

# Bureaucratic Deliberation and Performance: Evidence from a Field Experiment in Benin\*

Shan Aman-Rana<sup>†</sup>  
University of Virginia

Leonard Wantchekon<sup>‡</sup>  
Princeton University

Lazare Kovo<sup>§</sup>  
Emory University

August 18, 2023

## Abstract

Designing accountability systems for bureaucracies remains a key challenge for state capacity development. We carried out a randomized control trial at-scale with local bureaucrats in Benin to investigate whether performance improves with deliberation on external evaluation reports. In the treated municipalities, bureaucrats organized internal meetings to deliberate on issues highlighted in audit reports and surveys of bureaucrats and citizens. We find a positive treatment effect on internal performance of bureaucrats (6.8 pp or 8.7% of the control group mean), while there is no immediate effect on service delivery. We investigated the mechanisms and found that the meetings helped bureaucrats update their beliefs about corruption in the local government, and resulted in a fall in trust and cooperation between bureaucrats. These results suggest that in public sector bureaucracies in which the chance of collusion is high (Tirole, 1986), an intervention that leads to a lower level of cooperation and trust may still improve performance.

**JEL codes:** D73, H11, O1

**Keywords:** State capacity, performance, monitoring, accountability, audit, corruption, deliberation, Benin

---

\*We are grateful to Paola Sapienza, Ivo Welch, and participants at seminars and workshops for comments and suggestions. Sanjog Rajiv provided excellent research assistance. Financial support from J-PAL Governance Initiative is gratefully acknowledged. All mistakes are our own.

<sup>†</sup>Department of Economics, University of Virginia. sa8ey@virginia.edu.

<sup>‡</sup>Department of Politics, Princeton University and African School of Economics. lwantche@princeton.edu.

<sup>§</sup>Department of Politics, Emory University. lkovo@emory.edu.

# 1 Introduction

Corruption and state capture continue to pose significant challenges for governments in the developing world (Finan et al., 2017; Besley, 2006; Besley and Persson, 2014). In these contexts, bureaucrats and frontline service providers often exhibit absenteeism (Kremer et al., 2005; Banerjee and Duflo, 2006), or engage in the direct embezzlement of public funds (Niehaus and Sukhtankar, 2013a,b), leading to serious repercussions for efficiency and future state capacity formation (Olken and Pande, 2012; Banerjee et al., 2012b; Ferraz et al., 2012; Aman-Rana et al., 2023).

To overcome these challenges, attention has been mainly restricted to top-down accountability systems in which either politicians or central agencies monitor and control bureaucrats (Olken, 2007; Lichand et al., 2016; Gulzar and Pasquale, 2017). However, such top-down enforcement structures are equally vulnerable to encountering similar challenges (Raffler, 2022; Iyer and Mani, 2012).

An often overlooked alternative to top-down approaches are systems that rely on bureaucrats collectively deliberating on issues related to their own performance and overall service delivery. On the one hand, it can work well by leveraging the bureaucrats' concern for their reputation (Leaver, 2009; Aman-Rana, 2023), or by creating a sense of empowerment,<sup>1</sup> particularly when the reasons for poor performance are information or coordination failures. On the other hand, such accountability systems can fail if the agenda is captured by a smaller group to further their narrow interests or used as a way to waste time. If such collective accountability systems work, these can be a low cost lever in the hands of resource constraint developing countries to improve performance, with implications for the organization of bureaucracies.

In this paper we worked with local administrations in Benin to ask: can collective deliberation by bureaucrats improve bureaucratic performance. While important, studying this question is not straight forward. First, any such study needs a buy-in from the administration to carry out their own accountability through open discussion. Second, for the deliberation to be meaningful it would need to be structured with a clear agenda (Goold et al., 2012). More importantly, we have to be able to control the deliberation environment so that throughout the discussion bureaucrats remain in charge of the agenda (rather than politicians or citizens) and the discussion remains focused on public service delivery issues and the bureaucrats' role in it. Third, the discussion would need to be based on information that is incontrovertible and ideally generated by a third party, so that the effects could be clearly attributed to the deliberation process rather than to the quality of information.<sup>2</sup>

---

<sup>1</sup>There is a recent wave of studies that highlight the importance of autonomy of bureaucrats in public sector bureaucracies for project completion (Rasul and Rogger, 2018), environmental regulation (Duflo et al., 2018), procurement prices (Bandiera et al., 2020), and promotions (Aman-Rana, 2023).

<sup>2</sup>Studies have shown how bureaucrats seek to reduce the efficacy of monitoring and accountability mechanisms by lowering the quality of information generated (Banerjee et al., 2008; Duflo et al., 2013; Dhaliwal and Hanna, 2017).

We overcame these challenges through a field experiment carried out in direct partnership with the government in 20 out of the 77 municipal bureaucracies in Benin. Conducting the experiment in a representative sample of municipalities and at-scale, in collaboration with local governments, helped us effectively tackle a common concern related to external validity that often plagues many randomized control trials (Muralidharan and Niehaus, 2017).

The bureaucrats that work in these local governments are in charge of the delivery of civil status documents (e.g., birth certificates and ID card documents), water and sanitation, and construction and maintenance of local infrastructures (roads, schools, health, and local marketplaces). They work under the mayor and manage the local budget amounting to 23 billion FCFA (USD 38 million) in 2016.

In the ten treated municipalities bureaucrats held three structured monthly meetings starting from September 2016. All meetings had a very clear agenda and used external evaluation reports to deliberate on the ways to improve public service delivery.<sup>3</sup> The research team collected three types of data to support the meetings: previous year's audit reports of each municipality and independent surveys of bureaucrats and citizens from across the 20 municipalities to gather insights about constraints on public service delivery.

The first meeting was focused on the findings from the most recent audit report of the municipality, the second and third meetings (in October and November 2016 respectively) presented results from the bureaucrats and citizens survey.<sup>4</sup> Each meeting lasted on average 95 minutes and had the same format: there was a presentation of empirical results followed by open discussion on the topic discussed. Approximately 40% of the time was spent on formal presentations and 46% on discussion of what was presented.

Our main analysis is based on data that we digitized from the audit reports of municipalities published yearly by The National Commission for Local Finance (CoNaFiL) from 2014-2019. For each municipality auditors compute a performance score (out of a maximum possible score) that is aimed to capture the overall quality of local governance. The auditors emphasis is on the bureaucrats' record keeping, compliance with the regulations especially in relation to public procurement contracts, and the quality of their overall management. The auditors recommendations play a crucial role in ensuring accountability within the context. Multiple studies have described the oversight functions played by the organization as transparent and efficient (Bedasso, 2021; *Decentralisation and Communal Development Support Programme*, 2021; Riedl and Dickovick, 2014; Behanzin et al., 2018).

Our main outcome is the score (as a percentage of the maximum attainable) given by the auditors to

---

<sup>3</sup>The mayor was not invited to any of the meetings to ensure that the bureaucrats controlled the agenda and to ensure that the meetings remained an internal deliberation effort by the bureaucrats.

<sup>4</sup>The presentations of the audit reports in the first meeting highlighted issues reported by the auditors followed by their recommendations on how to overcome these. On the other hand, the second meeting that was focused on the bureaucrats highlighted the state of their knowledge of rules and procedures, their interpersonal relations and problems. The third meeting was focused on the results from the citizen survey and highlighted their experiences with public services and the various problems they faced.

the municipality.<sup>5</sup> The average baseline performance score of the municipalities in our sample is 76 (out of 100), with variation across years (standard deviation 14). There are incentives for the bureaucrats to receive higher scores since well-performing municipalities get extra funds from the central government in future budgets.

We find that after the meetings, auditors' evaluation of the performance of bureaucrats increased by 6.8 percentage points (8.7% of the control mean,  $p$ -value<0.05) in treated municipalities relative to control. Bureaucrats in the treated municipalities had a 9 percentage point (12% of the control mean) higher score in record keeping, 6 percentage point (7.4% of the control mean) higher score in transparency of public procurement and a 14 percentage point (17.2% of the control mean) higher score in the auditors' evaluation of their overall management (although the results for separate components are not statistically significant). This suggests that just three meetings in which issues highlighted by the auditors, bureaucrats and citizens were collectively discussed led to an improvement in how the bureaucrats carried out their internal business.

Using citizen survey data from the Afrobarometer for two waves 2014 and 2017,<sup>6</sup> we find no effects of the deliberation exercise on citizens' perceptions of the local government or their engagement with the bureaucracy, bribes or ease of access to services. It appears that while such accountability-based meetings can work to streamline the internal workings of the bureaucracy, translating these into better public service delivery and an improved experience of the citizen takes a lot more. These results are consistent with studies that have found that policies that succeed at improving specific organizational aspects might nevertheless fail to improve overall performance (Casey et al., 2012; Moreira and Pérez, 2021).

We investigated two main mechanisms behind the positive effects on audit-based performance: the effect of the meetings on bureaucrats' information and the changes in trust and cooperation within the bureaucracy.<sup>7</sup> Using data from an endline bureaucrat survey we carried out in December 2016, our intention-to-treat estimates show that the meetings had a positive effect on bureaucrats' information and a negative effect on their trust and cooperation with each other. Bureaucrats in treated municipalities were 0.4 standard deviations more likely to believe that the private vendors that deliver public services engage in corruption ( $p$ -value<0.01). They were 0.2 standard deviations less likely to believe that public procurement rules were followed ( $p$ -value <0.1). Such negative updating can be one reason why bureaucrats were noted

---

<sup>5</sup>There is also an assessment of the Council's performance by the auditors. These include: whether the council followed rules and regulations and whether the local government's financial situation improved.

<sup>6</sup>The Afrobarometer is an independent, non-partisan and nationally representative survey of Beninese, reaching both rural and urban areas. We preferred to

<sup>7</sup>Our measure of trust is the bureaucrat's trust in specific entities, like the municipal council, office and other colleagues in the City Hall. Anderson et al. (2004) find that such survey measures of trust are significant determinants of contribution levels in a canonical public-goods experiment. Glaeser et al. (2000) critique the more general attitudinal measures of trust like agreement with "most people can be trusted" but find that survey questions that are more specific in nature are predictive of trust in the experimental games.

by the auditors to have become more careful in record keeping, public procurement procedures and overall management. While their beliefs regarding corruption by the private vendors and the council deteriorated, the magnitude of the effects on their beliefs regarding corruption by other bureaucrats were of a lower magnitude and not statistically significant.

While the meetings helped bureaucrats update their beliefs regarding corruption and lack of enforcement of rules, these resulted in a *fall* in trust and cooperation between the bureaucrats. Bureaucrats in treated municipalities had 0.4 standard deviations lower trust in the council ( $p$ -value  $<0.05$ ) and perceived such meetings to negatively impact working conditions. Bureaucrats in treated municipalities believed meetings would increase conflict between colleagues (12 pp., 14% of the mean of the control group,  $p$ -value  $<0.1$ ) and result in lower cohesion at the workplace. Bureaucrats in treated municipalities also reported that deliberative meetings would increase administrative dysfunction (7.5 pp., 7% of the mean of the control group,  $p$ -value  $<0.05$ ) and lead to a deterioration of the working conditions (10 pp., 11% of the mean of the control group,  $p$ -value  $<0.01$ ). On the other hand, they were no more likely than the control group to consider such meetings important. They are also no more likely to want similar meetings to continue.

The idea that deliberation can worsen social interactions between participants has been documented before by studies in both political science and public administration.<sup>8</sup> In our case, a lower trust and cooperation along with an updating of their beliefs regarding corruption resulted in a positive effect on their performance. Taken together these results suggest that the deliberative meetings seem to have generated a strong sense of both vertical and horizontal accountability.

These results can also help explain why such low-cost collective deliberation exercises are not a regular part of the organizational culture in many public sector bureaucracies. If there is a buy-in from the elected representatives or top of the bureaucracy and the meetings are centered around objective information by a third party, then such collective deliberative meetings can be a low-cost way through which governments in developing countries can improve internal accountability at no cost to public service delivery. Such forums are easy to replicate in other contexts: other countries and private firms and can help lower the reliance on external (high paid) experts, making the accountability process more sustainable over time.

Additionally, it appears that in public sector bureaucracies in which the chance of collusion between bureaucrats can be high (Tirole, 1986; Milgrom and Roberts, 1988; de Janvry et al., 2023), a lower level of

---

<sup>8</sup>Deliberative meetings have been shown to increase conflict, result in power play and have been viewed by employees to be a costly activity (see Mendelberg (2002); Thompson (2008) for a detailed review). Roberts (1997) describe two cases in the education sector in the US in which public deliberation was used to significantly reduce a school district's budget and in crafting state educational policy. Although the deliberation process was successful at the end, it was rife with conflict and participants engaging in power play. Mendelberg and Oleske (2000) study discussions about race in New Jersey and find that in diverse groups deliberation failed to lessen conflict, increase mutual understanding and tolerance, or reduce the use of group-interested arguments. Hibbing and Theiss-Morse (2002) present results from focus groups and find that deliberations can fan emotions unproductively, exacerbate rather than diminish power differentials among those deliberating, and can make people feel frustrated with the system that made them deliberate.

cooperation and trust can be helpful to improve accountability and in this case auditors' evaluation of the municipal administration. These results speak to the literature in economics on the negative aspects of social capital (Basu, 1986; Dasgupta, 2005; Bloch et al., 2007; Alesina and Giuliano, 2014) and studies on mafia and "amoral familism" in the sociological literature that suggests that a high level of group cohesion may not always be a useful thing for the overall performance of organizations and societies (Banfield, 1958; Gambetta, 1993; Portes and Landolt, 1996).

