-Bureau × High autonomy	0.115	0.021				
	[0.130]	[0.042]				
RC-Woreda			0.092	0.106	-0.018	-0.003
			[0.422]	[0.358]	[0.049]	[0.003]
RC-Woreda × High autonomy			-0.176	-0.376	0.050	-0.003
			[0.721]	[0.535]	[0.075]	[0.005]
RC-Bureau			0.470	0.617	0.049	-0.004
			[0.475]	[0.537]	[0.056]	[0.005]
RC-Bureau × High autonomy			0.337	-0.349	-0.106	0.009*
			[0.680]	[0.666]	[0.066]	[0.005]
Treatment × High autonomy = 0	0.379	0.614	0.883	0.195	0.185	0.212
Control mean	0.44	0.24	3.34	3.46	0.45	0.09
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	112	170	580	580	290	301

Notes: Data from the endline admin and citizens survey. The unit of observation is the woreda-sector. *High autonomy* equals 1 if, at baseline, the WCA reported having some or a lot of influence over policy in that sector. Outcomes in (1)-(4) are absolute differences. The dependent variable in these columns is the gap between admin and citizen values, with negative coefficients indicating improved knowledge (columns 1-2) or alignment (columns 3-4). All models control for region and baseline values. In columns 1 and 2, all sectors with available outcome data in the woreda were treated, i.e., there are no 'spillover' woreda-sectors. Robust standard errors are shown in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

bias. Taken at face value, these reports suggest that bureaucrats invest substantially in their jobs and are content with their positions. But the RC treatment does not impact any of these measures of effort or job satisfaction (Table 9). There is also no heterogeneity in treatment effects for low versus high-performing bureaus.[22]

The monitoring experiment was conducted to tease out the role of outside monitoring alone in the RC-treatment impact, if any. Positive treatment effects on measures of effort or budgetary responses in the monitoring experiment would indicate that internal accountability systems failed to maximize bureaucrat performance. Since we did not conduct household or firm surveys for the set of woredas in this experiment, we are only able to test for treatment effects on self-reported bureaucrat outcomes and budget allocations.

[22]Note that the positive coefficient on RC-Woreda in Column (4) of Table 9 implies a substantial increase in the share of time interacting with the public in woredas with few low performing sectors. However, the overall effect is indistinguishable from zero for the average woreda.

Table 9: Effect on admin and manager effort by sector performance

	Work hours		% work hours interacting with public		Satisfied with job	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: WCAs						
RC-Woreda	-4.22	5.65	0.03	0.09**	-0.00	0.25
	[2.62]	[7.81]	[0.02]	[0.05]	[0.13]	[0.30]
RC-Woreda × # low performing sectors		-3.02		-0.02		-0.07
		[2.32]		[0.01]		[0.09]
Control mean	65.62	65.62	0.09	0.09	0.69	0.69
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
N	58	58	58	58	58	58
Panel B: Managers						
RC-Woreda	-0.17	-0.16	0.04	0.04	-0.04	-0.04
	[1.40]	[1.40]	[0.03]	[0.03]	[0.09]	[0.09]
RC-Bureau	-0.59	-1.20	-0.02	-0.02	0.07	0.14
	[1.41]	[1.81]	[0.02]	[0.02]	[0.08]	[0.10]
RC-Bureau × Low performing		1.11		-0.00		-0.13
		[2.37]		[0.04]		[0.11]
Control mean	54.74	54.74	0.09	0.09	0.61	0.61
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	516	516	515	515	516	516

Notes: Data from the endline admin and bureau manager surveys. The unit of observation is the local leader (WCA or bureau manager). In Panel A, *# low performing sectors* refers to the total number of low-performing sectors in a woreda. A woreda-sector (bureau) is considered low-performing if it falls below the median on at least two out of three performance indicators, based on citizen survey data. In Panel B, *Low performing* equals 1 if the sector meets the same criteria described above for Panel A. All models control for region and baseline values. Models in Panel B include a dummy flag for bureaus without performance data (i.e., those tagged by *RC-Woreda*). Standard errors are shown in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Unsurprisingly, given the results of the report card experiment, the monitoring announcements had no impact on measures of bureaucrat effort or sectoral budgets. Results in Table 10 show precise nulls on measures of administrator hours working with the public and total work hours using Equation 9. The monitoring treatment also had no impact on sectoral budgets or a sector's share of the woreda budget (Table E9). Together, the results support the conclusion that third-party monitoring does not influence bureaucrat behavior or effort.

Table 10: Monitoring-only treatment: Effects on leader effort

	Work hours	% work hours interacting with public	Satisfied with job
	(1)	(2)	(3)
Panel A: WCAs			
RC-Woreda	-2.66	0.01	0.14
	[2.26]	[0.02]	[0.10]
RC-Woreda + RC-Bureau = 0	0.242	0.701	0.151
Control mean	70.09	0.10	0.60
Baseline controls	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes
Woreda FEs	No	No	No
N	89	89	89
Panel B: Managers			
RC-Woreda	-1.62	0.03*	0.04
	[1.42]	[0.02]	[0.05]
RC-Bureau	-0.52	0.00	0.06
	[1.64]	[0.02]	[0.07]
RC-Woreda + RC-Bureau = 0	0.253	0.115	0.152
Control mean	60.42	0.07	0.62
Baseline controls	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes
Woreda FEs	No	No	No
N	707	706	706

Notes: The unit of observation is the local leader: the woreda chief administrator in Panel A and bureau manager in Panel B. All outcomes were self-reported by official directly. All models control for region and baseline values. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Even if measured accurately, measures of self-reported time use (hours worked and time interacting with the public) may not reflect bureaucrat effort if working hours are not used efficiently. We now consider tangible indicators of bureaucrat effort. Table A10 shows that that treated woredas did not increase an index measure of citizen engagement such as the frequency of site visits, reception of community feedback, or the number of sources of community feedback in either the RC or monitoring experiments.[23]

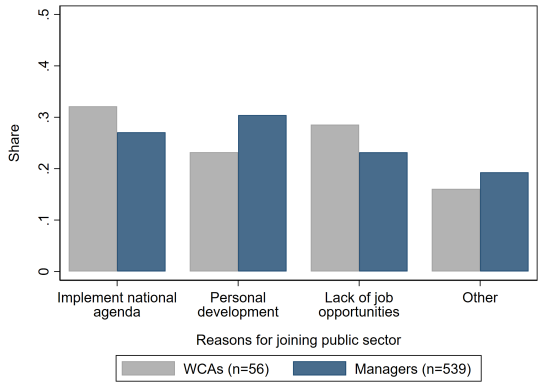
We also find no consistent evidence of increased lobbying efforts in RC-treated sectors: while there is a small increase in the probability that RC-treated bureaus requested more intra-woreda funds, there is a decrease in the probability of woreda officials requesting additional funds for treated bureaus (Table A11). This holds at the woreda-level: treated woredas were no more likely to submit supplementary budget requests, proposals targeting

[23]RC-treated woredas increase the frequency at which reports are shared between bureaus and chief administrators; the probability that a bureau reports at least bi-weekly increases by 73% though these effects are not concentrated in RC-treated bureaus. Thus, RCs may marginally improve intra-woreda information sharing, even if this reporting does not translate to key downstream outcomes. The monitoring announcements failed to impact any of these measures.

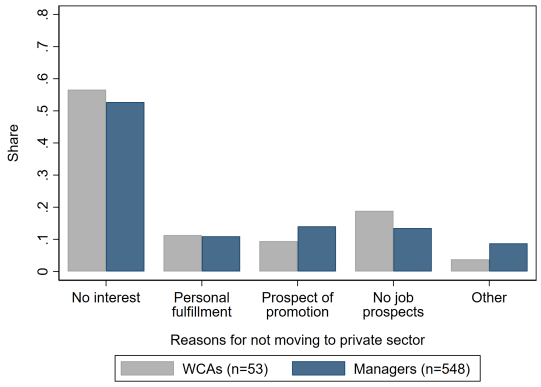
municipal projects, or informally negotiate their budget with higher authorities (Table A12).

Figure 8: Proxies for bureaucrats' intrinsic motivation at baseline

(a) Reasons for joining public sector



(b) Reasons for staying in public sector



Notes: The figures show the shares of WCAs and managers who reported different reasons for entering (Panel A) and staying in (Panel B) the public sector during the baseline surveys. We test whether the distributions differ between officials in the treatment and control groups, and find no significant differences.

Together, these results are consistent with an environment in which bureaucrats have few incentives to expend effort or improve the alignment of their woreda with public priorities. While a full characterization of the woreda incentive structure is saved for future work, Figure 8 shows bureaucrats self-reported reasons for joining and continuing to work in the civil service. If bureaucrats display sufficient intrinsic motivation towards meeting public demands, this motivation could compensate for a lack of direct incentives. We observe that less than 1/3 of respondents report joining for reasons related to mission alignment or motivation to service; respondents are equally likely to have joined due to a lack of jobs elsewhere or for personal development. For chief administrators interested in private sector employment, most stay in the civil service due to a lack of outside job prospects; bureau managers are more likely to mention the prospect of promotion as the reason they stay in the public sector. We take these qualitative findings to suggest that mission alignment, or broader intrinsic motivation towards public service, are unlikely to replace incentives as drivers of responsive public services in this context.

7.3 Turnover

Our final test of mechanisms investigates the role of bureaucrat turnover in explaining our results. If woredas have limited institutional memory, such that new bureaucrats enter with little information, then we would expect misalignment when turnover rates are high. Table 11 shows that 74% of local officials in the sample started their position following the baseline surveys.[24] We further note that at endline, around 50% of bureaucrats in treated woredas mentioned a report card during their survey. We take this as evidence of potential information frictions within the woreda.

Table 11: Turnover and institutional memory at endline

	WCAs	Managers	Total
Panel A: Turnover			
New since baseline	0.74	0.74	0.74
Months in position at baseline	17.14	16.96	16.98
Months in position at endline	17.00	18.51	18.36
N	58	523	581
Panel B: Institutional memory			
If treated: mentioned previous RCs	0.44	0.51	0.49
N	25	53	78

Notes: The table uses endline values only. In Panel A, woreda chief administrators (WCAs) and bureau managers are considered new if their reported tenure is less than the time since the baseline surveys. The outcome in Panel B was observed for demand experiment participants only.

High turnover could mask treatment effects concentrated within the set of bureaucrats who remained in their position. Therefore, in Table 12 we return to our primary estimating equations, allowing for heterogeneity by bureaucrat tenure. However, we again find no evidence of differential treatment effects for new versus existing woreda officials. If anything, returning administrators show marginally larger knowledge gaps and misalignment than new administrators. This further supports our conclusion that reducing information frictions alone is insufficient to address the misalignment of Ethiopian woredas.

[24] While we do not have reliable measures of historical turnover, we believe turnover rates in our study period to be abnormally high as it overlapped with a federal regime change. Treatment had no effect on the probability that a woreda official remained at their job by endline. New administrators look similar to tenured administrators across most observable characteristics (Appendix Table E10). While we cannot reject the joint test of no differences, new administrators have lived longer in the woreda and have a lower desired time in their position, on average.

Table 12: Heterogeneity by turnover status

	Gap in WCA knowledge of:		Misalignment between:			
	% citizens prioritizing a sector as of Wave 2	% citizens using public services as of Wave 2	WCA's perceptions and citizen priorities as of Wave 2	WCA's own and citizens' priorities as of Wave 2	% citizens satisfied with sector	Budget share
	(1)	(2)	(3)	(4)	(5)	(6)
RC-Bureau	0.248*	0.028				
	[0.140]	[0.041]				
RC-Bureau × New WCA	-0.291*	-0.006				
	[0.158]	[0.048]				
New WCA	0.039	-0.017	0.104	-0.289	-0.025	0.006
	[0.100]	[0.034]	[0.274]	[0.265]	[0.028]	[0.004]
RC-Woreda			0.103	0.051	0.053	0.007
			[0.312]	[0.413]	[0.042]	[0.006]
RC-Woreda × New WCA			-0.108	-0.050	-0.071	-0.003
			[0.396]	[0.457]	[0.048]	[0.006]
RC-Bureau			0.867	0.366	-0.061	-0.003
			[0.528]	[0.654]	[0.045]	[0.007]
RC-Bureau × New admin			-0.359	0.099	0.096**	-0.008
			[0.527]	[0.673]	[0.044]	[0.008]
Treatment × New WCA = 0	0.071	0.903	0.663	0.989	0.104	0.110
Control mean	0.44	0.24	3.35	3.45	0.50	0.08
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	112	170	812	812	348	415

Notes: Data from the endline admin survey. The unit of observation is the woreda-sector. The *New WCA* indicator equals 1 if the admin started working after the baseline survey. The outcome in (1)-(2) is the absolute gap (in ppt) between admins' and the true estimates. In (3) and (4), it is the absolute gap (in ranks) between admins and citizens. All models control for region and baseline values. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Conclusion

Our paper documents a striking disconnect between the theoretical promise of decentralization and its actual implementation in practice. In a low-income country with high levels of formal decentralization, Ethiopia, we find that local bureaucrats possess limited knowledge of local conditions and public preferences. Through our experiments and surveys, we demonstrate that this information deficit reflects not merely a data collection problem, but a fundamental lack of demand for such information within the bureaucratic structure.

Despite receiving detailed, actionable information about local needs and preferences—and despite possessing at least some discretionary authority to direct policy in response—bureaucrats showed minimal interest in acquiring this information and exhibited virtually no behavioral response when it was provided. This pattern suggests that local information is simply not a

relevant input into bureaucratic decision-making, even when decentralization reforms have in principle created the institutional space for responsive governance.

One likely explanation is that, in the Ethiopian context we consider, bureaucrat positions function primarily as employment opportunities rather than as vehicles for public service, characterized by low compensation and recruitment processes far less selective than the competitive civil services found in some other contexts. Indeed, the majority of bureaucrats in our sample report entering public service primarily due to the absence of viable private sector alternatives, rather than from a commitment to public administration.

The incentive structure reinforces this dynamic if promotion decisions are only weakly tied to performance metrics, while political affiliation is a salient factor in career advancement. Moreover, top-down directives from higher levels of government systematically override signals of public demand, further attenuating any incentive for bureaucrats to invest in understanding or responding to local preferences. In this environment, gathering information about citizen needs represents an investment with minimal expected returns.

These findings carry important implications for our understanding of decentralization reforms and for policy design more broadly. The *de jure* transfer of formal authority to lower levels of government is insufficient to generate responsive governance. Absent clear performance incentives that reward responsiveness to local conditions, and without meaningful mechanisms for public accountability that empower citizens to influence bureaucratic behavior, the promise of decentralization is unlikely to be realized. Decentralization may be a necessary condition for locally responsive governance, but our evidence suggests it is far from sufficient.

References

Abebe, Girum, Daniel Agness, Pascaline Dupas, Marcel Fafchamps, Tigabu Getahun, and Deivy Houeix, "Urban Development in Africa: Preliminary report on the Addis Ababa AUDRI study," Technical Report, Technical report 2018.