**Related literature.** Our paper contributes to the literature on state capacity building, focusing on the political economy of bureaucracies. There is a growing realization that bureaucrats that deliver services are an important determinant of state capacity and that there is a need to open up the black box of internal organization of public sector bureaucracies (Finan et al., 2017). There have been studies that focus on investigating the incentives of bureaucrats (Banerjee et al., 2012a; Callen et al., 2023; Ashraf et al., 2014; Khan et al., 2016; Bertrand et al., 2020; Khan et al., 2018; Khan, 2023; Dipoppa and Gulzar, 2023); and the way to recruit or promote pro-social and talented bureaucrats (Dal Bó et al., 2013; Ashraf et al., 2020; Jia et al., 2015; Riaño, 2021). There has been a small but growing body of recent work that shows the importance of the autonomy of bureaucrats (Rasul and Rogger, 2018; Duflo et al., 2018; Bandiera et al., 2020; Aman-Rana, 2023). Our paper shows a new way through which we can endogenize bureaucratic autonomy.

Apart from the studies already mentioned on accountability systems, others investigate the interaction between the politicians and the bureaucrats and how political incentives and turnover affect bureaucratic performance (Iyer and Mani, 2012; Gulzar and Pasquale, 2017; Akhtari et al., 2022; Callen et al., 2023; Dasgupta and Kapur, 2020). Our paper contributes to the existing literature by showing that there can be contexts and institutional arrangements, even in the public sector bureaucracies in developing countries, in which accountability can work from within the organization. Results show that there might be an active role that bureaucrats can play, at substantially lower costs, in setting their own house in order. The paper therefore, complements the standard approaches to reducing corruption in bureaucracies, (Becker and Stigler, 1974; Besley and McLaren, 1993; La Porta et al., 1999; Di Tella and Schargrodsky, 2003; Burgess et al., 2012; Niehaus and Sukhtankar, 2013b; Corbacho et al., 2016; Debnath et al., 2023) and adds collective deliberation forums organized by bureaucrats themselves as an additional tool available to governments in developing countries to improve accountability. Moreover, the study also highlights the social aspect of state capacity and how accountability mechanisms can lead to a negative effect on organizational cohesion with a positive impact on internal working of the organization.

The second strand of the literature that the paper contributes to is the rapidly expanding Economics, Political Science and Public Administration literature on public deliberation. Public reason and deliberation



have been considered key for establishing the legitimacy of governments (Mendelberg, 2002; Parkinson and Mansbridge, 2012; Hadfield and Macedo, 2012). Studies have shown that bringing voters and politicians together in town hall style meetings can help strengthen the political process, reduce the prevalence of clientelism, and increase parties' vote shares in Benin and Philippines (Fujiwara and Wantchekon, 2013; López-Moctezuma et al., 2022). Wantchekon and Guardado (Forthcoming) show that such meetings can lead to a more informed citizenry, higher electoral participation, lower policy polarization, and more importantly "ethical" voters who respond negatively to vote-buying. In Sierra Leone, Bidwell et al. (2020) find a positive impact of exposure to political candidates' debates on voting behavior, campaign spending, and politicians' performance. Similar positive effects have been observed in other contexts (Collier and Vicente, 2014). While the literature on public deliberation has focused on how increased agency of the voters can strengthen the political process, we complement the literature by presenting empirical evidence showing that agency of the bureaucrats can be just as important as the agency of voters, in improving accountability and performance.

Finally, our paper contributes to the management literature. Previous studies have shown the importance of management skills in improving firm productivity (Bloom and Van Reenen, 2007; Bloom et al., 2013) and the existence of passive waste in public sector organizations (Bandiera et al., 2009). This paper emphasizes the significance of institutional arrangements and highlights that the process is equally vital as the management's skills. Other studies show that there are frictions in communication and knowledge flows in private firms (Coffman, 2014; Papay et al., 2020; Battiston et al., 2021; Sandvik et al., 2021; Menzel, 2021). In a public sector context, this paper shows that collective communication is not just about exchange of skills, but also about improving accountability.

## 2 Context and research design

### 2.1 Context

**Local administration.** In 1999, Benin took a significant step towards decentralization by establishing local governments with political and financial autonomy. This resulted in the creation of 77 municipalities. Each municipality works under the leadership of an elected council and a mayor, who are responsible for their decisions and actions, accountable to the council members.

The local administration within each municipality operates with a hierarchical structure. Bureaucrats, who are employed by the municipality, report to the mayor and carry out various administrative tasks. There are four main offices in the local administration. The office of the General Secretary (GS) oversees

the entire administration and plays a central role in coordinating administrative tasks across different departments and monitoring the execution of the Annual Working Plan (AWP). The public procurement unit (CCMP) drafts and issues public procurement offers to selected private vendors while the office of financial affairs (SAF) focuses on bookkeeping and accountancy of local finances. The office of planning and local development (SPDL) and the the office of local infrastructures (ST) design and monitor the implementation of local development projects drafted by the mayor.

These bureaucrats collectively contribute to the design, execution, and monitoring of the municipality's development plans (PDC), managing local budgets amounting to approximately 23 billion FCFA (USD 38 million).<sup>9</sup> They are compensated with fixed salaries, and the mayor holds significant authority over the career paths of the bureaucrats within the municipality. They have the power to hire, fire, and transfer these bureaucrats between different departments at their own discretion.

According to the World Economic Forum's Global Competitiveness Report (2015), accountability remains a significant challenge in Benin. The report highlights that several provisions in the law regarding citizens' rights to access public information and the Transparency Act's provisions concerning public procurement are not being respected. This lack of adherence to the law contributes to a lack of documentation and transparency in public procurement contracts, which in turn creates an environment conducive to corruption.

Various studies have indicated that local government officials in Benin are often co-opted by private vendors, further exacerbating the issue of accountability. In fact, research has shown that 63% of private businesses prefer utilizing public procurement as a means for state capture (Ch et al., 2019; Canen et al., 2023). These findings underscore the extent to which the influence of private entities continue to undermine accountability in Benin's public procurement processes.

There is also rampant passive waste in the local government due to poor coordination and lack of information across multiple layers of the bureaucracy. This results in low execution rates of municipal budgets. Between 2015–2018 budget execution at the municipal level was only 60 percent, with an execution rate of only 35 percent for capital expenditures (Bedasso, 2021).

**Sources of municipal funds and their accountability.** To support municipalities in meeting their budgets and development needs, the central government introduced the Support Fund for the Development of Municipalities (FADeC) in 2008. This initiative, established through Decree no. 2008-276 on May 19, 2008, aimed to provide financial resources to assist municipalities in meeting their budgets and development needs.

---

<sup>9</sup>The PDC encompasses the mayor's policies, various development projects, the annual budget, and the Annual Working Plan (AWP).



Transfers from FADeC are categorized into two types: general-purpose grants and investment grants. The general-purpose grant is allocated to cover the recurrent expenditures of local governments. On the other hand, the investment grant is specifically designated for funding local public goods and services.

The investment grant further consists of two components: an earmarked grant and a formula-based grant. The earmarked grant is intended to target specific policy areas such as health and education. It is allocated to support initiatives and projects related to these specific sectors within the municipalities. The formula-based grant, however, is disbursed based on various socio-economic and performance indicators. These indicators include factors such as the municipality's geographic area, population size, poverty levels, as well as the administrative and financial performance of the municipalities in previous fiscal years. The assessment of administrative and financial performance is conducted by the National Commission on Local Finance (CoNaFiL), while the socio-economic indicators are provided by the National Institute of Statistics (INSAE). FADeC investment grants serve as a significant source of financing for local public goods and services, constituting nearly 90% of the funding.

The National Commission for Local Finance (CoNaFiL), in collaboration with the Office of the Auditor General, conducts yearly audits across all municipalities (see FADeC Procedural Manual, pp.69-73).<sup>10</sup> During these audits, auditors assess the management of funds received by municipalities, focusing on areas such as record-keeping, transparency in public procurement contracts, compliance with procurement regulations, and the quality of bureaucratic management. Additionally, auditors evaluate the financial situation of local governments and assess whether the council adhered to budget adoption rules, administrative account regulations, and voting deadlines, among other criteria.

Following the completion of the audits, auditors prepare reports that summarize the key findings. These reports also include scores assigned to each sub-component of the municipality's performance that was evaluated (refer to Appendix C.1 for a detailed list of sub-components and their corresponding maximum scores). These scores are of great importance as bureaucrats are incentivized to achieve high scores. Municipalities that perform well in the audits may receive additional funds for future transfers, as stated in the FADeC Procedural Manual (p.18). Appendix Table A.1 presents evidence for the municipalities in our sample. A one percentage point higher rating by the auditors results in a 2 million XOF higher investment transfer by the central government under FaDeC in the following year (2 pp. of the total FaDec investment transfers to a municipality in a year). This suggests that the bureaucracy faces high incentives to perform well on the auditors' criteria of performance evaluation.

Multiple studies have highlighted the importance and effectiveness of the National Local Finance Commission (CoNaFiL) in overseeing municipalities (Riedl and Dickovick, 2014; Behanzin et al., 2018).

---

<sup>10</sup>Since the audits are based on a municipality's utilization of FADeC disbursed by the central government yearly, CoNaFiL conducts the audits at the end of each administrative year (January-December).

CoNaFiL possesses the necessary technical capabilities and the authority granted by the central government to ensure accountability in the management of funds (Bedasso, 2021). According to the International Monetary Fund’s Public Investment Management Assessment (PIMA) of Benin, the country’s regulatory framework for public investment is of high quality, and Benin outperforms its peers in Sub-Saharan Africa in this regard (Imbert et al., 2018). Additionally, municipalities consider these reports to be more than just instruments for financial accountability. They view them as essential tools for assessing their overall performance and driving improvements.<sup>11</sup>

## 2.2 Experimental design

This study was implemented in collaboration with the Institute for Empirical Research in Political Economy (IERPE) and the African School of Economics (ASE) in 20 randomly chosen municipalities across 8 regions in Benin (see Figure 1). The experiment was the result of multiple meetings with the top bureaucrats in the municipal administration (the General Secretary) to create a buy-in for the deliberation exercises as a way forward for self-accountability of the bureaucracy. There was an agreement that the municipalities would engage in deliberation exercises and the deliberation environment would be controlled. This was to ensure that bureaucrats would remain in charge of the agenda (rather than politicians or citizens) and the entire discussion would remain focused on public service delivery issues and the bureaucrats’ role in it. The bureaucrats were, however, free to invite other participants such as councilors, civil society representatives, bureaucrats from the central government or media personnel if they deemed it appropriate for any particular meeting agenda.

Figure I describes the timeline of the experiment. In August 2016 we collected information that was to be presented during the meetings. In September, October and November 2016 the first, second and third collective deliberation meetings, respectively were held. In December 2016 the endline survey of bureaucrats was carried out. We describe each of these steps below.

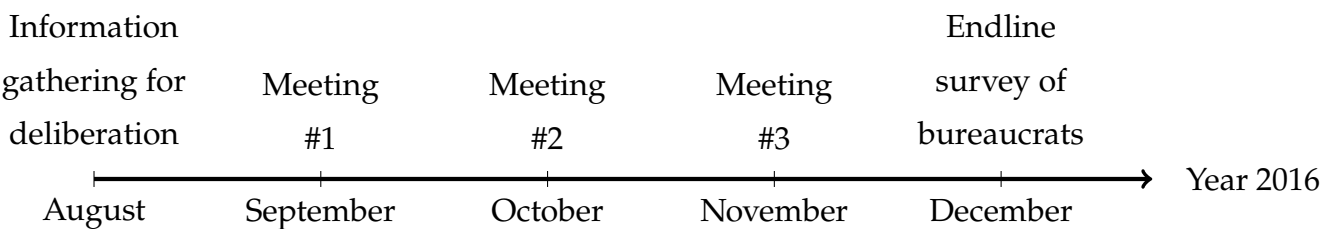


Figure I: Experiment timeline

<sup>11</sup>To follow up on the audit recommendations, a committee is established at the local level (*Decentralisation and Communal Development Support Programme, 2021; Bedasso, 2021*).

**Generating information for the deliberation exercises.** For the deliberation exercises to be meaningful they had to be based on information that is incontrovertible and generated by a third party, so that the effects could be clearly attributed to the deliberation process rather than to quality of information. To provide a holistic picture of the constraints on public service delivery we collected three types of information. The first set included audit reports of each municipality published by the central oversight body, CoNaFiL. These were then complemented by independent survey of bureaucrats and citizens from across the 20 municipalities.

The survey of bureaucrats included questions on their basic knowledge of rules and regulations, their administrations and ongoing projects in the municipality; their perceptions of the role of the administration in improving education and health institutions, the markets in the municipality, maintenance of roads, and providing safe water and sanitation to the area; and the nature of their interactions with their colleagues.

For the survey of citizens a representative sample of hundred individuals<sup>12</sup> were sampled from each municipality. The survey focused on the local populations' involvement in local politics and development affairs and their demand for better services, their perceptions of local governance and corruption, their trust in the local institutions, their satisfaction with performance of the local government.

**Meeting #1: Deliberation on audit reports and recommendations.** The first meeting was held in September, 2016. This meeting was based on the audit reports and the auditors' recommendations. Table 1 describes details. At the start of each meeting attendance was taken. The total number of participants in the first meeting were 31, out of which 22 were bureaucrats. Figure 2 shows details of the other participants that attended each of the three meetings. The General Secretary attended the meetings in all the municipalities except one. The total duration of the meeting was 100 minutes on average, 34% of this time was spent on presentation of the key findings and 47% was spent discussing those findings. Figure 3 and 4 describe the details of the time spent presenting and discussing different topics. Administration functioning and financial autonomy of the municipalities took the maximum time of the meeting. Enumerators reported approximately 7 suggestions and criticisms and 4 questions during the discussion. The meeting was considered interactive and information and there were no tensions reported.

**Meeting #2: Deliberation on the experiences of bureaucrats.** The second meeting was held in October, 2016. This meeting was based on the results from the survey of the bureaucrats across the 20 municipalities. Table 1 shows that the total number of participants were 32, out of which 24 were bureaucrats (see Figure 2 for details). The General Secretary attended the meeting in all the ten municipalities. The total duration of

---

<sup>12</sup>Exception: 112 citizens were surveyed in the municipality of Ouinhi, making the total number of observations for citizens' survey to be 2012.

the meeting was 92 minutes on average, 42% of this time was spent on presentation of the key findings and 41% was spent discussing those findings. Figure 3 shows that the bureaucrats' understanding of rules and procedures, including their perceptions of the public procurement procedures and availability of financial and administrative reports was presented for a longer time than other aspects. Figure 4 shows that the duration of the discussion was equally split among all the topics, with the least amount of time spent on their interpersonal relationships. There were three questions asked, and there were more criticisms (6) than suggestions (4).

**Meeting #3: Deliberation on service delivery and the experiences of citizens.** The third meeting was held in November, 2016. This meeting was based on the results from the survey of the citizens. Table 1 shows that the total number of participants were 34, out of which 23 were bureaucrats (see Figure 2 for details). The General Secretary attended the meeting in all the ten municipalities. The total duration of the meeting was 94 minutes on average, 42% of this time was spent on presentation of the key findings and 48% was spent discussing those findings. The discussion remained largely focused on citizen's trust in the state and corruption and their perceptions of service delivery by the local government (see Figures 3 and 4). Only 60% enumerators reported that the meeting was informative, making this the least informative of all the three. While the enumerators reported that the atmosphere was not tense, this meeting had the highest number of criticisms and questions during the discussion period and the least number of suggestions.

**Compliance.** 173 out of the 299 surveyed bureaucrats in the treated municipalities reported to have attended at least one of the three meetings i.e. a take up rate of 57.86%.<sup>13</sup> Given this first stage, we focus below on the intent-to-treat (ITT) estimates, which can be interpreted as the average treatment effects corresponding to an approximately 60% complete implementation. The ITT is the relevant policy parameter since deliberation exercises in any bureaucracy will remain voluntary. In Appendix Tables A.2 and A.3 we also report results using an instrumental variables strategy.

### 2.2.1 Integrity of the experiment

Figure 5 shows the balance tests for two categories of covariates: bureaucrat-level characteristics and municipality-level characteristics. We standardized the variables and plotted their mean difference w.r.t. treatment status along with 95% confidence intervals. Standard errors are clustered at the municipality level. All the covariates are balanced across the treated and control municipalities at the 5% significance

<sup>13</sup>Among the 299 surveyed, 116 (38.8%), 118 (39.46%) and 120 (40.13%) reported to have attended the first, second and third meeting respectively.

level. However, among the municipality-level characteristics, indicator variables for the presence of yearly accountability sessions and existence of procedure for acquisition of equipment were significantly lower in the treated municipality relative to the control municipality (significant at the 10% level). In the estimation sections below we explain how we control for these.

All the municipalities complied with the treatment protocol. Meetings were only held in the treated municipalities and there were no deliberation exercises in the control municipalities. Additionally, bureaucrats are not transferred outside the municipalities in which they are hired, however, 1.7% (5 out of 296) bureaucrats from control municipalities reported to have attended at least one meeting that was being held in treated municipalities.<sup>14</sup> All the analysis using bureaucrat surveys, therefore, reports the intention-to-treat estimates.

We ensured that the bureaucrats did not feel that their activity during the meeting was being monitored. Enumerators sat in the meetings discreetly, just like any other participant, and refrained from compiling their reports until after the meeting had concluded.

Additionally, bureaucrats were assured that their responses to the survey were confidential and would only be discussed in the aggregate. Responses from specific bureaucrats or even average responses in any particular municipality were never discussed during any meeting. Only the principal investigators have access to that data.

## 2.3 Data sources and key variables

**Bureaucrats' audit-based performance from audit reports.** We digitized data on the evaluations of the bureaucrats by the auditors using the annual reports published by CoNaFiL. This includes the raw score attained for each criterion on which the government is judged along with maximum achievable score. Following this digitization we created a panel of 20 municipalities between 2014-2016.<sup>15</sup> In a given year, each municipality is scored on a number of criteria including, record keeping, transparency of public procurement and overall management (see Appendix Tables C.1-C.4 for a detailed description of the criteria used across years along with their maximum possible scores in each year).<sup>16</sup>

To calculate the bureaucrats' audit-based performance in a municipality-year we summed the scores earned by a municipality as a percentage of maximum attainable scores. To calculate sub-components of

---

<sup>14</sup>Three bureaucrats from Za-Kpota claimed to have attended last two meetings, one bureaucrat each from Toviklin and Kouande reported to having attended only the first meeting and all meetings respectively.

<sup>15</sup>In case any audit report or a score for a criterion was missing for a particular year, to fill in the missing value we used the subsequent year's "Evolution Table" that describes how the performance on a particular criterion has changed across the years. There are audit reports on eight municipalities that are missing for the year 2018.

<sup>16</sup>Note that the criteria under Overall Management are not present in reports for 2014 and 2015.

the bureaucrats' audit-based performance we used a similar approach but restricted attention to the set of relevant criteria for that sub-component.

Figure 6 shows the baseline evaluation of the bureaucrats by auditors in treated and control municipalities.<sup>17</sup> The average score is 78% in control municipalities (14.7 standard deviation), while it is 75% in treated municipalities (14.4 standard deviation). Appendix Figures B.1 to B.3 shows the baseline scores in record keeping, transparency of public procurement and overall management.<sup>18</sup>

**Survey of bureaucrats' beliefs and their perceptions of the meetings.** In December 2016, we carried out a survey of a random sample of 30 bureaucrats in each of the 20 municipalities that are part of the study.<sup>19</sup> Our sampling frame were all bureaucrats working in the local government, rather than those that attended the meetings. This was to ensure that we have a representative sample of bureaucrats rather than a selected sample, since the choice to attend the meeting is correlated with unobserved characteristics of the bureaucrats.<sup>20</sup> The survey resulted in 299 bureaucrats in treated municipalities and 296 bureaucrats in control municipalities for a total of 595 bureaucrats. Figure B.4 presents the distribution of surveyed bureaucrats across each of the treated and control municipalities.

The survey had questions regarding the bureaucrats' beliefs about corruption by the Council, private vendors and other bureaucrats; their views on whether rules regarding public procurement and recruitment were enforced, and their trust in the Council and their colleagues. The bureaucrats were also asked if they found such meetings to be helpful in solving issues they faced at work. The survey included questions about their working environment including indicators for whether they believed that the meetings would help solve conflicts between colleagues, reduce administrative dysfunction, improve poor working conditions, and increase cohesion at the workplace.

We also created a cooperation index that included ordinal variables for how often the bureaucrats believed they engaged in the following activities with their co-workers: discuss work together, organize recreational activities, organize professional activities like workshops, offer to perform indisposed col-

---

<sup>17</sup>Appendix Figures C.1 to C.4 present the overall performance scores for the treated and control municipalities at baseline as well as the scores that were related to the Council's performance.

<sup>18</sup>At baseline, the auditors' score regarding the government's record keeping is 78% in the control municipality (18.4 standard deviation) and 73% in treated municipality (19.6 standard deviation). The auditors' score on transparency of public procurement is 78% in control municipality (18 standard deviation) and 76% in treated municipalities (17 standard deviation). The auditors' score on the bureaucrats' overall management is 83% in control municipalities (15 standard deviation) and 72.5% in treated municipalities (28 standard deviation).

<sup>19</sup>The average size of the bureaucracy is 63 bureaucrats, ranging from 30 in Toucountouna and 101 bureaucrats in Ze. We were not able to reach 30 bureaucrats in the municipalities of Pehunco (29), Toffo (29) and Ze (27).

<sup>20</sup>Within each municipality, thirty bureaucrats in each municipality were surveyed at baseline in July-August 2016. We were able to survey less than thirty in the following municipalities: Gogounou (26), Grand-Popo (25), Houeyogbe (24), Pehunco (29) and Toucountouna (19). Not all questions covered in the endline survey were included at baseline. Therefore, we carry out a cross-sectional analysis of the bureaucrats' response and use the baseline data to test for balance.



league's work, solve conflict with each other, and refer a work problem to the top managers to resolve it.<sup>21</sup>

**Afrobarometer data on public service delivery to the citizens** To investigate the effect of the intervention on the citizens' lives we used the Afrobarometer Survey data<sup>22</sup>. This is a repeated cross-section survey of citizens conducted every three years across several countries in Africa, Benin being one of them.<sup>23</sup> We use the 2014 and 2017 waves of the survey as pre and post treatment periods respectively for our analysis.<sup>24</sup>

The analysis is based on indicators of whether the citizens trust the councilors, feel that they are listened to, whether they approve of the performance of the councilors, believe that there is corruption in council and government officials and whether the latter go unpunished. We also included indicators of citizen engagement with the local councilor, govt officials, whether they complained or are willing to complain about poor services or any issues they face. Finally, the data also includes indicators of whether the citizens feel that their present living conditions are good and are better compared to others, and indicators of the citizens' experiences with obtaining services like accessing birth certificate, driver's license, passport, voter's card or permit from government and, household services (like water, electricity and sanitation), or had to pay bribes.

### **3 Did collective deliberation by bureaucrats improve the auditors' evaluation of their performance and public service delivery?**

We start the analysis with an investigation of whether bureaucrats can hold themselves accountable through collective deliberation meeting. We report results using both the data from the audit reports as well as the data from the Afrobarometer on citizens' experience with public service delivery. We find that while collective deliberation exercises can work to streamline the internal workings of the bureaucracy as reported by the auditors however, this is not enough to improve services for the citizens.

---

<sup>21</sup>This was measured on a Likert scale ranging from 1 to 4, 1 being never and 4 being always. For the variables for which we constructed an index we first took an average across sub-categories of a question for each bureaucrat. These were then standardized, using the mean of all bureaucrats and dividing by the standard deviation.

<sup>22</sup>Accessed at: <https://www.afrobarometer.org/countries/benin/>

<sup>23</sup>Figure B.5 presents the distribution of surveyed citizens across municipalities and over time.

<sup>24</sup>Appendix Table A.4 shows the results including data from the 2021 round in addition to 2014 and 2017.

### 3.1 Estimation

For municipality  $m$ , year  $t$  we estimate:

$$y_{mt} = \alpha_m + \alpha_t + \beta T_{mt} + e_{mt} \quad (1)$$

where  $y_{mt}$  are outcomes including the auditors' evaluation of the performance of the bureaucrats (measured as a percentage of the maximum score), and measures of the citizens' experiences with public service delivery like their perceptions of the local government or their engagement with the bureaucracy, their living conditions, bribes or ease of access to service.  $T_{mt}$  is a dummy variable that turns on one for the treated municipalities in 2017-2019, and remains zero otherwise.<sup>25</sup>  $\alpha_m$  are municipality fixed effects,  $\alpha_t$  are year fixed effects,  $e_{mt}$  is the error term clustered at the municipality level as that is the level of the treatment (Abadie et al., 2017).

For all estimates of  $\beta$  we report the municipality-level clustered standard errors and  $p$ -values, as well as the  $p$ -values from a two-sided randomization inference test as suggested by Young (2019).<sup>26</sup> Randomization inference procedures provide the exact tests of sharp (i.e. precise) hypotheses and have the advantage of providing inference with correct size regardless of the sample size, regression design or characteristics of the disturbance term. This test consists of reassigning the treatment and control status in the sample and reestimating  $\beta$  using this placebo assignment multiple times (10,000 replications as in Young (2019); Ashraf et al. (2020)). Under the null of zero treatment effects, the  $p$ -value is based on the proportion of reestimated  $\beta$ 's that are larger (in absolute value) than the actual  $\beta$ .

### 3.2 Results: Auditors' evaluation of bureaucratic performance

We first present the distribution of the bureaucrats' audit-based performance using the raw data and then present results from estimating Equation 1. Figure 7 presents the kernel density in treated (blue line) and control municipalities (red line), pre (left panel) and post treatment (right panel).  $p$ -values from Kolmogorov-Smirnov test are presented at the bottom of each panel. In the period before the treatment (2014-2016), the auditors' evaluation of the bureaucrats' performance is not systematically different from each other ( $p$ -value of the Kolmogorov-Smirnov test is 0.8). On the other hand, following the collective deliberation meetings, the distribution of the bureaucrats' performance is shifted to the right in treated municipalities as compared to the control municipalities (although  $p$ -value of the Kolmogorov-Smirnov

---

<sup>25</sup>We carried out the experiment in the last quarter of 2016, we therefore, consider 2016 as the period before the treatment since 8 out of 12 months on which the audit reports are based are from before any meetings.

<sup>26</sup>We computed these using `ritest` command in STATA with 10,000 iterations.

test is 0.493).

Table 2 presents results from estimating Equation 1. The dependent variables are in rows and the estimates of  $\beta$  are in Column (2). The key takeaway from Table 2 is that when bureaucrats hold meetings in which they discuss their own service provision their internal performance improves. Column (2) shows that after the meetings, auditors' evaluation of the performance of bureaucrats increased by 6.8 percentage points (8.7% of the control mean,  $p$ -value<0.05) in treated municipalities relative to control.