_ , Stefano Caria, Pascaline Dupas, Marcel Fafchamps, and Tigabu Getahun, "Competition and Management Upgrading: Experimental Evidence from Ethiopia," Technical Report, National Bureau of Economic Research 2025.

Agness, Daniel and Tigabu Getahun, "Housing and Human Capital: Condominiums in Ethiopia."," 2024.

Ahmad, Junaid, Stephen Commins, Shantayanan Devarajan, Deon Filmer, Jeffrey Hammer, Lant Pritchett, Ritva Reinikka, Shekhar Shah, and Agnès Soucat, "World Development Report 2004: Making services work for poor people," *Washington, DC : World Bank Group.*, 2003.

Aneja, Abhay and Guo Xu, "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," *American Economic Review*, August 2024, *114* (8), 2352–87.

Arora, Abhishek, Siddharth George, Nivedita Mantha, Vijayendra Rao, and MR Sharan, "(When) Do Politicians Govern Better than Bureaucrats? Evidence from a Natural Experiment in India," September 2024. Unpublished manuscript, draft.

Azfar, Omar, Jeffrey Livingston, and Patrick Meagher, "Decentralization in Uganda," *Decentralisation and local governance in developing countries: a comparative perspective*, 2006, pp. 223–55.

Badin, María Inés, Diana Moreira, Juan Francisco Santini, and Monica Singhal, "Politicians and Tax Policy: The Role of Preferences and Beliefs," *National Tax Journal*, 2025, *78* (2), 000–000.

Banuri, Sheheryar, Stefan Dercon, and Varun Gauri, "Biased policy professionals," *The World Bank Economic Review*, 2019, *33* (2), 310–327.

Bardhan, Pranab, "Decentralization of governance and development," *Journal of Economic Perspectives*, 2002, *16* (4), 185–205.

Bazzi, Samuel, Masyhur Hilmy, Benjamin Marx, Mahvish Shaukat, and Andreas Stegmann, "It Takes a Village Election: Turnover and Performance in Local Bureaucracies," Technical Report, National Bureau of Economic Research 2025.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu, "The Glittering Prizes: Career Incentives and Bureaucrat Performance," *The Review of Economic Studies*, 05 2019, *87* (2), 626–655.

Besley, Timothy and Anne Case, "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits," *The Quarterly Journal of Economics*, 1995, *110* (3), 769–798.

_ , Robin Burgess, Adnan Khan, and Guo Xu, "Bureaucracy and development," *Annual Review of Economics*, 2022, *14* (1), 397–424.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi, "Individuals and Organizations as Sources of State Effectiveness," *American Economic Review*, August 2023, *113* (8), 2121–67.

Bhavnani, Rikhil R and Alexander Lee, "Local embeddedness and bureaucratic performance: evidence from India," *The Journal of Politics*, 2018, *80* (1), 71–87.

Bó, Ernesto Dal, Frederico Finan, Nicholas Y Li, and Laura Schechter, "Government decentralization under changing state capacity: Experimental evidence from Paraguay," Technical Report, National Bureau of Economic Research 2018.

Boffa, Federico, Amedeo Piolatto, and Giacomo AM Ponzetto, "Political centralization and government accountability," *The Quarterly Journal of Economics*, 2016, *131* (1), 381–422.

Bongwa, Aloysius, Samson Kassahun, and Meine Pieter van Dijk, *Decentralization and service delivery in Ethiopia*, Shaker Publishing BV, 2011.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Brookman, David E. and Christopher Skovron, "Bias in Perceptions of Public Opinion among Political Elites," *American Political Science Review*, 2018, *112* (3), 542–563.

Butler, Daniel M, David W Nickerson et al., "Can learning constituency opinion affect how legislators vote? Results from a field experiment," *Quarterly Journal of Political Science*, 2011, *6* (1), 55–83.

Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee, "Data and policy decisions: Experimental evidence from Pakistan," *Journal of Development Economics*, 2020, *146*, 102523.

Faguet, Jean-Paul, "Decentralization and Governance," *World Development*, 2014, *53*, 2–13. Decentralization and Governance.

_ , Qaiser Khan, and Devarakonda Priyanka Kanth, "Decentralization's Effects on Education and Health: Evidence from Ethiopia," *Publius: The Journal of Federalism*, 06 2020, *51* (1), 79–103.

Finan, F., B.A. Olken, and R. Pande, "Chapter 6 - The Personnel Economics of the Developing State," in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 2 of *Handbook of Economic Field Experiments*, North-Holland, 2017, pp. 467–514.

Gadenne, Lucie and Monica Singhal, "Decentralization in developing economies," *Annu. Rev. Econ.*, 2014, *6* (1), 581–604.

Garcia, Marito and Andrew Sunil Rajkumar, *Achieving better service delivery through decentralization in Ethiopia*, Vol. 132, World Bank Publications, 2008.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 2000, *108* (5), 874–904.

Gemechu, Dereje Terefe, Meheret Ayenew Warota, Geert Bouckaert, and Bacha Kebede Debela, "Meritocracy, Career Development, and Promotion in Ethiopian Civil Service," in "Public administration in Ethiopia: Case studies and lessons for sustainable development," Leuven University Press, 2020, pp. 299–324.

Hatfield, John William, "Federalism, taxation, and economic growth," *Journal of urban economics*, 2015, *87*, 114–125.

Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini, "How research affects policy: Experimental evidence from 2,150 brazilian municipalities," *American Economic Review*, 2021, *111* (5), 1442–1480.

Kiringai, Jane Wangui, Michael Tobias Geiger, Mesfin Girma Bezawagaw, and Leif Jensen, "Ethiopia public expenditure review," Technical Report, The World Bank 2016.

Kosec, Katrina and Tewodaj Mogues, "Public Investment Choices by Local and Central Governments," *The World Bank Economic Review*, 11 2019, *34* (Supplement˙1), S52–S57.

_ **and** _ , "Decentralization without democracy," *World Politics*, 2020, *72* (2), 165–213.

Liaqat, Asad, "No Representation Without Information: Politician Responsiveness to Citizen Preferences," Technical Report, SSRN Working Paper 2023.

Maaskant, K., "Youth employment in urban Ethiopia," in E. W. Chapman and M. Vinez, eds., *Working Today for a Better Tomorrow in Ethiopia: Jobs for Poor and Vulnerable Households*, Washington, DC: World Bank, 2023, chapter 3, pp. 55–74.

McKenzie, David, "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 2012, *99* (2), 210–221.

Oates, Wallace E, "Fiscal federalism," *New York*, 1972, *1.*

Pereira, Miguel M, "Understanding and reducing biases in elite beliefs about the electorate," *American Political Science Review*, 2021, *115* (4), 1308–1324.

Prud'homme, Rémy, "THE DANGERS OF DECENTRALIZATION," *The World Bank Research Observer*, 08 1995, *10* (2), 201–220.

Rogger, Daniel and Ravi Somani, "Hierarchy and Information," *Journal of Public Economics*, 2023, *219*, 104823.

Tiebout, Charles M, "A pure theory of local expenditures," *Journal of political economy*, 1956, *64* (5), 416–424.

Toma, Mattie and Elizabeth Bell, "Understanding and increasing policymakers' sensitivity to program impact," *Journal of Public Economics*, 2024, *234*, 105096.

Ujhelyi, Gergely, "Civil service reform," *Journal of Public Economics*, 2014, *118*, 15–25.

Vivalt, Eva, Aidan Coville et al., "Local knowledge, formal evidence, and policy decisions," *Journal of Development Economics*, 2025, *174*, 103425.

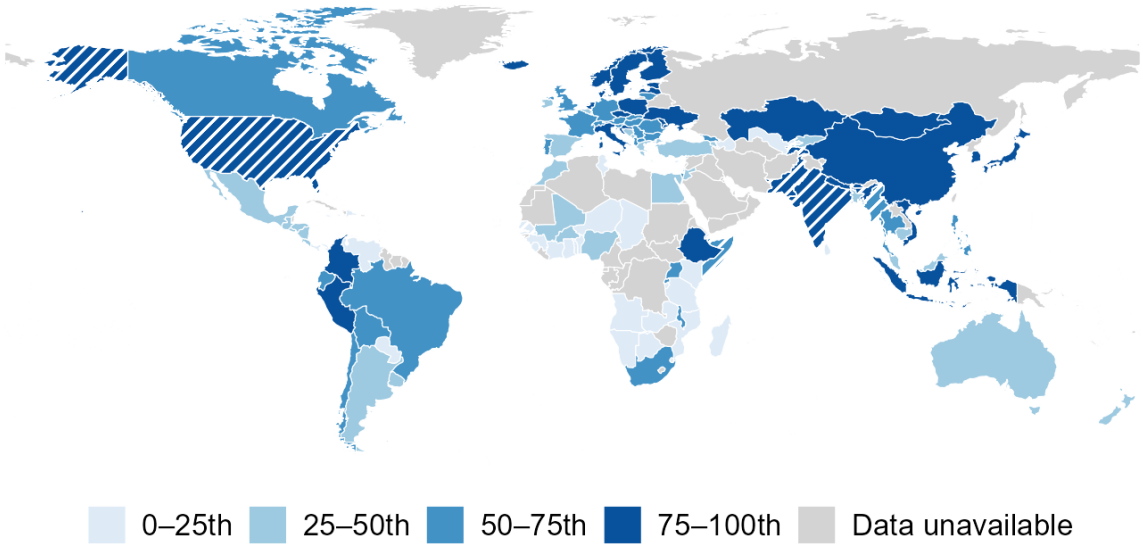
_ **and** _ , "How do policymakers update their beliefs?," *Journal of Development Economics*, 2023, *165*, 103121.

Wallis, John Joseph and Wallace E Oates, "Decentralization in the public sector: An empirical study of state and local government," *Fiscal federalism: Quantitative studies*, 1988, *5.*

Yimenu, Bizuneh, "Measuring and explaining fiscal de/centralization: Empirical evidence from Ethiopia, 1995–2020," *Public Administration and Development*, 2023, *43* (3), 245–259.

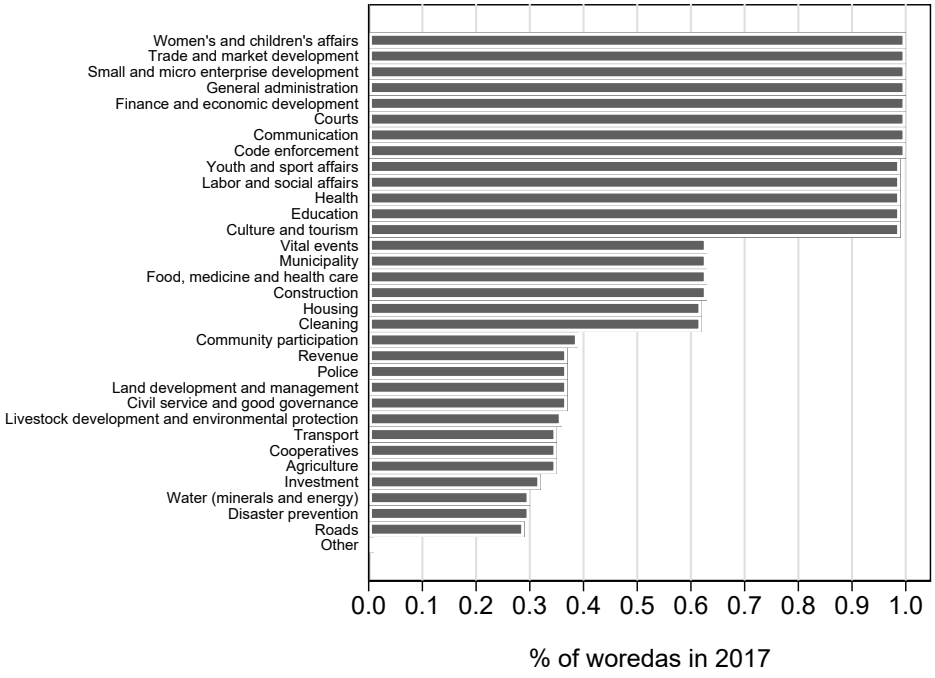
A Appendix Figures and Tables

Figure A1: Share of general government expenditures executed locally



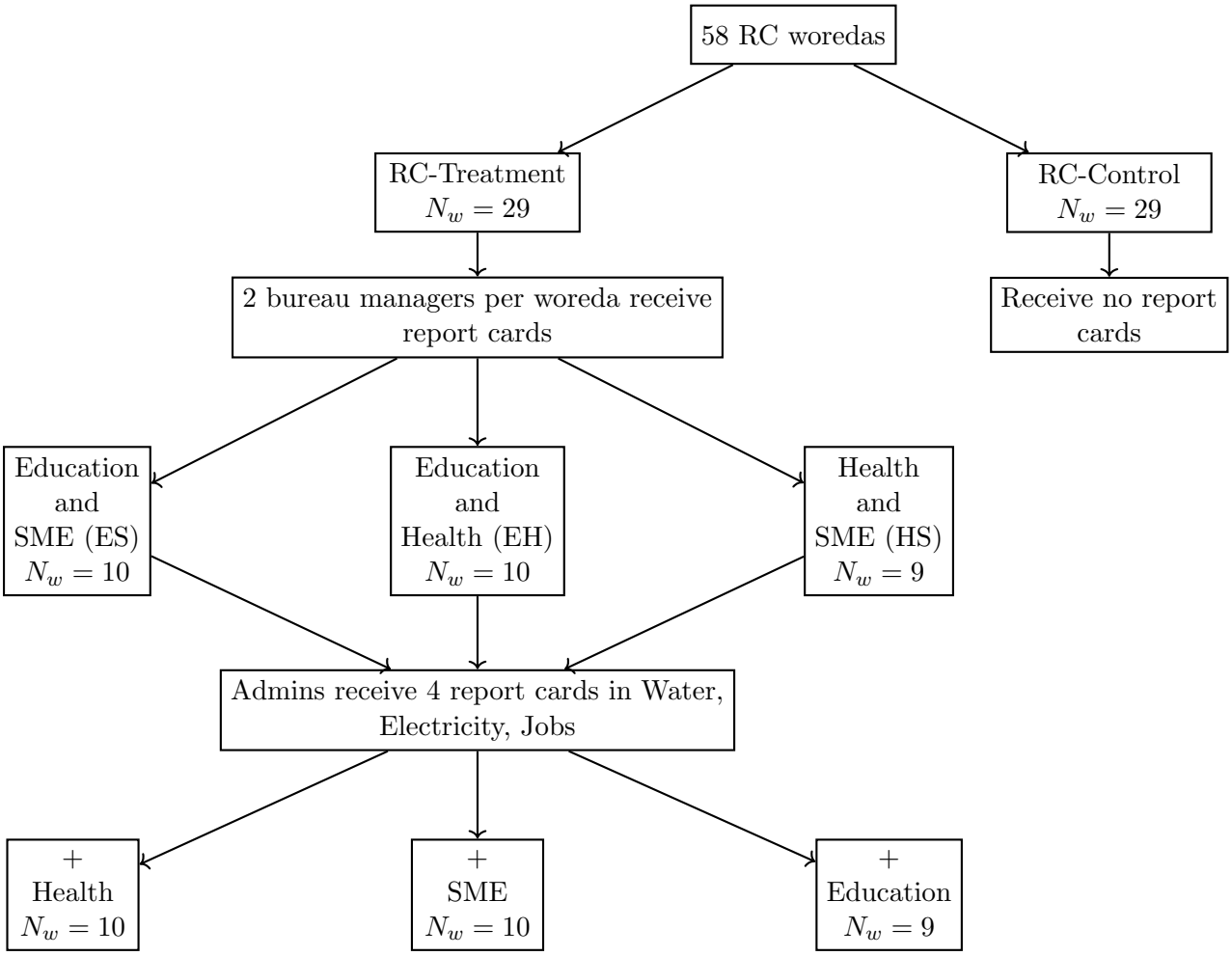
Notes: The figure shows the share of general government expenditure executed at the local government level, grouped into quartiles across countries. Darker shades represent higher shares of the national budget executed at the local level. The local government level includes municipalities, provinces, and other subnational entities whose authority is limited to specific parts of the national territory. Countries shown with white crosshatching correspond to cases where the lowest tier with available data is the state or regional level rather than the local level. Data are drawn from the OECD–UCLG World Observatory on Subnational Government Finance and Investment (SNG-WOFI) for 2020, except for Ethiopia. See Table E1 for data sources for Ethiopia.