We also unpack the overall effects to see which of the sub-components evaluated by the auditors is the main driver of the effects. Results in Table 2 show that the bureaucrats improved their audit-based performance in all the components (although the results for separate components are not statistically significantly). The effects were the highest for the bureaucrats' overall management. While they had a 9 percentage point (12% of the control mean) higher score in record keeping, 6 percentage point (7.4% of the control mean) higher score in transparency of public procurement, their score on overall management was 14 percentage point (17.2% of the control mean) higher.<sup>27</sup>

We explore these results further to see whether the effects of the meetings persist for multiple years or only materialize in the short-run. Since the experiment is based on 20 municipalities we are not powered to precisely detect separate effects for multiple years. We therefore, consider this as more of an exploratory exercise rather than a definitive test of the dynamic effects of the treatment between 2014-2019. For each municipality ( $m$ ) and year ( $t$ ) we estimate:

$$y_{mt} = \tau_m + \tau_t + \sum_{g \neq 2016} \rho_g T_m \times [\mathbb{1}(g = t)] + e_{mt} \quad (2)$$

where  $T_m$  is an indicator for treated municipalities interacted with year dummies from 2014-2019 (except 2016).  $\tau_t$  are the year fixed effects and  $\tau_m$  are the municipality fixed effects. All the rest of the variables are the same as in Equation 1. The coefficient  $\rho_g$  estimates the effect of belonging to the treated municipalities for each year from 2014-2019. Figure 8 plots  $\rho$  and 95% confidence intervals for each year. While the bureaucrats' audit-based performance is statistically indistinguishable from zero in the treated and control municipalities before the treatment, a substantial (but statistically insignificant) positive effect materializes after the meetings. The magnitude of the effects is the highest in the first year after the meetings and can be seen to be lower in the subsequent years, but never reverts to the same level as before the treatment.

<sup>27</sup>In Appendix Table A.5 we present the effect of the meetings on the auditors' overall evaluation of the municipalities as well as their evaluation of the local government council. Overall performance is based on the bureaucrats' and council's performance. Council's performance is a score constructed from 2 sub-components: (i) local government finance and (ii) rule following by the council. Results show that there is a positive yet statistically insignificant effect of the meetings on the overall performance of the treated municipalities. On the other hand, the effect of the meetings on the auditors' evaluation of the council's performance is negative but statistically insignificant.

These results suggest that collective deliberation meetings can be a low cost lever in the hands of governments in developing countries to improve the internal workings of public sector bureaucracies. The fact that the effect persisted over three years suggests that the meetings did not just work through increasing salience of the issues.

Next we test whether such accountability exercises can also lead to an improvement in service delivery and whether citizens experienced any positive changes in their outcomes following these.

### 3.3 Results: Citizens' experiences with public service delivery

Table 3 presents the results using the 2014 and 2017 waves of the Afrobarometer data. The rows are dependent variables while Column (2) presents the main estimates of interest. Results show that there was no effects of the meetings on the public service delivery to the citizens. The magnitude of the effects on citizens' perceptions of both the bureaucrats, the council and their engagement with the government (Panel A and B) remains small and not statistically significant.

Moreover, the magnitude of the effects on citizens' living conditions (Panel C), while large and positive remains statistically insignificant. Panel D and E report results looking at the effects of the meetings on the difficulty in accessing different services and the probability of paying bribes to get services, respectively. There is a drop in observations for these variables since these responses are only recorded for those citizens that have received a service. The effects on accessing services are small in magnitude, not precisely estimated and statistically insignificant. While there is a large and relatively precisely estimated fall in bribes paid for household services ( $p$ -value=0.11, randomization inference  $p$ -value=0.12), there does not appear to be any systematic evidence of collective deliberation by bureaucrats affecting a positive change in the experience of citizens.<sup>28</sup>

Therefore, it appears that while meetings can improve accountability and streamline the internal workings of the bureaucracy, converting this into better public service delivery and a better experience of the citizen is not guaranteed. The fact that these meetings improved one aspect of the organizational performance but not others resonates with results in previous work by Casey et al. (2012); Moreira and Pérez (2021).

One interpretation of these results in this context is that the information presented in the meeting on audit reports was relevant and the bureaucrats had the power and the incentives to bring about a change. There is an existing culture that takes the reports from CoNaFiL seriously. Municipalities have a committee that convenes regular meetings to review performance in the audit evaluations (Bedasso, 2021).

---

<sup>28</sup>In Appendix Table A.4 we report results including the 2021 wave of the Afrobarometer. Results remain largely unchanged.

Additionally, the audit reports carried weight for future funding received by the municipalities, while there were no such incentives for improving the experiences of the citizens. The intervention therefore, had all the three components that can make for an impactful intervention as suggested by [Kosec and Wantchekon \(2020\)](#). The presentation was on concrete issues reported by the auditors along with suggestions for the bureaucrats to address each of the issues. These concrete recommendations were missing from the presentation of the citizens survey results.

The results could also be due to intense debates during the third meeting that discussed the citizens' survey with few conclusions. As shown before, 40% of the enumerators attending the third meeting (discussing citizen survey) found the meeting to be less informative (see Table 1), this was not the case with the other meetings.

## 4 Mechanism: Effects on information or trust and cooperation within the bureaucracy

In this section, we examine the underlying mechanisms through which deliberation can influence performance. There are several potential channels through which deliberation can impact performance, and in this particular section, we focus on two key aspects: the effect on the bureaucrats' information regarding corruption and enforcement of rules and regulations, and their trust and cooperation with others in their work environment.

Following the meetings, bureaucrats might have updated their beliefs about corruption in the administration and become more careful about how they kept the official records, or whether public procurement rules were followed. These could have equally encouraged the General Secretary to take a more proactive role in effectively coordinating local services.

The meetings could have also impacted trust and cooperation among bureaucrats. On the one hand, deliberation can create a space for open dialogue, increase trust, exchange of ideas, and collaborative problem-solving. This enhanced cooperation can contribute to improved performance outcomes. On the other hand, deliberation could also result in a *fall* in trust and cooperation. This can be the case if the meetings take on an accountability flavor and the meetings allow a forum to evaluate the bureaucrats.

### 4.1 Estimation

To investigate the mechanisms we used the endline survey data from bureaucrats. For each individual bureaucrat  $i$  in municipality  $m$  we estimate:

$$y_{im} = \alpha + \phi T_m + \theta X_m + u_{im} \quad (3)$$

where  $y_{im}$  includes standardized variables measuring both information of the bureaucrats and their trust and cooperation with others.

More specifically, we test the effects of the meetings on the bureaucrats' information using data on the extent to which bureaucrats believe that councilors, other bureaucrats (that work in their own and other departments) or private vendors are engaged in corruption, and the extent to which they believe enforcement of rules is followed for public procurement and hiring staff.

Bureaucrats' trust and cooperation is measured using the extent to which they trust the councilors and other bureaucrats (that work in their own and other departments); whether bureaucrats perceived the meetings to improve their working environment; whether they engaged in the following activities with their colleagues: discuss work together, organize recreational activities, organize professional activities like strategic retreats or workshops, offer to perform indisposed colleague's work, solve conflict with each other, or refer any conflict to their top management. We also asked them whether they considered these meetings to be important and useful for the future.

$T_m$  is a dummy variable that turns on one for the treated municipalities and remains zero otherwise.  $X_m$  are municipality-level control variables that include average baseline age of bureaucrat, whether they possess a college degree, whether yearly accountability sessions were held at baseline, and whether there is an equipment acquisition procedure in place for the municipality.  $u_{im}$  is the error term clustered at the municipality level as that is the level of the treatment (Abadie et al., 2017).

As before in estimating Equation 1, for all estimates of  $\phi$  we report the municipality-level clustered standard errors and  $p$ -values, as well as the  $p$ -values from a two-sided randomization inference test as suggested by Young (2019).<sup>29</sup>  $\phi$  is the main intention-to-treat effect that shows the effect of being offered the collective deliberation meetings on the bureaucrats' beliefs and their trust in others.

## 4.2 Results

Table 4 presents the ITT estimates for the effects of the meetings on bureaucrats' beliefs regarding corruption and lack of enforcement of rules, while Table 5 presents the results for trust, working environment and cooperation index. The outcome variables are in the rows and Column (2) presents the treatment effects for each of these outcomes.

<sup>29</sup>We computed these using `ritest` command in STATA with 10000 iterations.



We find that following the meetings bureaucrats in the treated municipalities positively updated their beliefs about corruption by private vendors that deliver public services (Panel A). Bureaucrats in treated municipalities were 0.4 standard deviations more likely to believe that these vendors engaged in corruption ( $p$ -value $<0.01$ ). They also positively updated their beliefs about corruption by other bureaucrats and council members, however the effects are not statistically significant.

Results in Panel B show that after the meetings bureaucrats were 0.2 standard deviations less likely to believe that public procurement rules were followed ( $p$ -value $<0.1$ ). The magnitude of the effects on beliefs about enforcement of recruitment rules were similar but less precisely estimate ( $p$ -value $>0.1$ ).

In Table 5 we analyze the effects of the meetings on trust and cooperation at the workplace. Results show a significant decline in trust. Bureaucrats in the treated municipalities had a 0.4 standard deviations lower trust in the council ( $p$ -value $<0.05$ ). This fall is only 0.08 standard deviations in the case of other bureaucrats and it is statistically insignificant ( $p$ -value $>0.1$ ). These results suggest that the collective deliberation meetings seem to have created a strong sense of vertical accountability as compared to a sense of horizontal accountability. Such negative updating can be one reason why bureaucrats were noted by the auditors to have become more careful in record keeping, public procurement procedures and overall management.

Table 5 Panel B shows the effects of the meetings on their working environment. Bureaucrats in the treated municipalities considered such deliberation exercises to be 12 pp. (14% of the mean of the control group,  $p$ -value  $<0.1$ ) less likely to solve conflict between colleagues, 7 pp. (7.9% of the mean of the control group,  $p$ -value  $<0.05$ ) less likely to solve administrative dysfunction, 9 pp. (10% of the mean of the control group,  $p$ -value  $<0.05$ ) less likely to improve poor working conditions, and 11 pp less likely to create recreational activities for the workplace (not statistically significant) than the control municipalities. Bureaucrats in treated municipalities viewed their working environment index to be 0.4 standard deviations lower ( $p$ -value  $<0.05$ ), relative to control group.

In Panel C we test the effects of the meetings on cooperation between bureaucrats.<sup>30</sup> Bureaucrats in the treated municipalities had a fall in cooperation, however the effects are not statistically significant. Finally, they are not more likely than the control group to demand such meetings or consider these as a way to improve accountability.

This negative impacts of deliberation are not anomalous and are close to a large literature in both political science and public administration (Mendelberg, 2002; Thompson, 2008; Roberts, 1997; Mendelberg and Oleske, 2000; Hibbing and Theiss-Morse, 2002). Since such meetings were meant to improve

---

<sup>30</sup>We constructed an index using bureaucrats' responses to whether they discussed work together, organized recreational activity, organized workshops or strategic retreats, performed an indisposed colleague's work, solved conflict with each other, and reported conflict to hierarchy.

accountability in our context that may have been another reason why it negatively impacted collegiality between bureaucrats.

While the ITT is the policy parameter of interest, we also present estimates from instrumenting the bureaucrats' participation in the meetings with a dummy for whether the municipality was randomized to receive the treatment. These estimates are the local average treatment effects of the meetings and inform us about the magnitude of the effects on those bureaucrats that attended the meetings. Appendix Tables A.2 and A.3 show the results. The LATE are consistent with the ITT and, as expected, are larger in magnitude.

Together these results suggest that the meetings affected performance both through the provision of information and a *lowering* of cooperation between bureaucrats. The lack of demand for such meetings also suggests that for such meetings to be a part of the organizational culture there needs to be a push from the very top of the hierarchy. What is noteworthy is that these improvements in bureaucratic performance and accountability can be achieved without incurring additional costs to public service delivery, making it a valuable approach for enhancing governance and administrative effectiveness.

## 5 Conclusion

Meetings in which bureaucrats collectively deliberate on their performance can serve as an additional tool available to governments in developing countries to enhance accountability. In an at-scale field experiment with bureaucrats working in local governments of Benin we show that bureaucrats can hold themselves accountable through collectively deliberating on externally generated reports. At the end of just three meetings we see an improvement in internal working of the municipal governments, at no cost to public service delivery. The paper therefore presents a different lens of looking at problems of public service delivery, leveraging the agency of bureaucrats to overcome such issues. Since such deliberation exercises are low cost, institutionalizing such activities in resource starved countries can be a way forward to improve internal accountability.

The important components that can make it work are: a buy-in from the elected representatives or top of the bureaucracy, and structured meetings based on objective information by a third party.

A valuable avenue for future investigation would be to explore how such forums can be institutionalized and refined, leading to improved internal performance and better public service delivery. One potential refinement could involve transitioning from simple staff consultation to creating additional platforms that facilitate issue resolution and enable drafting and voting on resolutions.

## References

- Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge**, “When should you adjust standard errors for clustering?,” Technical Report, National Bureau of Economic Research 2017.
- Akhtari, Mitra, Diana Moreira, and Laura Trucco**, “Political turnover, bureaucratic turnover, and the quality of public services,” *American Economic Review*, 2022, 112 (2), 442–493.
- Alesina, Alberto and Paola Giuliano**, “Family ties,” in “Handbook of Economic Growth,” Vol. 2, Elsevier, 2014, pp. 177–215.
- Aman-Rana, Shan**, “Does Ability Matter for Discretionary Promotions in Bureaucracies? Evidence from Pakistan,” *Working Paper*, 2023.
- , **Clement Minaudier, and Sandip Sukhtankar**, “Informal fiscal systems in developing countries,” *Available at SSRN 4380447*, 2023.
- Anderson, Lisa R, Jennifer M Mellor, and Jeffrey Milyo**, “Social capital and contributions in a public-goods experiment,” *American Economic Review*, 2004, 94 (2), 373–376.
- Ashraf, Nava, Oriana Bandiera, and Scott S Lee**, “Awards unbundled: Evidence from a natural field experiment,” *Journal of Economic Behavior & Organization*, 2014, 100, 44–63.
- , —, **Edward Davenport, Scott Lee et al.**, “Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services,” *American Economic Review*, 2020, 110 (5), 1355–94.
- Bandiera, Oriana, Andrea Prat, and Tommaso Valletti**, “Active and passive waste in government spending: evidence from a policy experiment,” *American Economic Review*, 2009, 99 (4), 1278–1308.
- , **Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat**, “The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats,” Technical Report, National Bureau of Economic Research 2020.
- Banerjee, Abhijit and Esther Duflo**, “Addressing absence,” *Journal of Economic perspectives*, 2006, 20 (1), 117–132.
- , **Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh**, “Can institutions be reformed from within? Evidence from a randomized experiment with the Rajasthan police,” Technical Report, National Bureau of Economic Research 2012.