Figure A2: Representation of bureaus across study woredas at baseline



Notes: The figure shows the distribution of bureaus across study woredas, as proxied by the expense labels for the 2016/17 fiscal year.

Figure A3: Experimental design for the RC experiment



Notes: N_w denotes the number of woredas.

Figure A4: Study Timeline

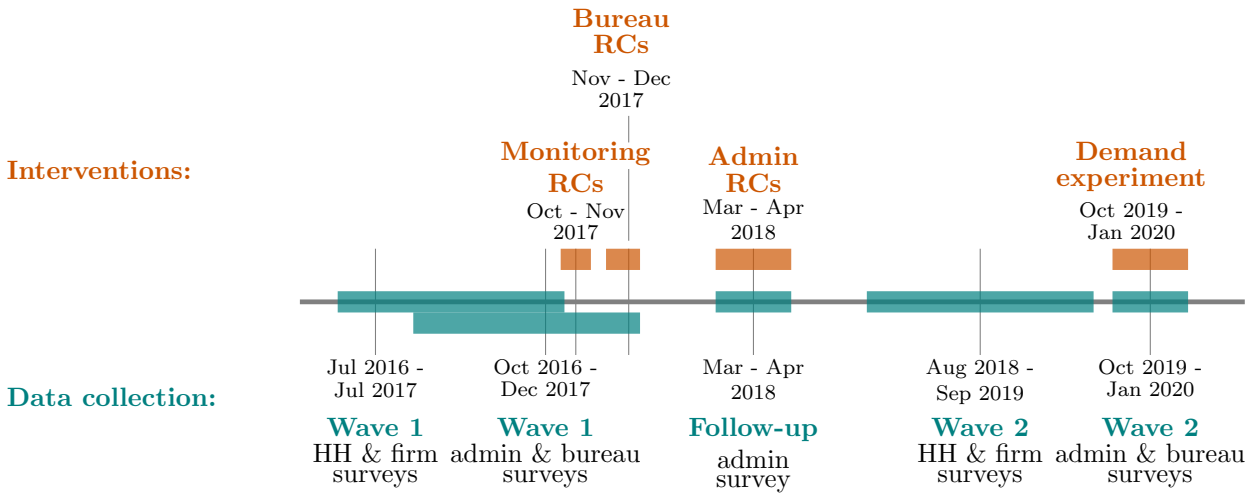
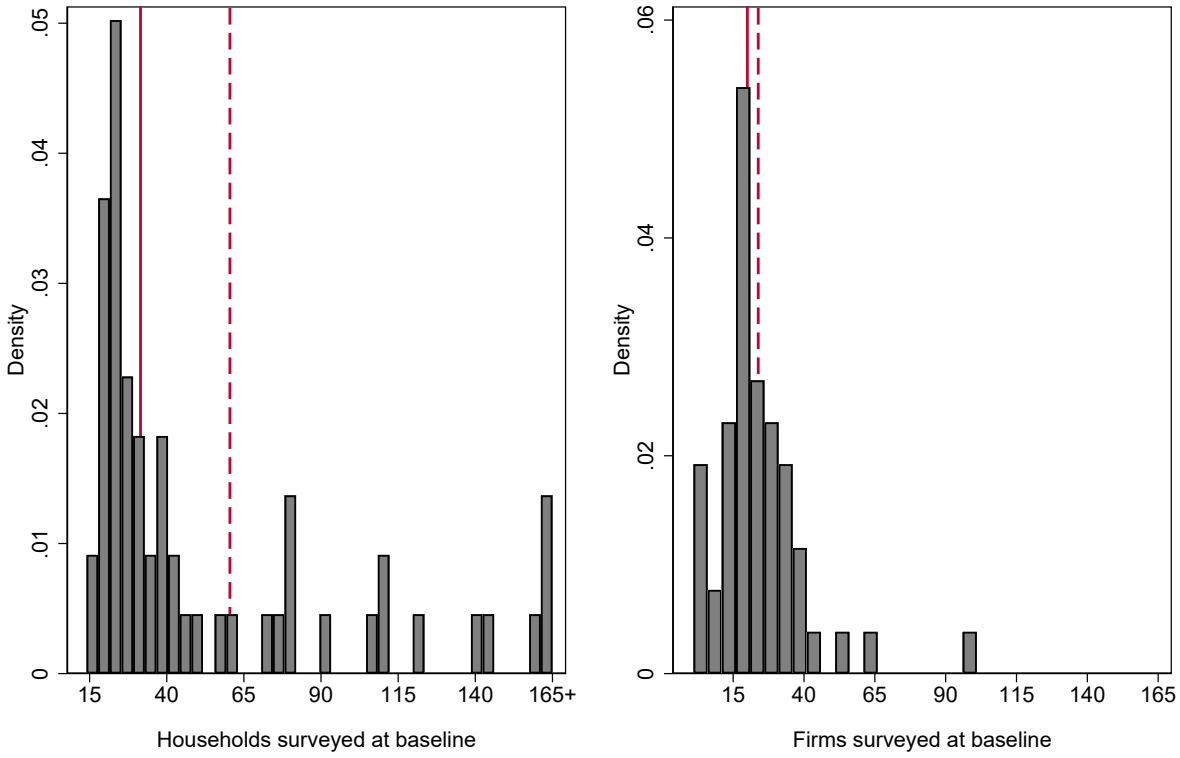


Figure A5: Number of sampled households and firms at baseline



Notes: The dashed lines indicate the mean and the solid lines, the median.

Figure A6: Experimental design for the demand experiment.

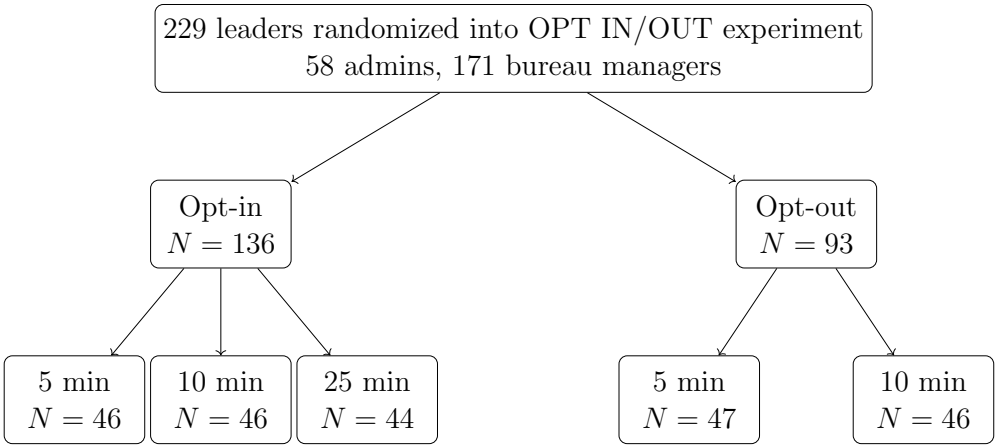
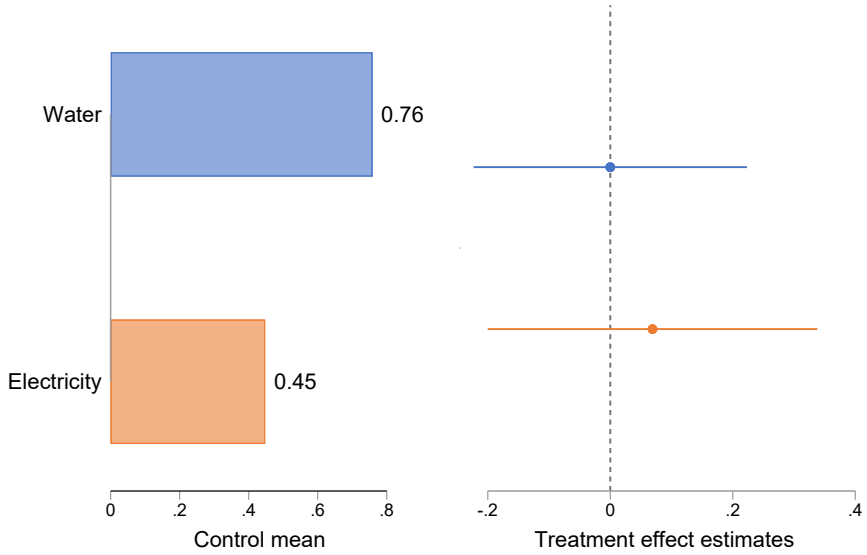
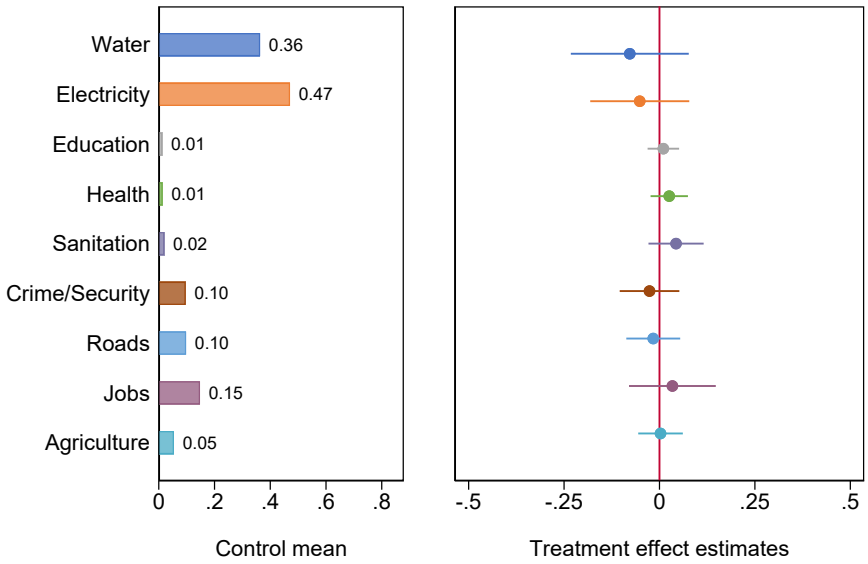


Figure A7: Null effect on local responsiveness

(a) Share of admins who've taken steps to improve basic services provision

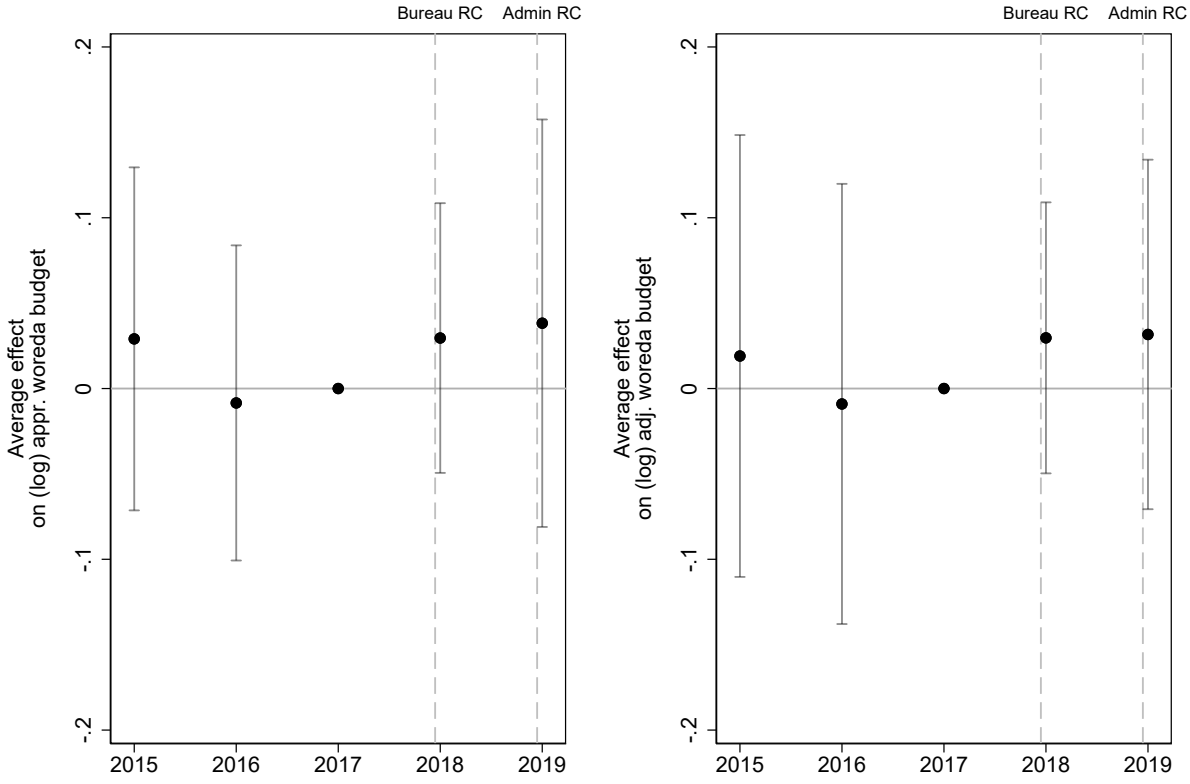


(b) Share of at least partly resolved complaints for basic services provision



Notes: Each coefficient is from a separate regression. In the top panel, admins were asked whether they had taken specific actions to improve the reliability of (i) water and (ii) electricity services. The bottom panel uses data from the citizen surveys. The outcome is the share of citizens who've reported some or complete improvement after lodging a complaint regarding water or electricity services. Controls include region and baseline values of the outcome. Standard errors clustered by woreda.

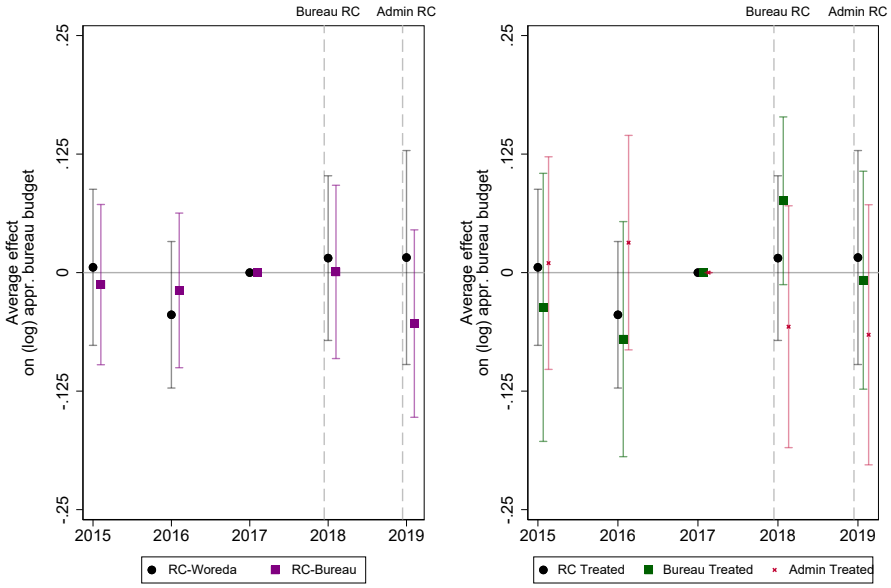
Figure A8: Effects on total woreda budget



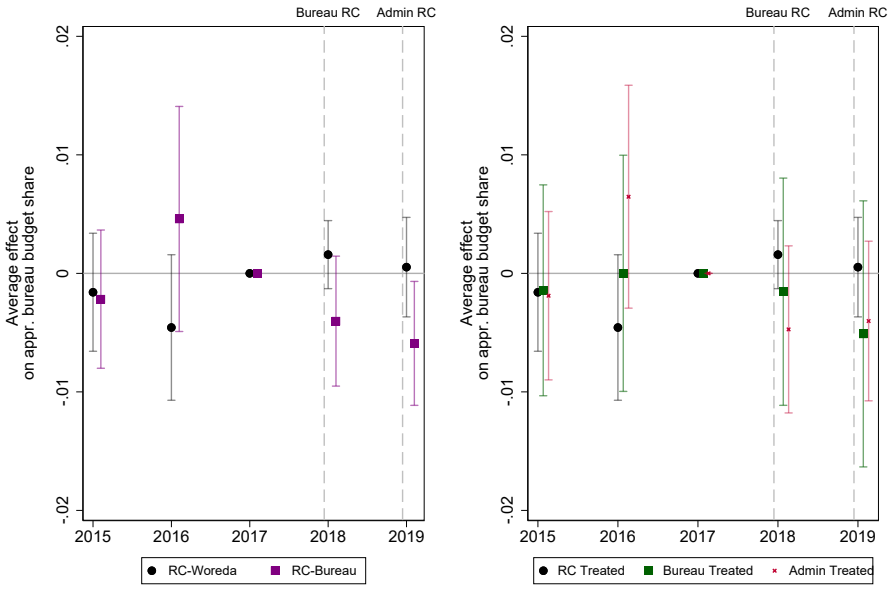
Notes: The figure shows event-study estimates of the report card treatment effect on woreda budgets. Each point represents the average effect in a given year relative to 2017, with 95% confidence intervals. Year and region fixed effects are included.

Figure A9: Effect on approved woreda-sector budget

(a) (Log) approved bureau budget



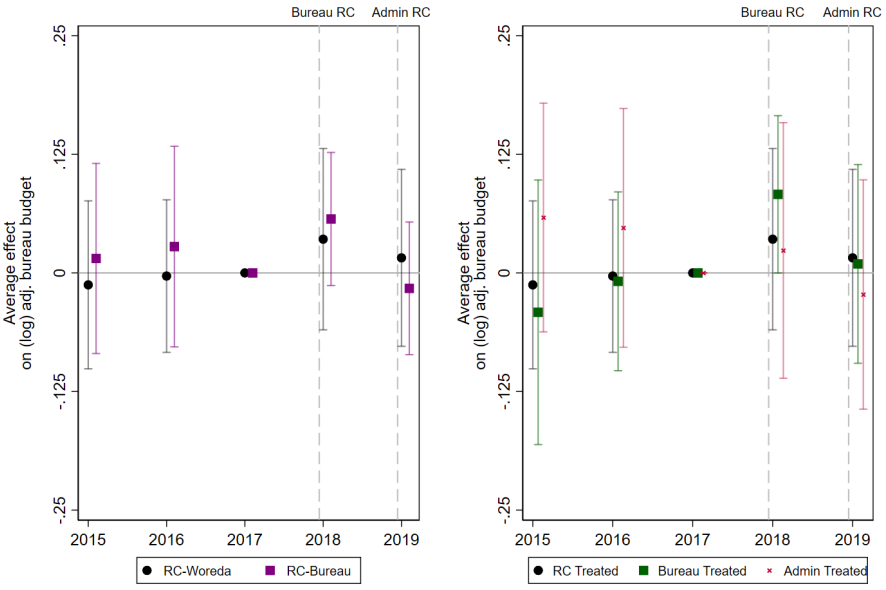
(b) Approved bureau budget share



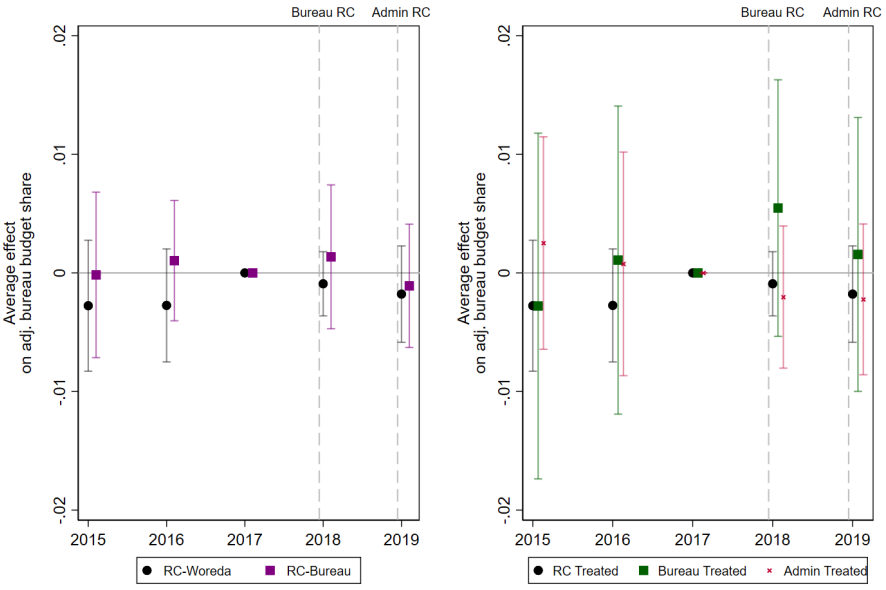
Notes: The figure shows event-study estimates of the report card treatment effect on *approved* woreda–sector budgets. Each point represents the average effect in a given year relative to 2017, with 95% confidence intervals. All models include woreda–sector fixed effects, sector–year fixed effects, and controls for region and total woreda budget. Panel (a) plots effects on the log of approved bureau budgets, Panel (b) on bureau budget shares. In each panel, the left plot estimates constant effects across Bureau and Admin RC treatments. On the right, treatment effects for Bureau and Admin RCs are estimated separately. Vertical dashed lines mark the introduction of Bureau and Admin RCs.

Figure A10: Effect on adjusted woreda-sector budget

(a) (Log) adjusted bureau budget



(b) Adjusted bureau budget share



Notes: The figure shows event-study estimates of the report card treatment effect on *adjusted* woreda–sector budgets. Each point represents the average effect in a given year relative to 2017, with 95% confidence intervals. All models include woreda–sector fixed effects, sector–year fixed effects, and controls for region and total woreda budget. Panel (a) plots effects on the log of adjusted bureau budgets, Panel (b) on bureau budget shares. In each panel, the left plot estimates constant effects across Bureau and Admin RC treatments. On the right, treatment effects for Bureau and Admin RCs are estimated separately. Vertical dashed lines mark the introduction of Bureau and Admin RCs.

Table A1: Accountability and performance incentives at baseline

	WCAs only		
	Treatment mean	Control mean	p-value
Share of work hours spent interacting with citizens	0.14	0.16	0.554
Receives at least monthly public feedback/opinion	0.83	0.79	0.743
Minimum monthly bureau evaluations	0.45	0.66	0.117
Uses complaint metrics to evaluate bureaus	0.54	0.62	0.525
Uses delivery/quality metrics to evaluate bureaus	0.79	0.93	0.118
Poor performing managers face severe consequences	0.62	0.55	0.601

Notes: Data from the baseline woreda chief administrator (WCA) survey.

Table A2: Baseline misalignment between WCA and citizen priorities by sector

	Aligned	Underestimated		Overestimated	
	%	%	Avg. gap	%	Avg. gap
Panel A: Admin beliefs v. citizen priorities					
Electricity	0.60	0.19	2.5	0.21	3.5
Jobs	0.55	0.12	2.9	0.33	4.6
Health	0.43	0.10	2.2	0.47	4.1
Water	0.33	0.60	6.2	0.07	2.5
Public transport	0.29	0.57	3.9	0.14	3.6
Security	0.28	0.47	3.1	0.26	5.4
Housing	0.21	0.02	2.0	0.78	5.6
Education	0.21	0.59	4.0	0.21	3.7
SME	0.19	0.05	2.3	0.76	7.2
Roads	0.19	0.67	5.9	0.14	3.0
Telecom	0.05	0.00		0.95	7.5
Financial services	0.00	0.00		1.00	8.8
Price of food	0.00	0.00		1.00	8.8
Recreational services	0.00	0.00		1.00	9.7
Panel B: Admin v. citizen priorities					
Security	0.55	0.29	3.0	0.16	4.2
Water	0.41	0.47	6.8	0.12	3.0
Public transport	0.36	0.53	3.4	0.10	4.8
Telecom	0.36	0.09	4.2	0.55	4.3
Housing	0.34	0.43	3.7	0.22	5.5
Roads	0.29	0.57	5.6	0.14	3.5
SME	0.22	0.03	3.5	0.74	5.9
Health	0.19	0.66	4.3	0.16	4.2
Electricity	0.17	0.76	5.7	0.07	3.2
Price of food	0.12	0.05	2.3	0.83	5.1
Financial services	0.12	0.00		0.88	4.3
Jobs	0.12	0.67	5.6	0.21	4.0
Education	0.09	0.57	3.2	0.34	4.2
Recreational services	0.07	0.00		0.93	6.2

Notes: Data from the baseline woreda chief administrator (WCA) and citizen surveys. Panel A shows how well admins' perceptions of public priorities align with actual public priorities. Panel B reports how closely admins' own priorities align with public priorities. Admins' beliefs were measured by the rank they assigned top 5 in Panel A (top 3 in Panel B) sectors when asked to identify citizens' priorities. Citizen priorities were determined based on the rank individuals gave 5 sectors in terms priority. These rankings were converted into scores, averaged, and re-ranked from 1 (lowest priority) to 14 (highest priority) at the woreda level. All sectors that were unranked by admins were given a value of 6 in Panel A and 4 in Panel B. An admin's ranking was considered aligned with public priorities if it was within 1 rank of the true public ranking.

Table A3: No difference in baseline WCA knowledge and alignment by autonomy

	Gap in midline WCA knowledge of:		Baseline misalignment between:	
	Share of citizens prioritizing a sector	Share of citizens using public services	WCA perceptions and citizen priorities	WCAs' own and citizens' priorities
	(1)	(2)	(3)	(4)
High autonomy	-0.048	-0.047	-0.174	-0.395*
	[0.065]	[0.029]	[0.270]	[0.231]
Control mean	0.45	0.27	3.76	3.74
Sector FEs	No	No	No	No
Woreda FEs	No	No	No	No
N	112	170	580	580

Notes: Data from the woreda chief administrator (WCA) survey. The unit of observation is the woreda-sector. All outcomes are absolute differences. Midline outcomes are used for (1) and (2) given the lack of baseline data on those outcomes. Robust standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Baseline balance: local leaders

	WCAs			Bureau managers		
	Treatment mean	Control mean	p-value	Treatment mean	Control mean	p-value
Age	34.34	36.24	0.270	32.48	33.37	0.190
	[5.76]	[6.94]		[5.88]	[6.19]	
Male	1.00	0.93	0.600	0.86	0.88	0.510
	[0.00]	[0.26]		[0.35]	[0.33]	
Ethnicity: Oromo	0.76	0.55	0.140	0.73	0.67	0.280
	[0.44]	[0.51]		[0.45]	[0.47]	
Share with university education	1.00	0.97	1.000	0.92	0.92	0.810
	[0.00]	[0.19]		[0.28]	[0.27]	
Monthly salary in birr	7726.59	7475.38	0.590	7121.52	7296.13	0.270
	[1791.34]	[704.94]		[1565.75]	[1745.68]	
Years lived in woreda	10.94	14.95	0.210	10.46	11.50	0.290
	[11.95]	[11.27]		[10.87]	[11.69]	
Years in position	1.51	1.35	0.560	1.36	1.43	0.510
	[1.28]	[0.95]		[1.25]	[1.50]	
Years in public sector	11.16	14.02	0.080	10.19	10.85	0.260
	[5.39]	[6.59]		[6.72]	[6.47]	
Previously served in political office	0.43	0.52	0.610	0.29	0.26	0.440
	[0.50]	[0.51]		[0.46]	[0.44]	
Ideal tenure in position (years)	8.96	3.16	0.010	13.70	12.24	0.320
	[12.08]	[6.84]		[15.47]	[14.42]	
Satisfied with job	0.59	0.72	0.330	0.69	0.71	0.730
	[0.50]	[0.45]		[0.46]	[0.46]	
Work hours per week	75.31	72.48	0.510	65.60	62.26	0.020
	[16.53]	[11.27]		[15.55]	[15.12]	
Joint F-test			0.378			0.280
N	29	29		133	435	

Notes: Data from the baseline woreda chief administrator (WCA) and bureau manager surveys. The treatment group includes admins and bureau managers who had received a report card by the endline. Standard deviations shown in brackets. We report randomization-inference p-values and the result of the omnibus F-test at the bottom.