- , **Sendhil Mullainathan, and Rema Hanna**, “Corruption,” Technical Report, National Bureau of economic research 2012.
- Banerjee, Abhijit V, Esther Duflo, and Rachel Glennerster**, “Putting a band-aid on a corpse: incentives for nurses in the Indian public health care system,” *Journal of the European economic association*, 2008, 6 (2-3), 487–500.
- Banfield, Edward C.**, *The Moral Basis of a Backward Society*, New York: Free Press, 1958.
- Basu, Kaushik**, “One kind of power,” *Oxford Economic Papers*, 1986, 38 (2), 259–282.
- Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier**, “Face-to-face communication in organizations,” *The Review of Economic Studies*, 2021, 88 (2), 574–609.
- Becker, Gary S and George J Stigler**, “Law enforcement, malfeasance, and compensation of enforcers,” *The Journal of Legal Studies*, 1974, 3 (1), 1–18.
- Bedasso, Biniam**, “Institutional and personnel capabilities for fiscal decentralisation in Benin,” Technical Report, Collaborative Africa Budget Reform Initiative (CABRI) 2021.
- Behanzin, Précieux Christian, Manuela Hounvi, Casilda Dossou-Yovo, Steven Mariano, and Stella Towanou**, “Social capital and the practices of governance in the municipality of Sèmè-Podji in Benin,” *Environmental and Water Sciences, Public Health and Territorial Intelligence Journal*, 2018, 2 (4), 31–40.
- Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu**, “The Glittering Prizes: Career Incentives and Bureaucrat Performance,” *The Review of Economic Studies*, 2020, 87 (2), 626–655.
- Besley, Timothy**, *Principled agents?: The political economy of good government*, Oxford University Press, USA, 2006.
- and **John McLaren**, “Taxes and bribery: the role of wage incentives,” *The economic journal*, 1993, 103 (416), 119–141.
- and **Torsten Persson**, “Why do developing countries tax so little?,” *Journal of Economic Perspectives*, 2014, 28 (4), 99–120.
- Bidwell, Kelly, Katherine Casey, and Rachel Glennerster**, “Debates: Voting and expenditure responses to political communication,” *Journal of Political Economy*, 2020, 128 (8), 2880–2924.
- Bloch, Francis, Garance Genicot, and Debraj Ray**, “Reciprocity in groups and the limits to social capital,” *American Economic Review*, 2007, 97 (2), 65–69.

- Bloom, Nicholas and John Van Reenen**, “Measuring and explaining management practices across firms and countries,” *The Quarterly Journal of Economics*, 2007, 122 (4), 1351–1408.
- , **Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts**, “Does management matter? Evidence from India,” *The Quarterly Journal of Economics*, 2013, 128 (1), 1–51.
- Bó, Ernesto Dal, Frederico Finan, and Martín A Rossi**, “Strengthening state capabilities: The role of financial incentives in the call to public service,” *The Quarterly Journal of Economics*, 2013, 128 (3), 1169–1218.
- Burgess, Robin, Matthew Hansen, Benjamin A Olken, Peter Potapov, and Stefanie Sieber**, “The political economy of deforestation in the tropics,” *The Quarterly journal of economics*, 2012, 127 (4), 1707–1754.
- Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee**, “The political economy of public sector absence,” *Journal of Public Economics*, 2023, 218, 104787.
- Canen, Nathan, Rafael Ch, and Leonard Wantchekon**, “Political uncertainty and the forms of state capture,” *Journal of Development Economics*, 2023, 160, 102972.
- Casey, Katherine, Rachel Glennerster, and Edward Miguel**, “Reshaping institutions: Evidence on aid impacts using a preanalysis plan,” *The Quarterly Journal of Economics*, 2012, 127 (4), 1755–1812.
- Ch, Rafael, Mathias Hounkpe, and Leonard Wantchekon**, “campaign Finance and State Capture,” Technical Report, Economic Development and Institutions 2019.
- Coffman, Katherine Baldiga**, “Evidence on self-stereotyping and the contribution of ideas,” *The Quarterly Journal of Economics*, 2014, 129 (4), 1625–1660.
- Collier, Paul and Pedro C Vicente**, “Votes and violence: Evidence from a field experiment in Nigeria,” *The economic journal*, 2014, 124 (574), F327–F355.
- Corbacho, Ana, Daniel W Gingerich, Virginia Oliveros, and Mauricio Ruiz-Vega**, “Corruption as a self-fulfilling prophecy: evidence from a survey experiment in Costa Rica,” *American Journal of Political Science*, 2016, 60 (4), 1077–1092.
- Dasgupta, Aditya and Devesh Kapur**, “The political economy of bureaucratic overload: Evidence from rural development officials in India,” *American Political Science Review*, 2020, 114 (4), 1316–1334.
- Dasgupta, Partha**, “Economics of social capital,” *Economic Record*, 2005, 81, S2–S21.

- de Janvry, Alain, Guojun He, Elisabeth Sadoulet, Shaoda Wang, and Qiong Zhang**, “Subjective Performance Evaluation, Influence Activities, and Bureaucratic Work Behavior: Evidence from China,” *American Economic Review*, 2023, 113 (3), 766–799.
- Debnath, Sisir, Mrithyunjayan Nilayamgode, and Sheetal Sekhri**, “Information Bypass: Using Low-cost technological innovations to curb leakages in welfare programs,” *Journal of Development Economics*, 2023, p. 103137.
- Decentralisation and Communal Development Support Programme*
- Decentralisation and Communal Development Support Programme*, Technical Report, KfW Development Bank 2021.
- Dhaliwal, Iqbal and Rema Hanna**, “The devil is in the details: The successes and limitations of bureaucratic reform in India,” *Journal of Development Economics*, 2017, 124, 1–21.
- Dipoppa, Gemma and Saad Gulzar**, “Administrative Incentives Impact Crop-Residue Burning and Health in South Asia,” *Working Paper*, 2023.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan**, “Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India,” *The Quarterly Journal of Economics*, 2013, 128 (4), 1499–1545.
- , —, —, and —, “The value of regulatory discretion: Estimates from environmental inspections in India,” *Econometrica*, 2018, 86 (6), 2123–2160.
- Ferraz, Claudio, Frederico Finan, and Diana B Moreira**, “Corrupting learning: Evidence from missing federal education funds in Brazil,” *Journal of Public Economics*, 2012, 96 (9-10), 712–726.
- Finan, Frederico, Benjamin A Olken, and Rohini Pande**, “The personnel economics of the developing state,” in “*Handbook of Economic Field Experiments*,” Vol. 2, Elsevier, 2017, pp. 467–514.
- Fujiwara, Thomas and Leonard Wantchekon**, “Can informed public deliberation overcome clientelism? Experimental evidence from Benin,” *American Economic Journal: Applied Economics*, 2013, 5 (4), 241–255.
- Gambetta, Diego**, “*The Mafia: A ruinous rationality*,” 1993.
- Glaeser, Edward L, David I Laibson, Jose A Scheinkman, and Christine L Soutter**, “Measuring trust,” *The Quarterly Journal of Economics*, 2000, 115 (3), 811–846.
- Goold, Susan Dorr, Michael A Neblo, SY Kim, R De Vries, Gene Rowe, and Peter Muhlberger**, “What is good public deliberation,” *Hastings Center Report*, 2012, 42 (2), 24–26.



- Gulzar, Saad and Benjamin J Pasquale**, *“Politicians, bureaucrats, and development: Evidence from India,” American Political Science Review*, 2017, 111 (1), 162–183.
- Hadfield, Gillian K and Stephen Macedo**, *“Rational reasonableness: Toward a positive theory of public reason,” The Law & Ethics of Human Rights*, 2012, 6 (1), 7–46.
- Hibbing, John R and Elizabeth Theiss-Morse**, *Stealth democracy: Americans’ beliefs about how government should work*, Cambridge University Press, 2002.
- Imbert, Bruno, Mouhamadou Sy, Bacari Koné, Louis d’Humières, Pierre Roumegas, Nicoletta Feruglio, Mathias Gogohounga, and Angelo Donou**, *“Benin Technical Assistance Report - Public Investment Management Assessment,” Technical Report, International Monetary Fund* 2018.
- Iyer, Lakshmi and Anandi Mani**, *“Traveling agents: political change and bureaucratic turnover in India,” Review of Economics and Statistics*, 2012, 94 (3), 723–739.
- Jia, Ruixue, Masayuki Kudamatsu, and David Seim**, *“Political selection in China: The complementary roles of connections and performance,” Journal of the European Economic Association*, 2015, 13 (4), 631–668.
- Khan, Adnan Q, Asim I Khwaja, and Benjamin A Olken**, *“Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors,” The Quarterly Journal of Economics*, 2016, 131 (1), 219–271.
- , **Asim Ijaz Khwaja, and Benjamin A Olken**, *“Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings,” Technical Report, National Bureau of Economic Research* 2018.
- Khan, Muhammad Yasir**, *“Mission motivation and public sector performance: Experimental evidence from pakistan,” Unpublished manuscript*, 2023.
- Kosec, Katrina and Leonard Wantchekon**, *“Can information improve rural governance and service delivery?,” World Development*, 2020, 125, 104376.
- Kremer, Michael, Nazmul Chaudhury, F Halsey Rogers, Karthik Muralidharan, and Jeffrey Hammer**, *“Teacher absence in India: A snapshot,” Journal of the European Economic Association*, 2005, 3 (2-3), 658–667.
- Leaver, Clare**, *“Bureaucratic minimal squawk behavior: Theory and evidence from regulatory agencies,” American Economic Review*, 2009, 99 (3), 572–607.
- Lichand, Guilherme, Marcos FM Lopes, and Marcelo C Medeiros**, *“Is corruption good for your health,” scholar.harvard.edu/glichand/publications/job-market-paper. Acessado em*, 2016, 25 (04), 2016.

- López-Moctezuma, Gabriel, Leonard Wantchekon, Daniel Rubenson, Thomas Fujiwara, and Cecilia Pe Lero**, “Policy deliberation and voter persuasion: Experimental evidence from an election in the Philippines,” *American Journal of Political Science*, 2022, 66 (1), 59–74.
- Mendelberg, Tali**, “The deliberative citizen: Theory and evidence,” *Political decision making, deliberation and participation*, 2002, 6 (1), 151–193.
- **and John Oleske**, “Race and public deliberation,” *Political Communication*, 2000, 17 (2), 169–191.
- Menzel, Andreas**, “Knowledge exchange and productivity spill-overs in Bangladeshi garment factories,” *Journal of Economic Behavior & Organization*, 2021, 185, 721–746.
- Milgrom, Paul and John Roberts**, “An economic approach to influence activities in organizations,” *American Journal of Sociology*, 1988, 94, S154–S179.
- Moreira, Diana and Santiago Pérez**, “Civil service exams and organizational performance: Evidence from the pendleton act,” *Technical Report, National Bureau of Economic Research* 2021.
- Muralidharan, Karthik and Paul Niehaus**, “Experimentation at scale,” *Journal of Economic Perspectives*, 2017, 31 (4), 103–124.
- Niehaus, Paul and Sandip Sukhtankar**, “Corruption dynamics: The golden goose effect,” *American Economic Journal: Economic Policy*, 2013, 5 (4), 230–69.
- **and —**, “The marginal rate of corruption in public programs: Evidence from India,” *Journal of Public Economics*, 2013, 104, 52–64.
- Olken, Benjamin A**, “Monitoring corruption: evidence from a field experiment in Indonesia,” *Journal of political Economy*, 2007, 115 (2), 200–249.
- **and Rohini Pande**, “Corruption in developing countries,” *Annu. Rev. Econ.*, 2012, 4 (1), 479–509.
- Papay, John P, Eric S Taylor, John H Tyler, and Mary E Laski**, “Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data,” *American Economic Journal: Economic Policy*, 2020, 12 (1), 359–388.
- Parkinson, John and Jane Mansbridge**, *Deliberative systems: Deliberative democracy at the large scale*, Cambridge University Press, 2012.
- Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny**, “The quality of government,” *Journal of Law, Economics, and Organization*, 1999, 15 (1), 222–279.

- Portes, Alejandro and Patricia Landolt**, *"The downside of social capital,"* 1996.
- Raffler, Pia J**, *"Does political oversight of the bureaucracy increase accountability? Field experimental evidence from a dominant party regime,"* *American Political Science Review*, 2022, 116 (4), 1443–1459.
- Rasul, Imran and Daniel Rogger**, *"Management of bureaucrats and public service delivery: Evidence from the nigerian civil service,"* *The Economic Journal*, 2018, 128 (608), 413–446.
- Riaño, Juan Felipe**, *"Bureaucratic nepotism,"* Available at SSRN 3995589, 2021.
- Riedl, Rachel Beatty and J Tyler Dickovick**, *"Party systems and decentralization in Africa,"* *Studies in Comparative International Development*, 2014, 49, 321–342.
- Roberts, Nancy**, *"Public deliberation: An alternative approach to crafting policy and setting direction,"* *Public Administration Review*, 1997, pp. 124–132.
- Sandvik, Jason, Richard Saouma, Nathan Seegert, and Christopher T Stanton**, *"Should workplace programs be voluntary or mandatory? evidence from a field experiment on mentorship,"* Technical Report, National Bureau of Economic Research 2021.
- Tella, Rafael Di and Ernesto Schargrodsky**, *"The role of wages and auditing during a crackdown on corruption in the city of Buenos Aires,"* *Journal of Law and Economics*, 2003, 46 (1), 269–292.
- Thompson, Dennis F**, *"Deliberative democratic theory and empirical political science,"* *Annual Review of Political Science*, 2008, 11, 497–520.
- Tirole, Jean**, *"Hierarchies and bureaucracies: On the role of collusion in organizations,"* *The Journal of Law, Economics, & Organization*, 1986, 2 (2), 181–214.
- Wantchekon, Leonard and Jenny Guardado**, *"Deliberation and Ethical Voting Behavior: Evidence from a Campaign Experiment in Benin,"* *American Political Science Review*, Forthcoming.
- Young, Alwyn**, *"Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results,"* *The Quarterly Journal of Economics*, 2019, 134 (2), 557–598.

# Tables

Table 1: Description of meetings held across ten treated municipalities

	Meeting 1		Meeting 2		Meeting 3		Overall	
	N	Mean	N	Mean	N	Mean	N	Mean
<b>Meetings' Characteristics</b>								
Total Participants	10	31.30	10	31.60	10	34.10	30	32.33
Bureaucrat Participants	10	21.80	10	23.80	10	22.90	30	22.83
General Secretary Attended	10	0.90	10	1.00	10	1.00	30	0.97
Duration (in minutes)	10	100.00	10	91.80	10	94.10	30	95.30
Duration of Presentation (in minutes)	10	32.70	10	36.40	10	38.20	30	35.77
Duration of Discussion (in minutes)	10	48.90	10	37.50	10	44.20	30	43.53
Duration of Presentation (in percentage)	10	0.34	10	0.42	10	0.42	30	0.39
Duration of Discussion (in percentage)	10	0.47	10	0.41	10	0.48	30	0.46
Number of Contributions	8	7.75	10	4.30	10	3.40	28	4.96
Number of Critics	5	7.20	10	5.60	6	9.67	21	7.14
Number of Questions	8	4.00	9	2.89	6	4.83	23	3.78
<b>Perception of Enumerators</b>								
Meeting is Interactive	10	1.00	9	1.00	10	1.00	29	1.00
Meeting is Informative	10	1.00	9	1.00	10	0.60	29	0.86
Meeting Atmosphere is Good	10	1.00	9	1.00	10	1.00	29	1.00
Meeting Atmosphere is Tense	10	0.00	9	0.00	10	0.00	29	0.00

Note: The data is compiled from the meeting notes taken by enumerators. Meeting 1 discussed the audit reports of the municipalities, while meetings 2 and 3 described results from a survey of bureaucrats and citizens, respectively. General Secretary attended is a dummy variable that turns on 1 if a General Secretary of the municipality attended the meeting. Criticism, Contribution, and Question are the number of times these items were recorded during each meeting across municipalities. The enumerators also reported their perceptions regarding the atmosphere of the meetings. These are dummies for whether the meetings were interactive or informative and whether the atmosphere was tense or good.

Table 2: Did Collective Deliberation Affect Bureaucrats' Performance?

Dependent Variables	Pre-Period Control Mean (1)	DID Estimate (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
Bureaucrats' Audit-Based Performance	77.835	6.8286**	2.9555	0.032	0.027	112
<b>Sub-Components of Bureaucrats' Audit-Based Performance</b>						
Record Keeping	78.493	9.3016	7.7031	0.242	0.209	112
Transparency of Public Procurement	78.297	5.7866	4.6223	0.226	0.182	112
Overall Management	83.500	14.4186	13.4617	0.298	0.235	72

Note: The unit of observation is municipality-year. The data is from the Audit Reports of 2014-2019. Bureaucrats' Audit-Based Performance is a score constructed from 3 sub-components: (i) Record Keeping (ii) Transparency of Public Procurement (iii) Overall Management. Record Keeping is in turn constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records. Transparency of Public Procurement is in turn constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders. Overall Management is in turn constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 1000 repetitions in Column (5). All regressions include municipality and year fixed effects. Standard Errors are clustered at the municipality level. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 3: Did Collective Deliberation Affect Citizens' Outcomes?

Dependent Variables	Pre-Period Control Mean (1)	DID Estimate (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
<b>Panel A : Citizen Perception</b>						
Trust in Council	0.788	-0.0695	0.0904	0.453	0.492	454
Do Councilors Listen	0.731	-0.0505	0.0898	0.581	0.578	451
Approve Performance of Councilors	0.596	0.0801	0.0859	0.364	0.393	456
Corrupt among Councilors	0.922	-0.0080	0.0557	0.888	0.895	447
Corrupt among Govt Officials	0.960	0.0005	0.0407	0.990	0.988	444
Officials Go Unpunished	0.892	-0.0196	0.0472	0.683	0.698	449
<b>Panel B : Citizen Engagement</b>						
Contacted Local Councilor	0.288	-0.0098	0.1327	0.942	0.942	454
Contacted Govt Official	0.058	-0.0808	0.0673	0.246	0.322	456
Complained	0.106	0.0149	0.0614	0.811	0.814	456
Willing to Complain	0.865	-0.1014	0.0811	0.228	0.283	456
<i>Citizen Engagement Index</i>	-0.042	-0.2359	0.2265	0.312	0.369	456
<b>Panel C : Present Living Conditions</b>						
Are Good	0.279	0.0516	0.1127	0.653	0.685	456
Better Compared to Others	0.233	0.0961	0.0705	0.191	0.228	455
<i>Living Conditions Index</i>	-0.062	0.0792	0.2629	0.767	0.780	456
<b>Panel D : Difficulty in Access to</b>						
Identity Document Services	0.652	-0.0675	0.1595	0.678	0.663	206
Household Services	0.857	0.0159	0.1434	0.913	0.910	88
<i>Difficulty in Access to Services Index</i>	0.201	0.0146	0.2879	0.960	0.958	237
<b>Panel E : Paid Bribes for</b>						
Identity Document Services	0.304	-0.0141	0.1250	0.912	0.907	206
Household Services	0.143	-0.4373	0.2576	0.110	0.141	88
<i>Likelihood of Bribes for Services Index</i>	0.121	-0.1116	0.2806	0.696	0.750	237

Note: The unit of observation is a citizen-year. The analysis is based on Afrobarometer data for the years 2014 and 2017. Panel A consists of indicators of whether the citizens trust the councilors, feel that they listen to them, approve of their performance, believe that there is corruption in council and government officials and whether the latter go unpunished. Panel B consists of indicators of whether the citizens contacted local councilor, govt officials, complained and willing to do so. Panel C consists of indicators of whether the citizens feels that their present living conditions are good and are better compared to others. Panel D & E consist of indicators of whether citizens faced difficulties in accessing (i) identity documents like a birth certificate, driver's license, passport, voter's card or permit from government and (ii) household services (which includes water, electricity and sanitation and had to pay bribes for it respectively). We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 1000 repetitions in Column (5). Standard Errors are clustered at the municipality level. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



Table 4: Mechanism: Effects of the meetings on information within the bureaucracy

Dependent Variables	Control Mean	Treatment Effect	Standard Error	P-Value	Randomization Inference P-Value	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Are the following engaged in corruption</b>						
Council	-0.096	0.2076	0.1435	0.164	0.220	455
Other Bureaucrats	-0.109	0.1731	0.1882	0.369	0.260	510
Vendors	-0.216	0.3881***	0.1121	0.003	0.000	374
<b>Panel B: Enforcement of Rules for</b>						
Public Procurement	0.115	-0.2279*	0.1315	0.099	0.160	454
Hiring Staff	0.126	-0.2126	0.1567	0.191	0.260	411

Note: The unit of observation is a bureaucrat. Panel A consists of standardized variables measuring the extent to which bureaucrats believe that councilors, other bureaucrats(own and other departments) and vendors are engaged in corruption. Panel B consists of standardized variables measuring extent to which bureaucrats believe enforcement of rules is followed for public procurement and hiring staff. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). Standard Errors are clustered at the municipality level. Control variables include average of : (i) Bureaucrats' age, and indicators for (ii) College Degree, (iii) Yearly Accountability Sessions and (iv) Existence of Equipment Acquisition Procedure at the municipality level in the baseline survey. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

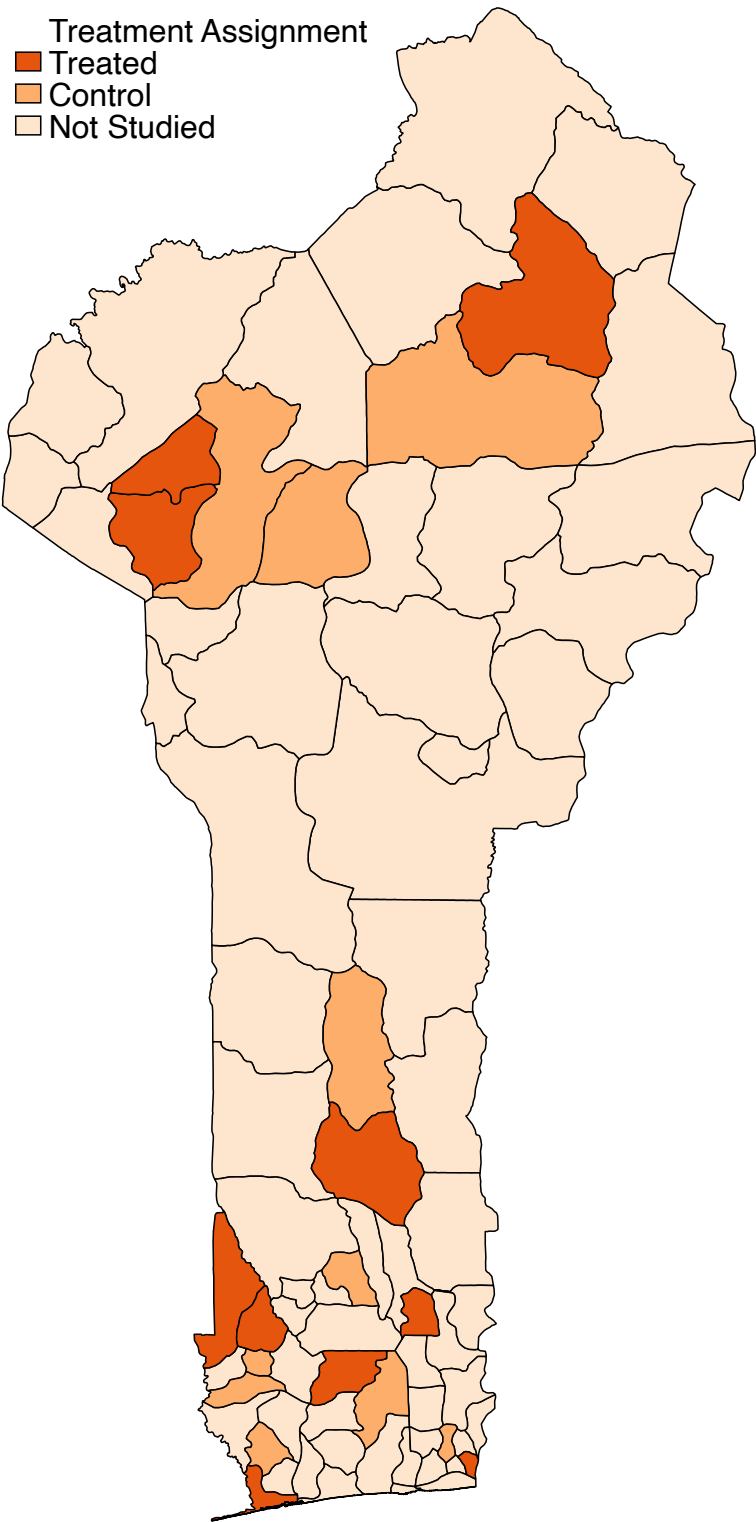
Table 5: Mechanism: Effects of the meetings on trust and cooperation in the bureaucracy

Dependent Variables	Control Mean	Treatment Effect	Standard Error	P-Value	Randomization Inference P-Value	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Bureaucrats' Trust in</b>						
Council	0.182	-0.3855**	0.1795	0.045	0.120	581
Other Bureaucrats	0.137	-0.0802	0.1520	0.604	0.680	595
<b>Panel B: Issues Meetings Solve</b>						
Conflicts between Colleagues	0.839	-0.1202*	0.0586	0.054	0.100	527
Administrative Dysfunctions	0.947	-0.0752**	0.0285	0.016	0.100	534
Poor Working Conditions	0.947	-0.0945**	0.0345	0.013	0.000	536
Lack of Recreational Activities	0.773	-0.1149	0.1181	0.343	0.160	524
<i>Working Environment Index</i>	0.176	-0.4151**	0.1745	0.028	0.000	541
<b>Panel C: Cooperation between Bureaucrats</b>						
Cooperation Index	0.164	-0.0785	0.2493	0.756	0.800	593
<b>Panel D: Do Colleagues</b>						
Meetings are Important	0.855	-0.0640	0.0622	0.316	0.280	500
Meetings Should Continue	0.964	-0.0089	0.0219	0.689	0.660	500

Note: The unit of observation is a bureaucrat. Panel A consists of standardized variables measuring the extent to which bureaucrats trust the councilors and other bureaucrats (own office, own department and other departments). Panel B consists of indicators of whether the bureaucrats feel such meetings would help in solving (i) conflicts between colleagues (ii) administrative dysfunctions (absence or non-compliance with texts, dependence on some agents, absences, lateness, lack of communication, etc.) (iii) poor working conditions and (iv) lack of cohesion (lack of recreational activities, lack of confidence, professional development, etc.). Working Environment Index is a standardized index of these indicators. Panel C consists of Cooperation Index. It is a standardized index of how often the bureaucrats engage in the following with their colleagues : (i) discuss work together (ii) organize recreational activities (iii) organize professional activities like workshops (iv) offer to perform indisposed colleague's work (v) solve conflict with each other and (vi) refer to hierarchy to resolve it. In Panel D, we have two variables: Meetings are important and meetings should continue are indicators of whether bureaucrats believe that the meetings were important and similar ones should continue in the future respectively. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). Standard Errors are clustered at the municipality level. Control variables include average of : (i) Bureaucrats' age, and indicators for (ii) College Degree, (iii) Yearly Accountability Sessions and (iv) Existence of Equipment Acquisition Procedure at the municipality level in the baseline survey. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

# Figures

Figure 1: Spatial Distribution of Treatment Assignment



Notes. Treated and control municipalities are in red and orange, respectively. White-colored ones are non-studied municipalities.