Table A5: Baseline balance: local autonomy

	WCAs			Managers		
	Treatment mean	Control mean	p-value	Treatment mean	Control mean	p-value
Share reporting having some or a lot of influence *deciding* policy in:						
Education	0.41	0.45	1.000	0.24	0.29	0.740
Health	0.45	0.45	1.000	0.17	0.25	0.550
SME	0.45	0.41	1.000	0.14	0.14	0.940
Jobs	0.45	0.41	1.000	0.19	0.30	0.320
Water	0.45	0.48	1.000	0.13	0.13	1.000
Electricity	0.07	0.10	1.000	0.13	0.13	1.000
Share reporting having some or a lot of influence *implementing* policy in:						
Education	0.86	0.86	1.000	0.76	0.86	0.430
Health	0.97	0.86	0.310	0.76	0.82	0.580
SME	0.93	0.83	0.500	0.76	0.89	0.280
Jobs	0.90	0.86	1.000	0.74	0.85	0.360
Water	0.83	0.76	0.720	0.93	0.93	1.000
Electricity	0.21	0.21	1.000	0.93	0.93	1.000
Reports having authority to reward high performers	0.69	0.72	1.000	0.87	0.81	0.170
Reports having authority to punish poor performers	0.97	0.86	0.360	0.97	0.95	0.400
Can influence budget allocation between sectors	0.69	0.79	0.560	0.97	0.95	0.400
Relative importance in deciding woreda policy:						
Higher-level authorities	0.81	0.71	0.230			
Woreda or kebele officials	0.16	0.22	0.470			
Public priorities	0.02	0.07	0.220			
N	29	29		133	435	

Notes: Data from the baseline woreda chief administrator (WCA) and bureau manager surveys. The treatment group includes admins and bureau managers who had received a report card by the endline. Standard deviations shown in brackets. We report randomization-inference p-values.

Table A6: Sampled households v. general population

	Study sample (2016)	UEUS sample (2014)
# HH members	4.28	8.25
	(1.95)	(4.58)
# under 18 HH members	1.71	2.71
	(1.42)	(2.37)
Female	0.54	0.53
	(0.50)	(0.50)
Married or cohabitating	0.63	0.50
	(0.48)	(0.50)
Age	33.32	34.72
	(11.90)	(14.93)
Moved within last 3 years	0.20	0.10
	(0.40)	(0.30)
Secondary schooling or above	0.35	0.22
	(0.48)	(0.42)
Monthly wage	3195.86	9415.88
	(6834.86)	(75495.99)
Self-employed	0.21	0.24
	(0.41)	(0.43)
Unemployed	0.31	0.36
	(0.46)	(0.48)
N	3540	17459

Notes: Data in column 1 uses baseline citizens survey (2016-17). Data in column 2 are from the Urban Employment Unemployment Surveys (2014). Standard deviations shown in parentheses.

Table A7: Null effect on admin knowledge of citizen priorities by sector

	Gap in WCA knowledge of:					
	4 months after baseline		24 months after baseline			
	Share of citizens prioritizing a sector	Share of citizens using public services	Share of citizens prioritizing a sector		Share of citizens using public services	
			as of Wave 1	as of Wave 2	as of Wave 1	as of Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Education						
RC-Bureau	0.02	0.00	0.02	0.08	-0.01	0.06
	[0.07]	[0.05]	[0.08]	[0.08]	[0.04]	[0.04]
Control mean	0.46	0.23	0.49	0.47	0.28	0.23
Midline controls	No	No	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	56	56	56	56	56	56
Panel B: Health						
RC-Bureau	-0.05	-0.04	-0.05	-0.02	0.05	-0.00
	[0.08]	[0.05]	[0.07]	[0.07]	[0.06]	[0.05]
Control mean	0.45	0.30	0.40	0.37	0.29	0.25
Midline controls	No	No	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No	No	No
N	56	56	56	56	56	56
Panel C: SME						
RC-Bureau		-0.05			0.01	0.00
		[0.06]			[0.04]	[0.03]
Control mean		0.25			0.14	0.15
Midline controls		No			Yes	Yes
Woreda FEs		No			No	No
N		58			58	58

Notes: The unit of observation is the woreda-sector. All outcomes are absolute differences between admin estimates and data from the citizens' survey. Columns 1 and 3-4 show treatment effects on the difference between admins' estimates of the share of citizens that prioritize a sector and the true share; this is available for only Education and Health. The rest of the columns show effects on the difference between admins' estimates of the share of citizens using public services and the true services coverage; this is available for Education, Health, and SME. All true values were shared in the report card, calculated using the baseline citizen survey. Columns 1–2 compare admins' midline estimates to baseline citizens' data. Columns 3 and 5 compare admins' endline responses to baseline citizens' data while columns 4 and 6 compare to endline citizens' data. All models control for region. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: No evidence of systematic updating

	Overestimated Wave 1 citizen priority at endline	Underestimated Wave 1 citizen priority at endline	Overestimated Wave 2 citizen priority at endline	Underestimated Wave 2 citizen priority at endline
	(1)	(2)	(3)	(4)
Panel A: Underestimated at baseline				
RC-Woreda	-0.00	-0.00	0.00	-0.04
	[0.06]	[0.06]	[0.08]	[0.07]
RC-Bureau	-0.06	0.07	-0.13	0.17
	[0.10]	[0.11]	[0.11]	[0.11]
RC-Woreda + RC-Bureau = 0	0.316	0.309	0.061	0.077
Control mean	0.19	0.76	0.43	0.48
Baseline controls	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No
N	355	355	355	355
Panel B: Overestimated at baseline				
RC-Woreda	-0.02	0.04	-0.05	0.04
	[0.03]	[0.03]	[0.04]	[0.05]
RC-Bureau	0.04	-0.08	0.02	-0.01
	[0.08]	[0.08]	[0.09]	[0.09]
RC-Woreda + RC-Bureau = 0	0.869	0.557	0.721	0.661
Control mean	0.80	0.18	0.55	0.37
Baseline controls	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No
N	379	379	379	379

Notes: The unit of observation is the woreda–sector. Columns 1 and 2 report treatment effects on whether the woreda chief administrator (WCA), at endline, over- or under-estimated a sector's priority relative to citizens' true rankings from Wave 1. Columns 3 and 4 repeat this using true citizen rankings from Wave 2. Results are shown separately for sectors the WCA initially under-estimated (Panel A) and over-estimated (Panel B) at baseline. All models control for region and the baseline absolute gap between WCA perceived rankings and true citizen rankings. Standard errors, clustered by woreda, are shown in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: No effect on woreda budget

	Appr. woreda budget (log)		Adj. woreda budget (log)	
	(1)	(2)	(3)	(4)
Treated × Post	0.033		0.031	
	[0.049]		[0.044]	
Treated × 2018		0.029		0.030
		[0.041]		[0.042]
Treated × 2019		0.038		0.032
		[0.062]		[0.053]
Control mean	17.27	17.27	17.39	17.39
Year FEs	Yes	Yes	Yes	Yes
N	284	284	284	284

Notes: The unit of observation is the woreda-year. Treated woredas received report cards starting in 2018. All models control for region. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Citizen engagement and reporting practices across bureaus

	Index (ICW)	Conducts at least 1 site visit bi-weekly	Tracks user satisfaction	# sources of community feedback	Receives community feedback bi-weekly	Reports to WCA bi-weekly
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: RC sample						
RC-Woreda	0.10	0.02	0.07	-0.42	0.02	0.09**
	[0.17]	[0.07]	[0.06]	[0.40]	[0.10]	[0.04]
RC-Bureau	-0.02			0.08	-0.02	0.00
	[0.09]			[0.21]	[0.07]	[0.02]
RC-Woreda + RC-Bureau = 0	0.509	0.823	0.310	0.331	0.933	0.124
Control mean	-0.11	0.39	0.86	4.05	0.73	0.02
Number of sectors	15	3	2	15	15	15
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	523	167	97	523	521	523
Panel B: Monitoring sample, Addis only						
RC-Woreda	0.02	0.03	-0.04	-0.08	0.04	0.00
	[0.05]	[0.09]	[0.09]	[0.22]	[0.06]	[0.02]
RC-Bureau	-0.01	0.06	0.01	0.05	0.07	-0.02
	[0.06]	[0.08]	[0.09]	[0.20]	[0.06]	[0.03]
RC-Woreda + RC-Bureau = 0	0.920	0.266	0.652	0.928	0.152	0.625
Control mean	-0.26	0.52	0.88	3.17	0.71	0.05
# sectors	15	3	2	15	15	15
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	701	260	136	701	691	701

Notes: The unit of observation is the woreda-sector. In column 2, the outcome variable equals 1 if the bureau head reported visiting at least one of the facilities or firms under its oversight bi-weekly (e.g., public schools for Education, health clinics for Health, and firms for SME). In column 3, the outcome variable flags whether a bureau reports collecting service user satisfaction as part of its evaluation tool. In column 4, the outcome is the number of distinct sources of community feedback used by a bureau. These include suggestion boxes, community meetings, hotlines, and six additional options—for a total of nine possible sources. In column 5, the outcome is an index ranging from 0 to 1, measuring the share of selected feedback sources reviewed by the bureau on a biweekly basis. Lastly, column 6 shows effects on the probability of bureau heads reporting to the woreda chief administrator (WCA) bi-weekly or more frequently. The number of sectors for which outcome data is available is indicated at the bottom of each panel. All models control for region and baseline values of the outcome. Robust standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Lobbying and project development across bureaus

	Index (ICW)	Lobbying attempts during budgeting	Requested more funds from woreda	Requested more funds from higher-ups	Proposed project using joint funds
	(1)	(2)	(3)	(4)	(5)
Panel A: RC sample					
RC-Woreda	0.00	0.30	-0.11	0.03	0.00
	[0.04]	[0.35]	[0.09]	[0.04]	[0.01]
RC-Bureau	-0.03	-0.41	0.03	-0.07*	0.02
	[0.07]	[0.32]	[0.06]	[0.03]	[0.02]
RC-Woreda + RC-Bureau = 0	0.698	0.722	0.156	0.152	0.284
Control mean	-0.11	1.34	0.25	0.04	0.02
Number of sectors	30	11	14	14	30
Baseline controls	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes
N	1,543	545	478	478	1,410
Panel B: Monitoring sample, Addis only					
RC-Woreda	0.01	0.03	-0.01	0.08**	0.00
	[0.03]	[0.14]	[0.05]	[0.04]	[0.00]
RC-Bureau	0.02	0.03	0.01	-0.00	-0.00
	[0.04]	[0.07]	[0.06]	[0.04]	[0.00]
RC-Woreda + RC-Bureau = 0	0.535	0.738	0.971	0.120	0.565
Control mean	-0.03	0.96	0.22	0.08	0.03
# sectors	30	11	14	14	30
Baseline controls	Yes	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes	Yes
N	2,654	857	614	614	2,640

Notes: The unit of observation is the woreda-sector. Column 1 reports the number of times officials from a bureau lobbied on behalf of their sector during the budgeting process over the past two years. In column 4, the outcome indicates whether the woreda proposed a project—under the joint citizen/government funding system—for the given sector in the past two years. The number of sectors for which outcome data is available is indicated at the bottom of each panel. All models control for region and the baseline value of the outcome. Robust standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Lobbying and project development across woredas

	Proposals relying on municipal or construction budget	Submitted supplementary budget requests	Informally negotiated budget with higher authorities
	(1)	(2)	(3)
Treated	-2.523	0.150	0.142
	[3.476]	[0.109]	[0.112]
Control mean	8.05	0.14	0.14
Baseline controls	Yes	Yes	Yes
N	42	56	56

Notes: Data from the municipality and finance modules of the bureau manager surveys. The unit of observation is the woreda. In column 1, the outcome is the number of projects proposed by the woreda in the past two years that relied on municipal or construction budgets. In columns 2 and 3, the outcomes are binary indicators equal to 1 if the woreda (i) submitted additional budget requests or (ii) informally negotiated with higher-level officials about the budget, respectively. All models control for region and baseline values of the outcome. Robust standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

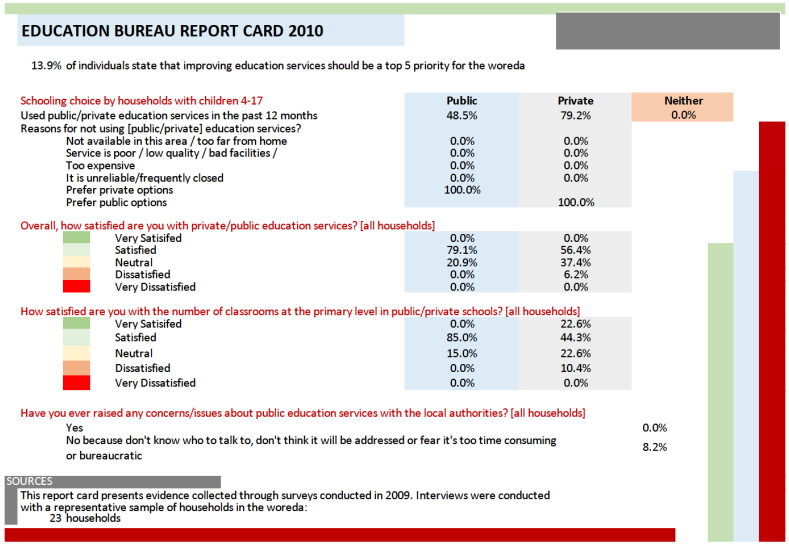
B Appendix B: Report Cards

B.1 Bureau Report Cards

These are examples of report cards delivered as part of Experiment 1 to bureau chiefs in the Health, Education, and Small and Micro Enterprise bureaus.