Figure 2: Proportion of Meeting Attendees By Organizations

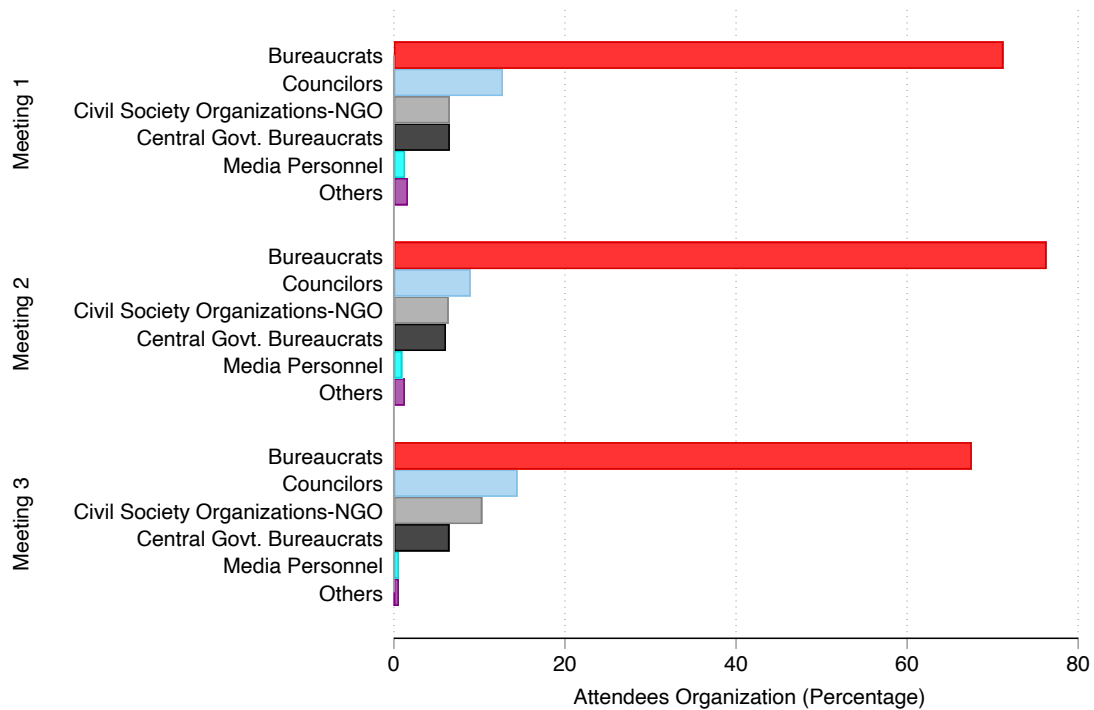


Figure 3: Average duration of presentation

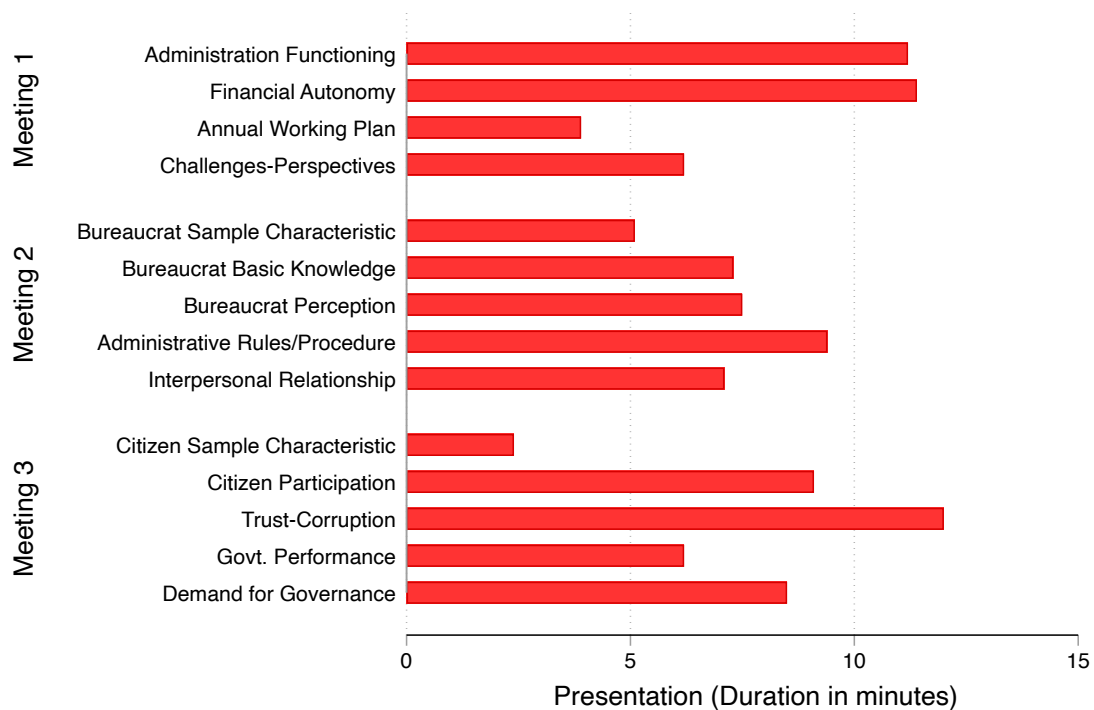


Figure 4: Average duration of discussion after the presentation

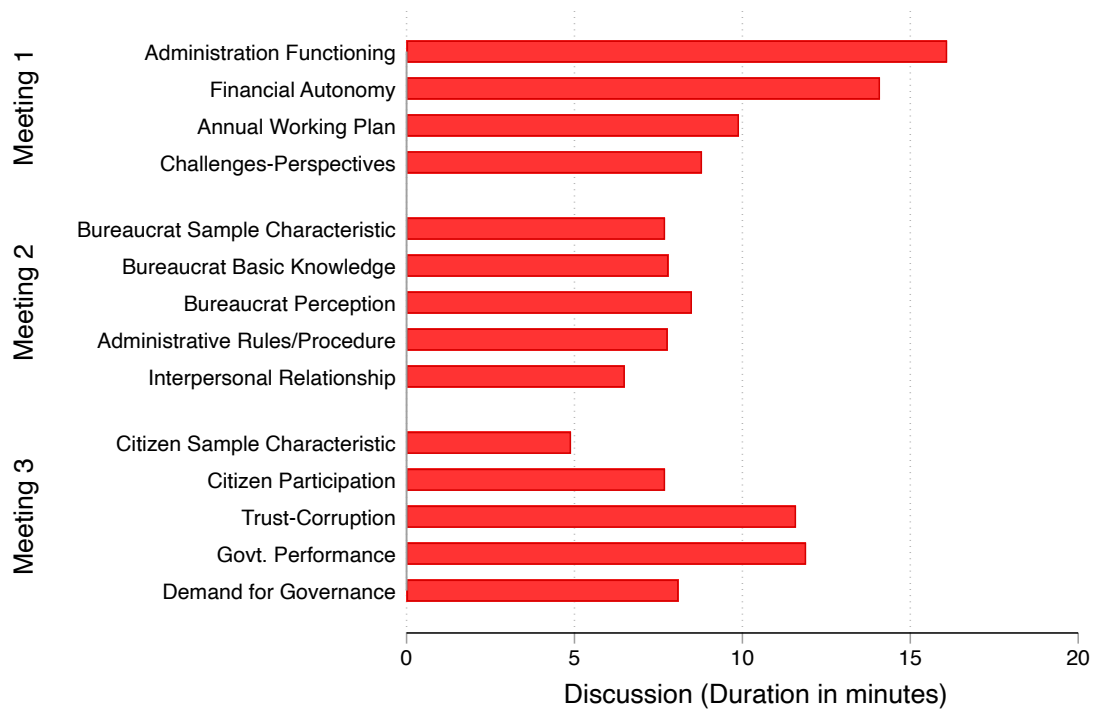
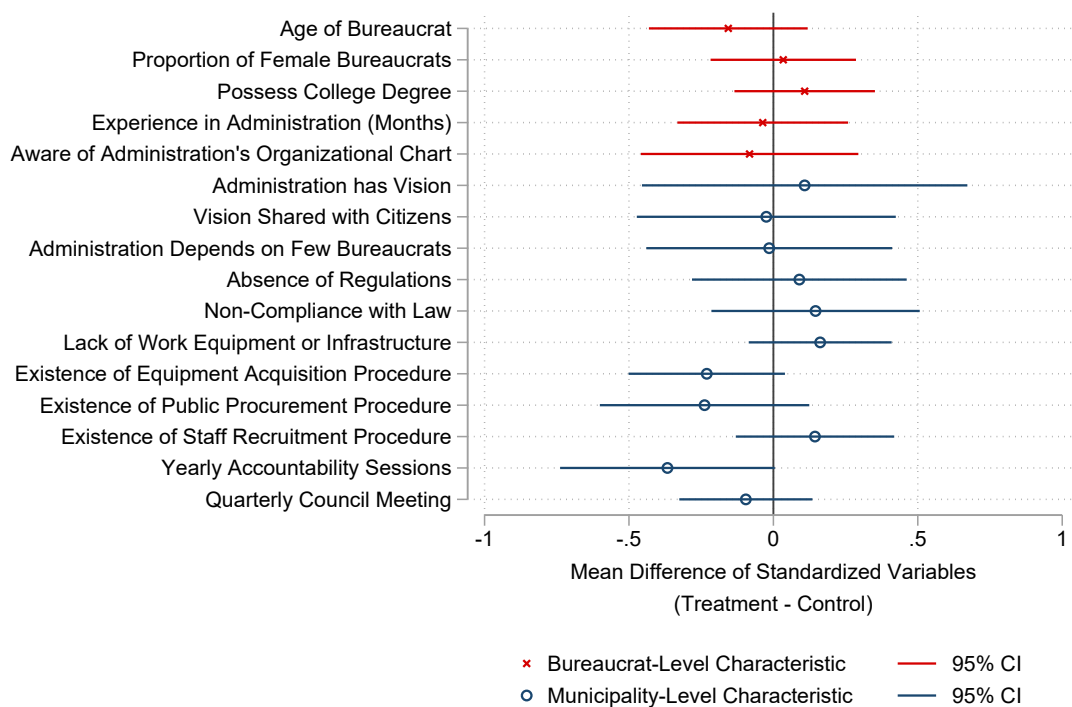
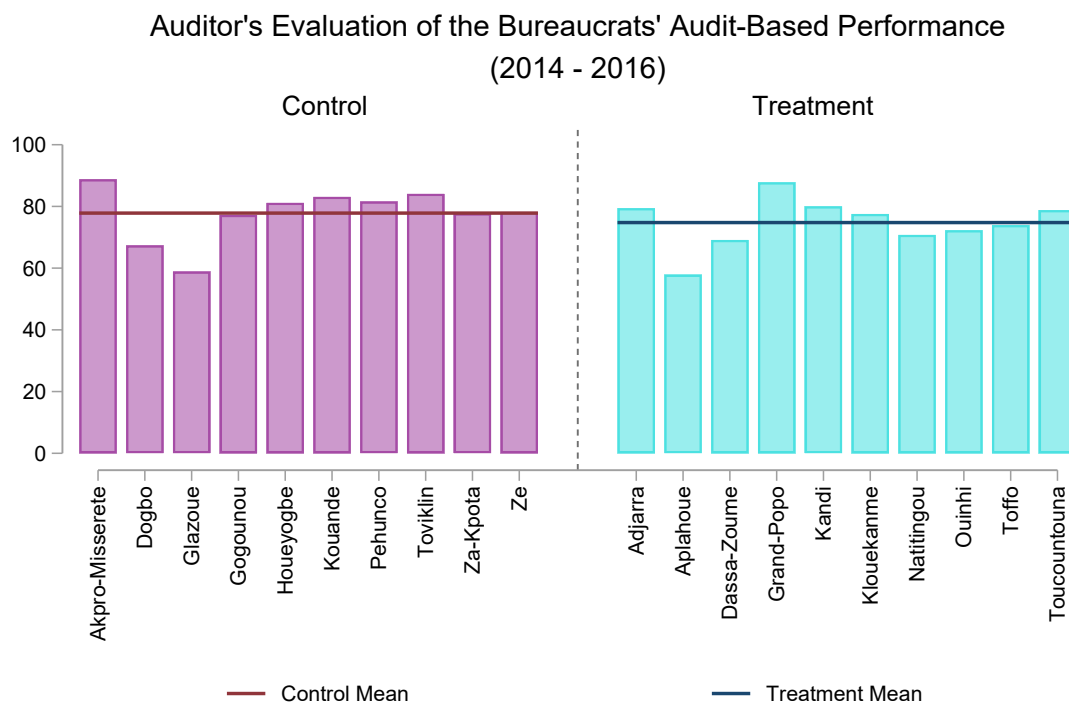