Figure B.1: Bureau report cards by sector

(a) Education



EDUCATION BUREAU REPORT CARD 2010

13.9% of individuals state that improving education services should be a top 5 priority for the woreda

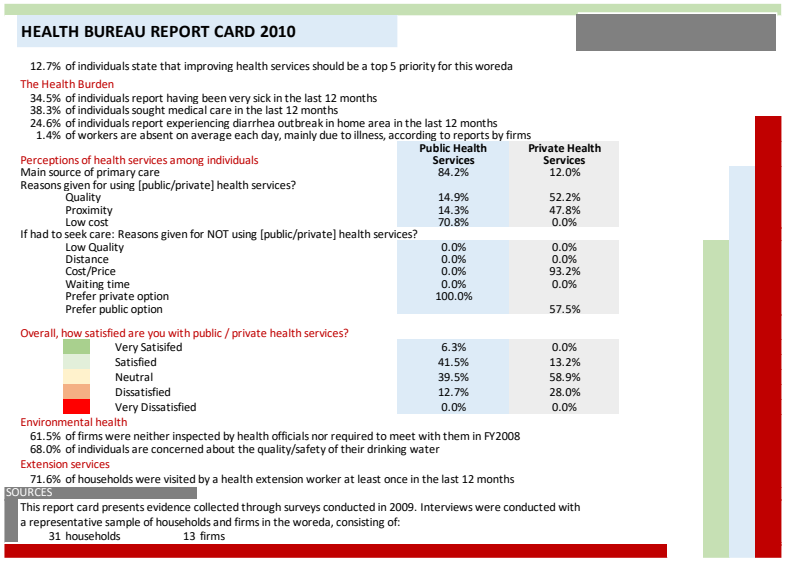
Schooling choice by households with children 4-17	Public	Private	Neither
Used public/private education services in the past 12 months	48.5%	79.2%	0.0%
Reasons for not using [public/private] education services?			
Not available in this area / too far from home	0.0%	0.0%	
Service is poor / low quality / bad facilities /	0.0%	0.0%	
Too expensive	0.0%	0.0%	
It is unreliable/frequently closed	0.0%	0.0%	
Prefer private options	100.0%		
Prefer public options		100.0%	
Overall, how satisfied are you with private/public education services? [all households]			
Very Satisifed	0.0%	0.0%	
Satisfied	79.1%	56.4%	
Neutral	20.9%	37.4%	
Dissatisfied	0.0%	6.2%	
Very Dissatisfied	0.0%	0.0%	
How satisfied are you with the number of classrooms at the primary level in public/private schools? [all households]			
Very Satisifed	0.0%	22.6%	
Satisfied	85.0%	44.3%	
Neutral	15.0%	22.6%	
Dissatisfied	0.0%	10.4%	
Very Dissatisfied	0.0%	0.0%	
Have you ever raised any concerns/issues about public education services with the local authorities? [all households]			
Yes			0.0%
No because don't know who to talk to, don't think it will be addressed or fear it's too time consuming or bureaucratic			8.2%

SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a representative sample of households in the woreda:
23 households

Figure B.1: Bureau report cards by sector

(b) Health



HEALTH BUREAU REPORT CARD 2010

12.7% of individuals state that improving health services should be a top 5 priority for this woreda

The Health Burden

34.5% of individuals report having been very sick in the last 12 months
38.3% of individuals sought medical care in the last 12 months
24.6% of individuals report experiencing diarrhea outbreak in home area in the last 12 months
1.4% of workers are absent on average each day, mainly due to illness, according to reports by firms

	Public Health Services	Private Health Services
Perceptions of health services among individuals		
Main source of primary care	84.2%	12.0%
Reasons given for using [public/private] health services?		
Quality	14.9%	52.2%
Proximity	14.3%	47.8%
Low cost	70.8%	0.0%
If had to seek care: Reasons given for NOT using [public/private] health services?		
Low Quality	0.0%	0.0%
Distance	0.0%	0.0%
Cost/Price	0.0%	93.2%
Waiting time	0.0%	0.0%
Prefer private option	100.0%	
Prefer public option		57.5%
Overall, how satisfied are you with public / private health services?		
Very Satisifed	6.3%	0.0%
Satisfied	41.5%	13.2%
Neutral	39.5%	58.9%
Dissatisfied	12.7%	28.0%
Very Dissatisfied	0.0%	0.0%

Environmental health

61.5% of firms were neither inspected by health officials nor required to meet with them in FY2008
68.0% of individuals are concerned about the quality/safety of their drinking water

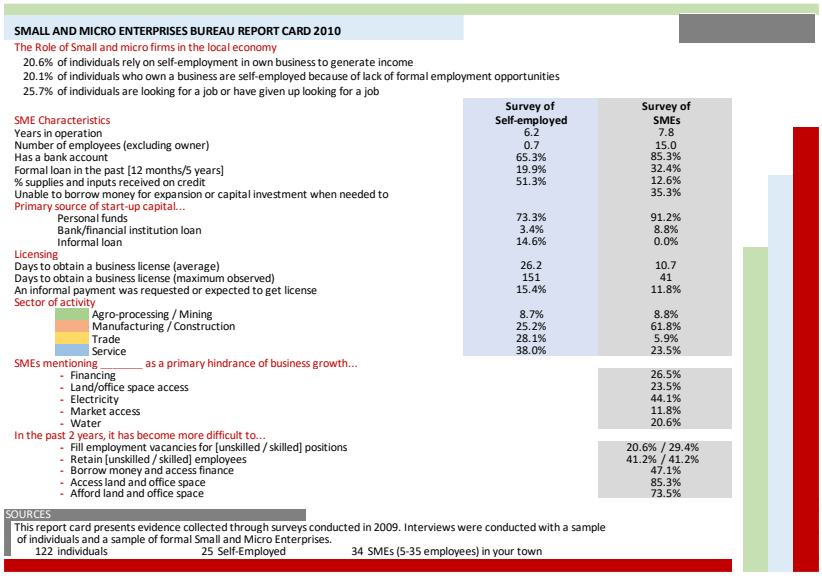
Extension services

71.6% of households were visited by a health extension worker at least once in the last 12 months

SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a representative sample of households and firms in the woreda, consisting of:
31 households 13 firms

(c) SME



SMALL AND MICRO ENTERPRISES BUREAU REPORT CARD 2010

The Role of Small and micro firms in the local economy

20.6% of individuals rely on self-employment in own business to generate income
20.1% of individuals who own a business are self-employed because of lack of formal employment opportunities
25.7% of individuals are looking for a job or have given up looking for a job

	Survey of Self-employed	Survey of SMEs
SME Characteristics		
Years in operation	6.2	7.8
Number of employees (excluding owner)	0.7	15.0
Has a bank account	65.3%	85.3%
Formal loan in the past [12 months/5 years]	19.9%	32.4%
% supplies and inputs received on credit	51.3%	12.6%
Unable to borrow money for expansion or capital investment when needed to		35.3%
Primary source of start-up capital…		
Personal funds	73.3%	91.2%
Bank/financial institution loan	3.4%	8.8%
Informal loan	14.6%	0.0%
Licensing		
Days to obtain a business license (average)	26.2	10.7
Days to obtain a business license (maximum observed)	151	41
An informal payment was requested or expected to get license	15.4%	11.8%
Sector of activity		
Agro-processing / Mining	8.7%	8.8%
Manufacturing / Construction	25.2%	61.8%
Trade	28.1%	5.9%
Service	38.0%	23.5%
SMEs mentioning _____ as a primary hindrance of business growth…		
- Financing		26.5%
- Land/office space access		23.5%
- Electricity		44.1%
- Market access		11.8%
- Water		20.6%
In the past 2 years, it has become more difficult to…		
- Fill employment vacancies for [unskilled / skilled] positions		20.6% / 29.4%
- Retain [unskilled / skilled] employees		41.2% / 41.2%
- Borrow money and access finance		47.1%
- Access land and office space		85.3%
- Afford land and office space		73.5%

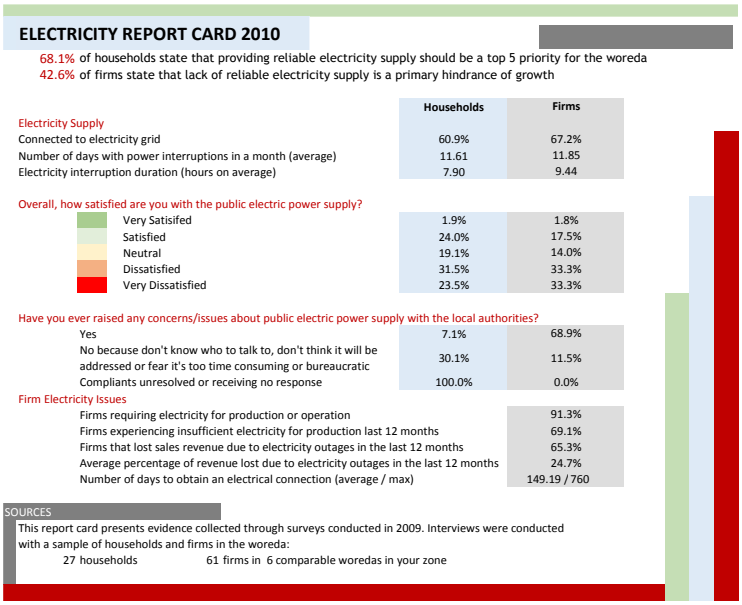
SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a sample of individuals and a sample of formal Small and Micro Enterprises.
122 individuals 25 Self-Employed 34 SMEs (5-35 employees) in your town

B.2 Admin Report Cards

Figure B.2: Admin report cards by sector

(a) Electricity



ELECTRICITY REPORT CARD 2010

68.1% of households state that providing reliable electricity supply should be a top 5 priority for the woreda
42.6% of firms state that lack of reliable electricity supply is a primary hindrance of growth

	Households	Firms
Electricity Supply		
Connected to electricity grid	60.9%	67.2%
Number of days with power interruptions in a month (average)	11.61	11.85
Electricity interruption duration (hours on average)	7.90	9.44
Overall, how satisfied are you with the public electric power supply?		
Very Satisifed	1.9%	1.8%
Satisfied	24.0%	17.5%
Neutral	19.1%	14.0%
Dissatisfied	31.5%	33.3%
Very Dissatisfied	23.5%	33.3%
Have you ever raised any concerns/issues about public electric power supply with the local authorities?		
Yes	7.1%	68.9%
No because don't know who to talk to, don't think it will be addressed or fear it's too time consuming or bureaucratic	30.1%	11.5%
Compliants unresolved or receiving no response	100.0%	0.0%
Firm Electricity Issues		
Firms requiring electricity for production or operation		91.3%
Firms experiencing insufficient electricity for production last 12 months		69.1%
Firms that lost sales revenue due to electricity outages in the last 12 months		65.3%
Average percentage of revenue lost due to electricity outages in the last 12 months		24.7%
Number of days to obtain an electrical connection (average / max)		149.19 / 760

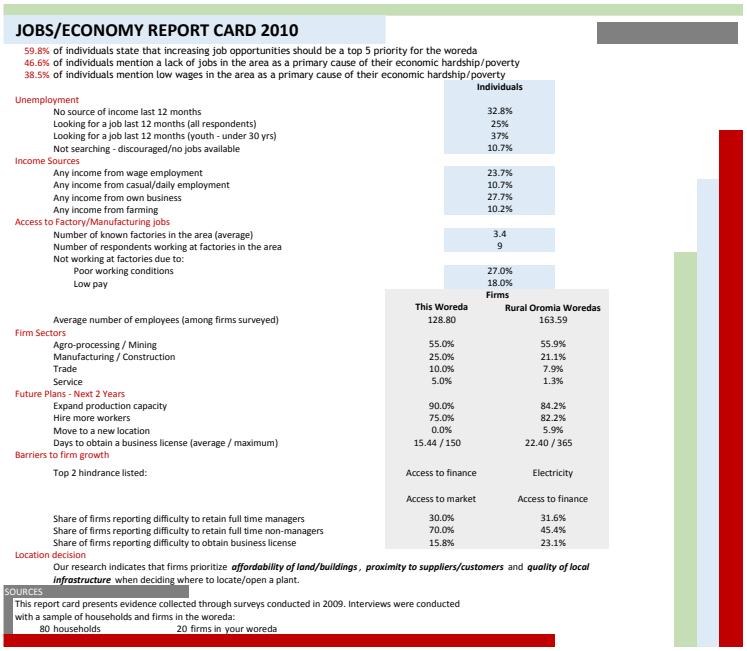
SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a sample of households and firms in the woreda:

27 households 61 firms in 6 comparable woredas in your zone

Figure B.2: Admin report cards by sector

(b) Jobs



JOBS/ECONOMY REPORT CARD 2010

59.8% of individuals state that increasing job opportunities should be a top 5 priority for the woreda
46.6% of individuals mention a lack of jobs in the area as a primary cause of their economic hardship/poverty
38.5% of individuals mention low wages in the area as a primary cause of their economic hardship/poverty

	Individuals
Unemployment	
No source of income last 12 months	32.8%
Looking for a job last 12 months (all respondents)	25%
Looking for a job last 12 months (youth - under 30 yrs)	37%
Not searching - discouraged/no jobs available	10.7%
Income Sources	
Any income from wage employment	23.7%
Any income from casual/daily employment	10.7%
Any income from own business	27.7%
Any income from farming	10.2%
Access to Factory/Manufacturing jobs	
Number of known factories in the area (average)	3.4
Number of respondents working at factories in the area	9
Not working at factories due to:	
Poor working conditions	27.0%
Low pay	18.0%

	Firms	
	This Woreda	Rural Oromia Woredas
Average number of employees (among firms surveyed)	128.80	163.59
Firm Sectors		
Agro-processing / Mining	55.0%	55.9%
Manufacturing / Construction	25.0%	21.1%
Trade	10.0%	7.9%
Service	5.0%	1.3%
Future Plans - Next 2 Years		
Expand production capacity	90.0%	84.2%
Hire more workers	75.0%	82.2%
Move to a new location	0.0%	5.9%
Days to obtain a business license (average / maximum)	15.44 / 150	22.40 / 365
Barriers to firm growth		
Top 2 hindrance listed:	Access to finance Access to market	Electricity Access to finance
Share of firms reporting difficulty to retain full time managers	30.0%	31.6%
Share of firms reporting difficulty to retain full time non-managers	70.0%	45.4%
Share of firms reporting difficulty to obtain business license	15.8%	23.1%

Location decision

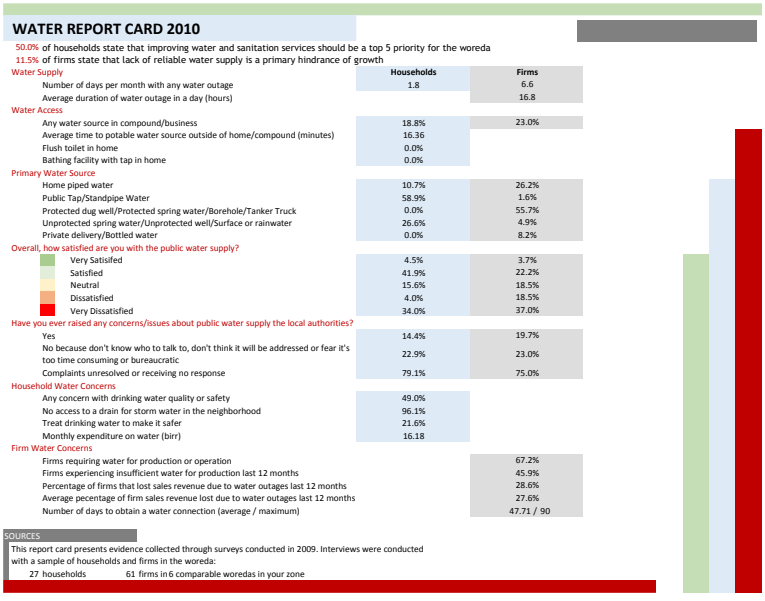
Our research indicates that firms prioritize ***affordability of land/buildings***, ***proximity to suppliers/customers*** and ***quality of local infrastructure*** when deciding where to locate/open a plant.

SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a sample of households and firms in the woreda:

80 households 20 firms in your woreda

(c) Water



WATER REPORT CARD 2010

50.0% of households state that improving water and sanitation services should be a top 5 priority for the woreda
11.5% of firms state that lack of reliable water supply is a primary hindrance of growth

	Households	Firms
Water Supply		
Number of days per month with any water outage	1.8	6.6
Average duration of water outage in a day (hours)		16.8
Water Access		
Any water source in compound/business	18.8%	23.0%
Average time to potable water source outside of home/compound (minutes)	16.36	
Flush toilet in home	0.0%	
Bathing facility with tap in home	0.0%	
Primary Water Source		
Home piped water	10.7%	26.2%
Public Tap/Standpipe Water	58.9%	1.6%
Protected dug well/Protected spring water/Borehole/Tanker Truck	0.0%	55.7%
Unprotected spring water/Unprotected well/Surface or rainwater	26.6%	4.9%
Private delivery/Bottled water	0.0%	8.2%
Overall, how satisfied are you with the public water supply?		
Very Satisifed	4.5%	3.7%
Satisfied	41.9%	22.2%
Neutral	15.6%	18.5%
Dissatisfied	4.0%	18.5%
Very Dissatisfied	34.0%	37.0%
Have you ever raised any concerns/issues about public water supply the local authorities?		
Yes	14.4%	19.7%
No because don't know who to talk to, don't think it will be addressed or fear it's too time consuming or bureaucratic	22.9%	23.0%
Complaints unresolved or receiving no response	79.1%	75.0%
Household Water Concerns		
Any concern with drinking water quality or safety	49.0%	
No access to a drain for storm water in the neighborhood	96.1%	
Treat drinking water to make it safer	21.6%	
Monthly expenditure on water (birr)	16.18	
Firm Water Concerns		
Firms requiring water for production or operation		67.2%
Firms experiencing insufficient water for production last 12 months		45.9%
Percentage of firms that lost sales revenue due to water outages last 12 months		28.6%
Average pecentage of firm sales revenue lost due to water outages last 12 months		27.6%
Number of days to obtain a water connection (average / maximum)		47.71 / 90

SOURCES

This report card presents evidence collected through surveys conducted in 2009. Interviews were conducted with a sample of households and firms in the woreda:

27 households 61 firms in 6 comparable woredas in your zone

B.3 Report Card Scripts

Before I leave, I would like to share these handouts with you. It explains what the study is doing. As I said when we started the survey, the Ethiopian Development Research Institute (EDRI) is collaborating with Stanford University in the United States on a research project. The project attempts to understand the opportunities and challenges that economic growth and urbanization bring about, in terms of local governance and public service provision. Theise handouts explains this and provides some information. It says that it is a multi-year study that intends to track the changes in the interactions of citizens and businesses with services provided by the federal, city, and sub-city governments in Addis Ababa and the Oromia region. We wanted you to have this as a record and also so that you have our contact information if you have any questions.

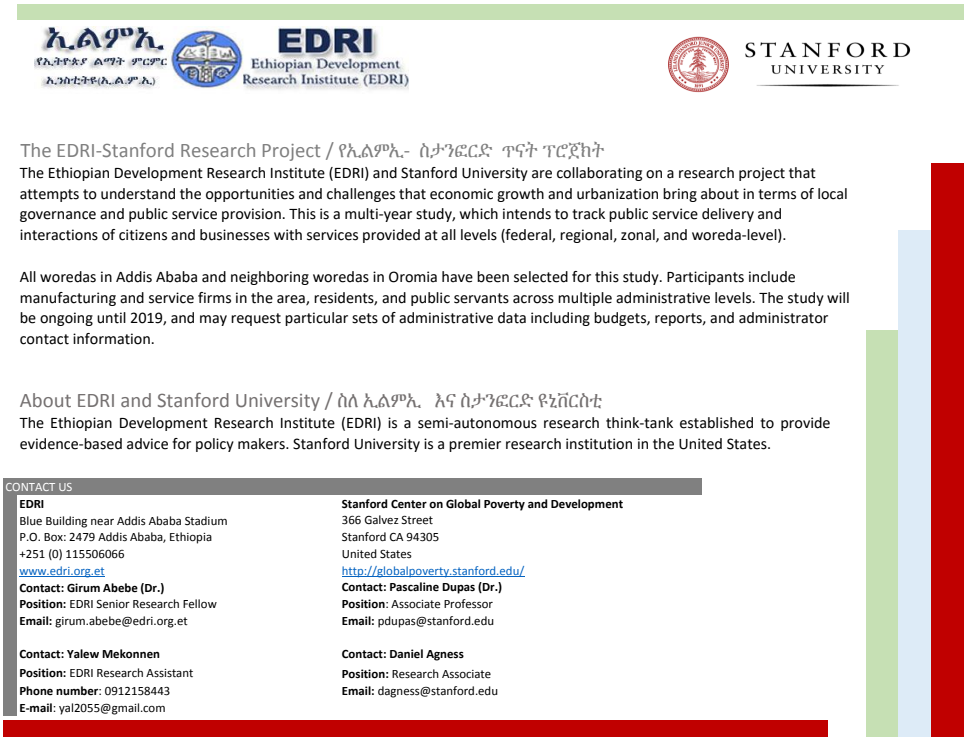
All woredas in Addis Ababa and around 50 neighboring woredas in the Oromia region have been selected for this study. The study includes surveys to administrators as well as citizens and firms.

FOR WOREDAS SAMPLED FOR RC EXPERIMENT: The back page of thisThese handouts includes statistics that we put together for you. These statistics are based on the surveys we conducted in this area in the recent past. Specifically, we randomly selected and surveyed some individuals in your woreda from June to December 2008. We did the same with some firms from January to May 2009. We followed gold standard research methods to obtain a representative sample. The number of individuals and firms surveyed is different in each woreda and depends on the woreda's population size and the number of ketenas/kebeles. Only adults and firms with more than 5 employees were surveyed.

As a local leader, we are aware that you already know a lot about your woreda. But we also know that the local officials have sometimes too little resources on hand to conduct surveys like those we did, so we thought that we should share the information we gathered with you.

FOR WOREDAS SAMPLED FOR MONITORING EXPERIMENT: If you have any questions, feel free to ask me now, I will do my best to answer. You would also like to contact the EDRI and Stanford researchers using the phone number and/or the email specified in the "Contact Us" section if you need any further information. Thanks a lot again for your time today and for valuing this research project.

B.4 Monitoring Announcement



C Appendix C: Demand Experiment

C.1 Demand Experiment Script

C.1.1 Opt-In

PSI has partnered with Stanford University to conduct a longitudinal study of the opportunities and challenges facing dwellers of Addis Ababa and the Oromia region. We randomly selected and surveyed some individuals in your woreda from June to December 2008. We did the same with some firms from January to May 2009. The number of individuals and firms surveyed is different in each woreda and depends on the woreda's population size and the number of ketenas/kebeles. Only adults and firms with more than 5 employees were surveyed. We conducted a new round of surveys with the same individuals and firms, exactly two years later from September 2010 to February 2011 for the households survey and from March to July 2011 for the firm survey. We also added new individuals and firms to the sample to reflect inflows into your woreda.

FOR 10 & 25 MIN VERSION, ADD: We have sampled 54 in a doughnut shape around the Addis Ababa City Center. The study includes woredas adjacent to the city limits from the inside, as well as woredas just outside the administrative boundary of the city. We also sampled areas just outside the city that are in the process of urbanizing; the sample covers all woredas sharing a border with Addis Ababa and represents a full census of medium-to-large towns within 50km of the capital boundaries.

FOR 25 MIN VERSION, ADD: We conducted (1) a listing exercise that collected information on basic dwelling and household characteristics for a representative sample of 5,200 households within the 54 selected woredas; (2) a detailed survey with a representative sample of around 3,618 individuals drawn from these households; (3) a detailed survey with 1,243 firms operating in the 54 selected woredas or in Addis city center; (4) interviews with 1,570 local government officials (woreda mayors and their bureau chiefs) from 147 woredas (in addition to the 54 sampled, we added all woredas within Addis city center); (5) revenue sheets and budgets for the 147 woredas.

We will be happy to share the findings of our research, in the form of these report cards, with you, in order to support your mission of quality public service delivery in this community. We can produce report cards on health, education, small and medium entreprise development, water, electricity, jobs, and possibly other topics depending on requests.

As a local leader, we are aware that you already know a lot about your woreda. But we also know that the local officials have sometimes too little resources on hand to conduct surveys like those we did, so we thought that we could share the information we gathered with you.

In order to get the report card, we humbly ask that you fill in the request form below. You must fully fill out the following form for us to proceed with the production of the report cards. Any incomplete request form will be discarded.

C.1.2 Opt-Out

PSI has partnered with Stanford University to conduct a longitudinal study of the opportunities and challenges facing dwellers of Addis Ababa and the Oromia region. We randomly selected and surveyed some individuals in your woreda from June to December 2008. We did the same with some firms from January to May 2009. The number of individuals and firms surveyed is different in each woreda and depends on the woreda's population size and the number of ketenas/kebeles. Only adults and firms with more than 5 employees were surveyed. We conducted a new round of surveys with the same individuals and firms, exactly two years later, from September 2010 to February 2011 for the households survey and from March to July 2011 for the firm survey. We also added new individuals and firms to the sample to reflect inflows into your woreda.

FOR 10 MIN VERSION, ADD: We have sampled 54 woredas (the administrative unit below the region and zone) in a doughnut shape around the Addis Ababa City Center. The study includes woredas adjacent to the city limits from the inside, as well as woredas just outside the administrative boundary of the city. We also sampled areas just outside the city that are in the process of urbanizing; the sample covers all woredas sharing a border with Addis Ababa and represents a full census of medium-to-large towns within 50km of the capital boundaries. We conducted (1) a listing exercise that collected information on basic dwelling and household characteristics for a representative sample of 5,200 households within the 54 selected woredas; (2) a detailed survey with a representative sample of around 3,618 individuals drawn from these households; (3) a detailed survey with 1,243 firms operating in the 54 selected woredas or in Addis city center; (4) interviews with 1,570 local government officials (woreda mayors and their bureau chiefs) from 147 woredas (in addition to the 54 sampled, we added all woredas within Addis city center; (5) revenue sheets and budgets for the 147 woredas.

We plan to share the findings of our research, in the form of these report cards, with you, in order to support your mission of quality public service delivery in this community. We will produce report cards on health, education, small and medium entreprise development, water, electricity, jobs, and possibly other topics depending on requests.

As a local leader, we are aware that you already know a lot about your woreda. But we also know that the local officials have sometimes too little resources on hand to conduct surveys like those we did, so we thought that we should share the information we gathered with you.

If you DO NOT WANT the report cards, we humbly ask that you fill in the cancellation form below. You must fully fill out the following form for us to cancel the production of the report cards. Any incomplete request form will be discarded.

C.2 Opt-In/Opt-Out Forms

Figure C.1: Opt-in forms

(a) Part 1

1.	WOREDA NAME:
2.	YOUR NAME:
3.	YOUR TITLE:
4.	TODAY'S DATE:
5.	WHICH REPORT CARDS DO YOU WANT TO RECEIVE? LIST THE SECTORS OF CONCERN BELOW 1. ---------- 2. ---------- 3. ---------- 4. ---------- 5. ---------- 6. ---------- 7. ----------
6.	Would like these report cards to be share with your superior the Administrator of the woreda or the Sub-city? ☐ Yes, I would like the reports cards to be shared with my superiors. ☐ No, I would like the information to be only disclosed to myself.

(b) Part 2

FOR 10 & 25 MIN VERSION, ADD:

7.	List the 8 [16 IF 25 MIN VERSION] statistics/pieces of information that would be most helpful for you to get data on, by order of importance. Be as clear as possible. Make sure to list no less than 8 [16]. 1. _____ 9. _____ 2. _____ 10. _____ 3. _____ 11. _____ 4. _____ 12. _____ 5. _____ 13. _____ 6. _____ 14. _____ 7. _____ 15. _____ 8. _____ 16. _____
8.	Because we surveyed people and firms twice (2009 and 2011), we can look at changes over time. We can compare trends in your woreda to those of nearby woredas. Which woredas would like to see comparisons to? List the top 3. *If you do not want comparisons with other woredas, please explain why here: (4 lines minimum)* ---------- ---------- ---------- ----------

Figure C.2: Opt-out form

1.	WOREDA NAME:
2.	YOUR NAME:
3.	YOUR TITLE:
4.	TODAY'S DATE:
5.	WHICH REPORT CARDS DO YOU **NOT** WANT TO RECEIVE? LIST THE SECTORS BELOW 1. ---------- 2. ---------- 3. ---------- 4. ---------- 5. ---------- 6. ---------- 7. ----------

FOR 10 MIN VERSION, ADD:

6.	*Please explain your choice to cancel the Report Cards production: (4 lines minimum)* ---------- ---------- ---------- ----------

D Appendix D: Model Derivations and Extensions

D.1 Proofs

Proof of Lemma 1 Substitute (1) into (4):

$$[L \mid \mathcal{I}] = \beta(p - \theta_c)^2 + \alpha((p - m)^2 + V) + ce.$$

Terms V and ce are constant in p. Differentiating the remaining terms and setting the derivative to zero yields

$$2\beta(p - \theta_c) + 2\alpha(p - m) = 0 \quad \Longleftrightarrow \quad (\beta + \alpha)p = \beta\theta_c + \alpha m,$$

which implies (5). The objective is strictly convex in p when $\beta + \alpha > 0$, so the optimizer is unique. □

Proof of Proposition 1 Under $e = 0$, $(m, V) = (0, \sigma^2)$ and the optimal policy is (6). The resulting expected loss is

$$\begin{aligned} L_0 &= \beta(p_0^* - \theta_c)^2 + \alpha[(p_0^* - \theta_p)^2] \\ &= \beta\left(\frac{-\alpha}{\beta + \alpha}\theta_c\right)^2 + \alpha\left(\left(\frac{\beta}{\beta + \alpha}\theta_c\right)^2 + \sigma^2\right) = \frac{\alpha\beta}{\beta + \alpha}\theta_c^2 + \alpha\sigma^2. \end{aligned}$$

Under $e = 1$, the optimal policy is (7). Conditional on θ_p,

$$p_1^*(\theta_p) - \theta_c = \frac{\alpha}{\beta + \alpha}(\theta_p - \theta_c), \qquad p_1^*(\theta_p) - \theta_p = \frac{\beta}{\beta + \alpha}(\theta_c - \theta_p),$$

so the realized loss (excluding c) is

$$\beta(p_1^* - \theta_c)^2 + \alpha(p_1^* - \theta_p)^2 = \frac{\alpha\beta}{\beta + \alpha}(\theta_p - \theta_c)^2.$$

Taking expectations, $[(\theta_p - \theta_c)^2] = \sigma^2 + \theta_c^2$, hence

$$L_1 = \frac{\alpha\beta}{\beta + \alpha}(\sigma^2 + \theta_c^2) + c.$$

Therefore, $e^* = 1$ iff $L_1 \le L_0$, which is equivalent to

$$c \le \alpha\sigma^2 - \frac{\alpha\beta}{\beta + \alpha}\sigma^2 = \frac{\alpha^2}{\beta + \alpha}\sigma^2,$$

giving (8). □

D.2 Extensions and robustness

Finite policy bounds

Suppose the proposed policy must lie in an interval $p \in [-\bar{p}, \bar{p}]$ for some $\bar{p} \in (0, \infty)$. Then the optimal policy is the projection of (5) onto the feasible set:

$$p^*_{\text{bd}}(m) = \Pi_{[-\bar{p},\bar{p}]}\left(\frac{\beta\theta_c + \alpha m}{\beta + \alpha}\right).$$

All main comparative statics remain: increasing α increases responsiveness on the interior and weakly increases the value of information by making public responsiveness more important when feasible.

Other distance functions

If $D(p, \theta)$ is any convex distance function (e.g., absolute loss), the policy problem remains a convex minimization and the solution continues to interpolate between θ_c and the belief about θ_p, with the weights governed by β and α. The quadratic case yields closed forms and isolates the variance-reduction value of information in Proposition 1.

Information aversion and willingness to pay to avoid information

In the baseline model, observing θ_p cannot raise the local's loss because the local can always ignore the signal and choose the same policy rule it would have chosen when uninformed (a standard "value of information" argument). However, in our second experiment we introduce the opportunity for costly information avoidance. Below, we present two simple variants that microfound such avoidant behavior.

As before, let $\theta_p \sim \mathcal{N}(0, \sigma^2)$ and $\alpha \equiv \alpha_\ell + \alpha_c$. Recall that in the baseline quadratic-normal environment, the gross value of observing θ_p (before any observation-specific cost) is

$$\Delta_0 \;=\; \frac{\alpha^2}{\beta + \alpha}\,\sigma^2. \tag{13}$$

This is the reduction in the local's expected loss from conditioning policy on θ_p rather than on the prior mean.

Suppose observing any signal (regardless of its content) triggers a fixed cost $\psi > 0$:

$$L \;=\; \beta(p - \theta_c)^2 + \alpha(p - \theta_p)^2 + ce \;+\; \psi \cdot \mathbf{1}\{\text{observe}\}, \qquad \psi \geq 0. \tag{14}$$

This captures settings in which observing creates an obligation to respond, document, or report, or where exposure to citizen feedback is intrinsically unpleasant.

In this case the expected observation-specific cost is ψ, so the net value of observing is

$$\Delta_{fixed} \;=\; \Delta_0 - \psi. \tag{15}$$

Binary effort / observation rule: If observing requires effort at cost c, the local

observes if and only if

$$e^* = 1 \quad \Longleftrightarrow \quad c \leq \Delta_{fixed} \;=\; \frac{\alpha^2}{\beta + \alpha}\sigma^2 - \psi. \tag{16}$$

Willingness to pay to avoid information: If the signal is provided by default and the local can pay d to avoid, then the willingness to pay is

$$\text{WTP}_B \;=\; \max\{0,\; \psi - \Delta_0\}. \tag{17}$$

E Appendix E: Additional Figures and Tables

Table E1: Expenditure data on Ethiopia

	Regional, woreda, and national expenditures (in millions)				
	Region	Woreda	Source	Latest year available	Notes
Harai	1,000.67	170.11	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
Afar	3,976.36	1,471.25	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
Gambella	1,751.18	700.47	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
BG	2,409.52	1,011.99	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
Somali	14,775	12,780.38	World Bank (2020). *Performance assessment report: Somali Regional State Government*	2017/18	–
Oromia	57,355	33,839.45	UNICEF (2021). *Expenditure analysis for Oromia Regional State (2012/13–2020/21)*	2018/19	–
SNNPR	35,593	22,779.52	UNICEF (2021). *Expenditure analysis for Southern Nations, Nationalities, and Peoples' Region (2012/13–2019/20)*	2018/19	–
Amhara	41,147	27,157.02	UNICEF (2021). *Expenditure analysis for Amhara Regional State (2012/13–2020/21)*	2018/19	–
Tigray	7,939.56	4,446.15	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
Dire Dawa	1,158.67	1,158.67	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	Subsidy amounts; expenditure data unavailable
Addis Ababa	30,013	20,678.96	World Bank (2019). *Performance assessment report: The City of Addis Ababa*	2017/18	–
Ethiopia	411,809	135,604.73	World Bank (2019). *Performance assessment report: Federal Democratic Republic of Ethiopia*	2018/19	–

Notes: The share of expenditures at the woreda level is then calculated as the sum of woreda expenditures across regions, divided by the national budget.

Table E2: Leaders' characteristics v. broader labor force

	Surveyed local leaders (1)	2014 UEUS full sample (2)	2014 UEUS private sector only (3)
Male	0.85	0.47	0.63
	(0.36)	(0.50)	(0.48)
Age	32.80	34.72	31.29
	(5.83)	(14.93)	(10.81)
University-educated	0.96	0.06	0.07
	(0.19)	(0.25)	(0.26)
Monthly salary in birr	7314.12	9415.88	8832.93
	(1726.79)	(75495.99)	(49943.23)
Work hours per week	66.05	39.89	42.55
	(17.10)	(22.71)	(22.21)
Satisfied with job	0.68	0.21	0.57
	(0.47)	(0.41)	(0.49)
Observations	1520	17459	2924

Notes: Data in column 1 from the baseline admin and bureau manager surveys (2016-17). In columns 2-3, the data are from the 2014 Urban Employment and Unemployment Survey (UEUS), conducted by Ethiopia's Central Statistical Agency. We report results for the UEUS sub-sample, aged 18 and above and living in Oromia or Addis Ababa.

Table E3: Budget descriptive stats FY 2017 (Addis Ababa)

	Mean	SD	Min	Max
Panel A: Woreda-level outcomes				
Approved budget (in 100k, 2025)	9.52	1.73	6.46	14.96
Net adjustment (2025)	545.37	383.40	-850.84	1712.54
Share of expenditures by function:				
Capital	0.04	0.03	0.00	0.13
Contingency	0.01	0.01	0.00	0.05
Operations	0.26	0.06	0.14	0.40
Salary	0.70	0.06	0.54	0.83
Panel B: Top 10 sectors by budget share				
Bank	0.14	0.06	0.04	0.37
Cleaning	0.09	0.04	0.02	0.24
Code enforcement	0.06	0.01	0.04	0.09
Construction	0.04	0.02	0.01	0.10
Education	0.10	0.02	0.06	0.16
General (woreda) administration	0.14	0.03	0.07	0.24
Job	0.09	0.01	0.06	0.13
SME	0.06	0.01	0.03	0.09
Women's and children's affairs	0.04	0.01	0.02	0.08
Youth and sport affairs	0.04	0.02	0.01	0.08

Notes:

Table E4: Budget descriptive stats FY 2017 (Oromia)

	Mean	SD	Min	Max
Panel A: Woreda-level outcomes				
Approved budget (in 100k, 2025)	40.48	9.69	14.55	81.50
Net adjustment (2025)	2066.64	4015.80	416.59	24825.79
Share of expenditures by function:				
Capital	0.09	0.04	0.02	0.19
Contingency	0.02	0.01	0.00	0.07
Operations	0.10	0.03	0.00	0.17
Salary	0.79	0.06	0.63	0.93
Panel B: Top 10 sectors by budget share				
Agriculture	0.06	0.03	0.00	0.17
Bank	0.08	0.03	0.04	0.19
Civil	0.03	0.01	0.01	0.05
Courts	0.04	0.01	0.01	0.10
Crime	0.05	0.02	0.03	0.13
Education	0.33	0.05	0.18	0.46
General (woreda) administration	0.06	0.03	0.02	0.22
Health	0.12	0.02	0.08	0.20
Livestock development and environmental protection	0.03	0.01	0.00	0.06
Roads	0.03	0.02	0.01	0.07

Notes:

Table E5: Null effect on WCA alignment with citizen priorities

	WCA beliefs about citizen priorities are correct		WCAs and citizens agree on woreda priorities	
	(1)	(2)	(3)	(4)
RC-Woreda	-0.002	0.112	-0.002	-0.033
	[0.046]	[0.080]	[0.061]	[0.155]
RC-Bureau		-0.164		0.042
		[0.103]		[0.184]
Control mean	0.51	0.51	0.37	0.37
Baseline controls	Yes	Yes	Yes	Yes
Sector FEs	Yes	Yes	Yes	Yes
Woreda FEs	No	No	No	No
N	288	288	171	171

Notes: Outcome data from the endline woreda chief administrator (WCA) survey. The unit of observation is the woreda-sector. In columns 1-2, the dependent variable is an indicator equal to 1 if the admin believes a sector (e.g., Education) is a top public priority, conditional on citizens in that woreda identifying it as a top priority. In columns 3-4, the dependent variable is an indicator equal to 1 if the admin names as a woreda priority a sector which citizens identified as a top priority. Citizen priorities were determined based on the average rank individuals gave 5 and 3 sectors, respectively, in terms of priority. In all models, citizen priorities were fixed at their baseline values. All models control for baseline values of the outcome and region. Robust standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E6: Null effect on overall WCA alignment with citizen priorities

	WCA second-order beliefs			WCA preferences		
	Share of citizen priorities recognized as such (1)	$\geq$ 2 priorities recognized (2)	$\geq$ 3 priorities recognized (3)	Share of citizen priorities also admin priorities (4)	$\geq$ 1 priority in common (5)	$\geq$ 2 priorities in common (6)
Treated	0.01	-0.10	0.06	-0.02	0.08	-0.12
	[0.05]	[0.10]	[0.12]	[0.06]	[0.10]	[0.10]
Control mean	0.50	0.90	0.52	0.37	0.76	0.28
N	58	58	58	58	58	58

Notes: Outcome data from the endline woreda chief administrator (WCA) survey. The unit of observation is the admin. In column 1, the dependent variable is the share of citizens' top 5 sector priorities that admins correctly identify as citizen priorities (i.e., share of citizen priorities that matched admins' second-order beliefs). Columns 2 and 3 show effects on how likely admins are to recognize at least 2 or 3 of citizens' top 5 priorities, respectively. Column 4 shows effects on the share of citizens' top 3 sector priorities that matched admins' own priorities. In columns 5 and 6, the dependent variable equals 1 if at least 1 or 2 citizen priorities are also admin priorities, respectively. Citizen priorities were determined based on the average rank individuals gave 5 sectors in terms priority. In all models, citizen rankings were fixed at their baseline values. All models control for baseline values of the outcome and region. Standard errors are clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E7: Opt-in or -out experiment results by leader

	Completed form					
	Opt-in			Opt-out		
	WCA	Bureau manager	p-value	WCA	Bureau manager	p-value
5 min	0.50	0.32	0.29	0.00	0.03	0.56
10 min	0.08	0.03	0.44	0.08	0.06	0.77
25 min	0.00	0.03	0.59			
Total	0.21	0.13	0.27	0.04	0.04	0.97
N	34	102		24	69	

Notes: Outcome data from the demand experiment. The unit of observation is the local leader. We group those who submitted incomplete forms with those who didn't submit at all.

Table E8: Demand for information by tenure

	Completed form					
	Opt-in			Opt-out		
	New	Returning	p-value	New	Returning	p-value
RC-treated only						
5 min	0.33	0.20	0.61	0.00	0.00	
10 min	0.10	0.14	0.80	0.00	0.00	
25 min	0.00	0.00				

Notes: Outcome data from the demand experiment. The unit of observation is the local leader. New refers to admins or bureau managers who started after the baseline survey. We grouped those who submitted incomplete forms with those who didn't submit any.

Table E9: Monitoring-only treatment effect on bureau budget

	(Log) appr. bureau budget		Appr. bureau budget share	
	(1)	(2)	(3)	(4)
Panel A: Monitoring sample, Addis only				
RC-Woreda × Post	-0.007	-0.006	0.001	0.001
	[0.015]	[0.015]	[0.001]	[0.001]
RC-Bureau × Post	-0.035	-0.035	-0.003	-0.003
	[0.032]	[0.031]	[0.002]	[0.002]
RC-Woreda + RC-Bureau = 0	0.175	0.167	0.126	0.132
Control mean	13.61	13.61	0.07	0.07
Woreda FEs	No	Yes	No	Yes
Sector × Year FEs	Yes	Yes	Yes	Yes
N	6,189	6,189	6,197	6,197

Notes: The unit of observation is the woreda-sector-year. Funds earmarked for *General woreda administration* are excluded. The outcome data span 2015–2019. *Post* is defined as 2018 and onwards. All models control for total woreda budget. Standard errors in brackets, clustered by woreda. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E10: Baseline characteristics by admin type

	WCAs				Managers			
	Returning	New	Full BL sample	Returning = New	Returning	New	Full BL sample	Returning = New
Age	37.00	34.70	35.29	0.270	33.43	33.12	33.20	0.530
	[6.06]	[6.46]	[6.39]		[6.12]	[6.17]	[6.15]	
Male	0.93	0.98	0.97	1.000	0.87	0.88	0.87	0.900
	[0.26]	[0.15]	[0.18]		[0.34]	[0.33]	[0.33]	
Ethnicity: Oromo	0.73	0.63	0.66	0.540	0.74	0.68	0.70	0.320
	[0.46]	[0.49]	[0.48]		[0.44]	[0.47]	[0.46]	
University education	0.93	1.00	0.98	0.240	0.92	0.92	0.92	1.000
	[0.26]	[0.00]	[0.13]		[0.27]	[0.27]	[0.27]	
Works in Oromia	0.60	0.49	0.52	0.590	0.60	0.58	0.59	0.830
	[0.51]	[0.51]	[0.50]		[0.49]	[0.49]	[0.49]	
Years lived in woreda	7.92	14.70	12.95	0.020	11.20	11.51	11.43	0.710
	[6.32]	[12.65]	[11.69]		[11.72]	[11.50]	[11.54]	
Years in position	1.41	1.44	1.43	0.940	1.44	1.40	1.41	0.810
	[1.20]	[1.10]	[1.12]		[1.51]	[1.47]	[1.48]	
Years in public sector	13.82	12.16	12.59	0.430	11.07	10.63	10.74	0.440
	[6.33]	[6.09]	[6.14]		[6.48]	[6.52]	[6.51]	
Previously served in political office	0.40	0.50	0.47	0.610	0.20	0.30	0.27	0.040
	[0.51]	[0.51]	[0.50]		[0.40]	[0.46]	[0.45]	
Work hours per week	72.93	74.23	73.90	0.810	63.58	62.49	62.77	0.580
	[14.93]	[13.96]	[14.10]		[14.69]	[15.31]	[15.15]	
Ideal tenure in position (years)	11.96	4.20	6.06	0.030	13.92	12.08	12.53	0.200
	[15.09]	[7.36]	[10.14]		[15.33]	[14.33]	[14.59]	
Satisfied with job	0.53	0.70	0.66	0.360	0.75	0.69	0.70	0.190
	[0.52]	[0.46]	[0.48]		[0.44]	[0.46]	[0.46]	
Joint F-test				0.338				0.586
N	15	43	58		135	388	523	

Notes: Data from baseline surveys. Returning woreda chief administrators (WCAs) started their position before baseline. We report randomization-inference p-values and the result of the omnibus F-test at the bottom.