Figure 5: Balance test



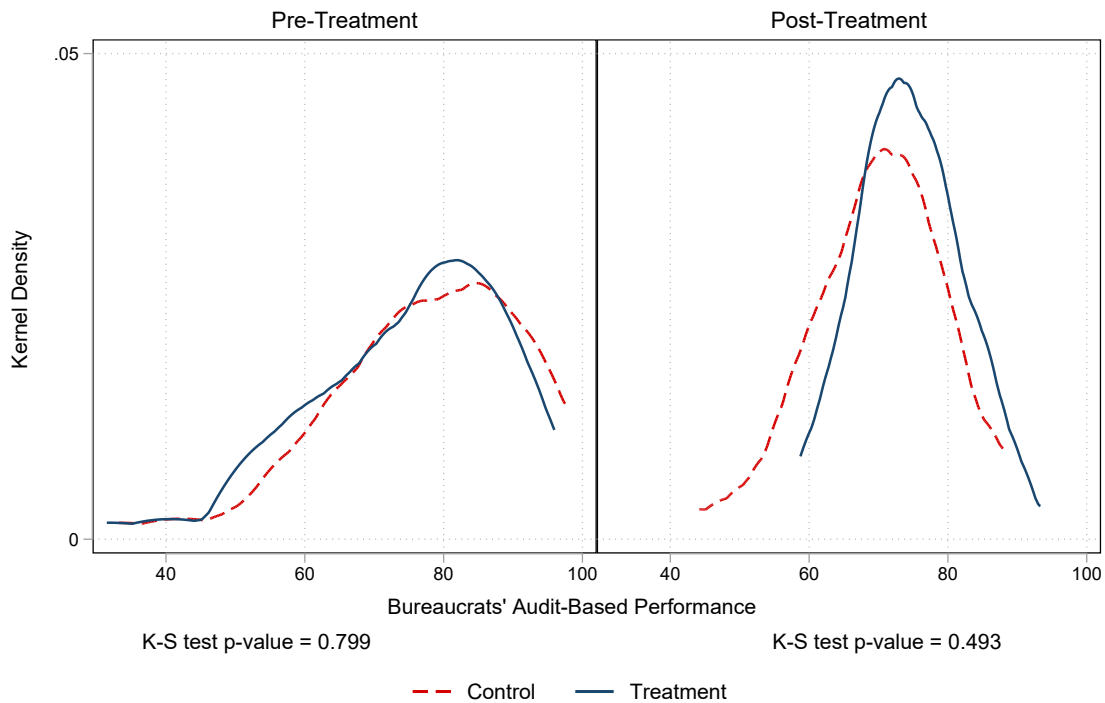
Notes. The variables have been standardized and their mean difference between treatment and control group is plotted along with 95% CIs. Yearly Accountability Sessions and Existence of Equipment Acquisition Procedure are statistically different (at 10% level of significance) between treatment and control municipalities.

Figure 6: Baseline Average of Bureaucrats' Audit-Based Performance



Notes. Bureaucrats' Audit-Based Performance is a score constructed from three sub-components: (i) Record Keeping (ii) Transparency of Public Procurement (iii) Overall Management. Record Keeping is constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records. Transparency of Public Procurement is constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders. Overall Management is constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents.

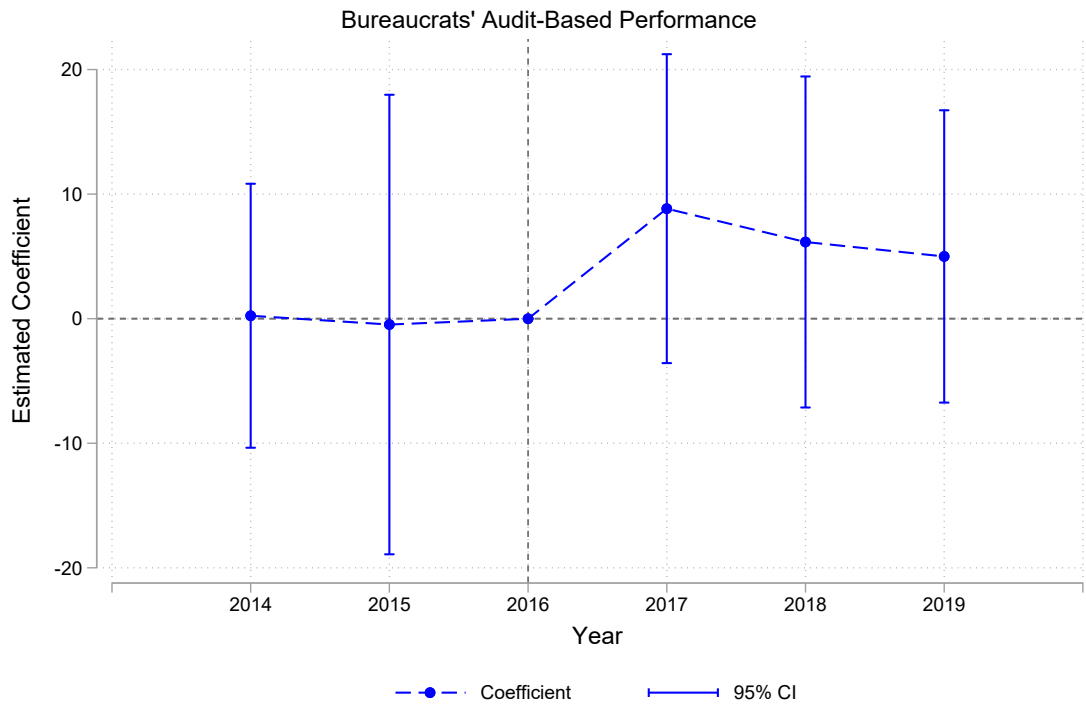
Figure 7: Bureaucrats' Audit-Based Performance



Notes. Pre-Treatment refers to years 2014-2016, while post treatment refers to years 2017-2019. Bureaucrats' Audit-Based Performance is a score built from the audit reports of FADeC. It score is constructed from three sub-components: (i) Record Keeping (ii) Transparency of Public Procurement (iii) Overall Management. Record Keeping is constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records. Transparency of Public Procurement is constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders. Overall Management is constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents.



Figure 8: Event Study Plot of Bureaucrats' Audit-Based Performance



Notes. Bureaucrats' Audit-Based Performance is a score built from the audit reports of FADeC. Its score is constructed from three sub-components: (i) Record Keeping (ii) Transparency of Public Procurement (iii) Overall Management. Record Keeping is constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records. Transparency of Public Procurement is constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders. Overall Management is constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents.

# Appendices

## A Appendix Tables

Table A.1: Relation between previous year's audit-based performance and current year's FaDeC investment transfers

	Dependent Variable: FaDeC Investment Transfers (t)			
	(1)	(2)	(3)	(4)
Audit-Based Performance (t-1)	1615995** (735236) [0.0406]	1564939* (754951.5) [0.0520]	1546852* (763319.3) [0.0570]	2029669** (905340.7) [0.0371]
Control for Population Size	No	Yes	Yes	Yes
Control for Poverty Rate	No	No	Yes	Yes
Municipality Fixed Effects	No	No	No	Yes
Observations	97	97	97	97

Note: The unit of observation is municipality-year. The data is from the Audit Reports of 2014-2019. The dependent variable is the amount of FaDeC investment transfers in year t. *Audit – based performance (t – 1)* is last year's performance score of the municipality as evaluated by the auditors. These are measured in percentages. Column (1) has no controls. Column (2) controls for population size of the municipalities in year t, while Column (3) controls for their poverty rates. Finally, Column (4) includes municipality fixed effects. The standard errors are reported in parentheses and are clustered at the municipality level. The p-values are reported in square brackets. Significance levels are denoted as: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

Table A.2: IV results mechanism: Effects of the meetings on information within the bureaucracy

Dependent Variables	Control Mean (1)	Local Average Treatment Effect (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
<b>Panel A: Are the following engaged in corruption</b>						
Council	-0.096	0.3890	0.2492	0.118	0.900	455
Other Bureaucrats	-0.109	0.3307	0.3592	0.357	0.780	510
Vendors	-0.216	0.7209***	0.2164	0.001	0.580	374
<b>Panel B: Enforcement of Rules for</b>						
Public Procurement	0.115	-0.4035*	0.2374	0.089	0.800	454
Hiring Staff	0.126	-0.3501	0.2473	0.157	0.860	411

Note: The unit of observation is a bureaucrat. Panel A consists of standardized variables measuring the extent to which bureaucrats believe that councilors, other bureaucrats(own and other departments) and vendors are engaged in corruption. Panel B consists of standardized variables measuring extent to which bureaucrats believe enforcement of rules is followed for public procurement and hiring staff. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). The Treatment indicator is used as an instrument for actual participation in the meeting for the bureaucrat. Standard Errors are clustered at the municipality level. Control variables include average of : (i) Bureaucrats' age, and indicators for (ii) College Degree, (iii) Yearly Accountability Sessions and (iv) Existence of Equipment Acquisition Procedure at the municipality level in the baseline survey. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A.3: IV results mechanism: Effects of the meetings on trust and cooperation in the bureaucracy

Dependent Variables	Control Mean (1)	Local Average Treatment Effect (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
<b>Panel A: Bureaucrats' Trust in</b>						
Council	0.182	-0.7392**	0.3372	0.028	0.780	581
Other Bureaucrats	0.137	-0.1565	0.2884	0.587	0.920	595
<b>Panel B: Issues Meetings Solve</b>						
Conflicts between Colleagues	0.839	-0.2192*	0.1130	0.052	0.740	527
Administrative Dysfunctions	0.947	-0.1380**	0.0569	0.015	0.760	534
Poor Working Conditions	0.947	-0.1767***	0.0602	0.003	0.580	536
Lack of Recreational Activities	0.773	-0.2106	0.2097	0.315	0.820	524
<i>Working Environment Index</i>	0.176	-0.7815**	0.3287	0.017	0.580	541
<b>Panel C: Cooperation between Bureaucrats</b>						
Cooperation Index	0.164	-0.1524	0.4641	0.743	0.880	593
<b>Panel D: Do Colleagues</b>						
Meetings are Important	0.855	-0.1087	0.1034	0.293	0.880	500
Meetings Should Continue	0.964	-0.0158	0.0376	0.675	0.920	500

Note: The unit of observation is a bureaucrat. Panel A consists of standardized variables measuring the extent to which bureaucrats trust the councilors and other bureaucrats (own office, own department and other departments). Panel B consists of indicators of whether the bureaucrats feel such meetings would help in solving (i) conflicts between colleagues (ii) administrative dysfunctions (absence or non-compliance with texts, dependence on some agents, absences, lateness, lack of communication, etc.) (iii) poor working conditions and (iv) lack of cohesion (lack of recreational activities, lack of confidence, professional development, etc.). Working Environment Index is a standardized index of these indicators. Panel C consists of Cooperation Index. It is a standardized index of how often the bureaucrats engage in the following with their colleagues : (i) discuss work together (ii) organize recreational activities (iii) organize professional activities like workshops (iv) offer to perform indisposed colleague's work (v) solve conflict with each other and (vi) refer to hierarchy to resolve it. In Panel D, we have two variables: Meetings are important and meetings should continue are indicators of whether bureaucrats believe that the meetings were important and similar ones should continue in the future respectively. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). The Treatment indicator is used as an instrument for actual participation in the meeting for the bureaucrat. Standard Errors are clustered at the municipality level. Control variables include average of : (i) Bureaucrats' age, and indicators for (ii) College Degree, (iii) Yearly Accountability Sessions and (iv) Existence of Equipment Acquisition Procedure at the municipality level in the baseline survey. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A.4: Afrobarometer Data - Round 6 (2014), 7 (2017) and 8 (2021)

Dependent Variables	Pre-Period Control Mean (1)	DID Estimate (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
<b>Panel A : Citizen Perception</b>						
Trust in Council	0.788	-0.0345	0.0818	0.678	0.705	693
Do Councilors Listen	0.731	-0.0283	0.0789	0.724	0.733	691
Approve Performance of Councilors	0.596	-0.0023	0.0815	0.978	0.976	696
Corrupt among Councilors	0.922	-0.0190	0.0461	0.685	0.710	686
Corrupt among Govt Officials	0.960	0.0290	0.0371	0.445	0.465	683
Officials Go Unpunished	0.892	0.0408	0.0546	0.465	0.474	685
<b>Panel B : Citizen Engagement</b>						
Contacted Local Councilor	0.288	0.0040	0.0942	0.967	0.969	694
Contacted Govt Official	0.058	-0.0808	0.0673	0.246	0.329	456
Complained	0.106	0.0149	0.0614	0.811	0.810	456
Willing to Complain	0.865	-0.1014	0.0811	0.228	0.286	456
<i>Citizen Engagement Index</i>	-0.042	-0.1641	0.1711	0.351	0.436	696
<b>Panel C : Present Living Conditions</b>						
Are Good	0.279	0.0147	0.1170	0.901	0.908	696
Better Compared to Others	0.233	0.0961	0.0705	0.191	0.228	455
<i>Living Conditions Index</i>	-0.062	0.0141	0.2805	0.961	0.960	696
<b>Panel D : Difficulty in Access to</b>						
Identity Document Services	0.652	-0.0517	0.1648	0.757	0.759	269
Household Services	0.857	0.0159	0.1434	0.913	0.912	88
<i>Difficulty in Access to Services Index</i>	0.201	0.0279	0.2654	0.918	0.919	300
<b>Panel E : Paid Bribes for</b>						
Identity Document Services	0.304	0.0103	0.0995	0.919	0.911	269
Household Services	0.143	-0.4373	0.2576	0.110	0.133	88
<i>Likelihood of Bribes for Services Index</i>	0.121	-0.0494	0.2002	0.808	0.840	300

Notes: The unit of observation is a citizen-year. The analysis is based on Afrobarometer data for the years 2014, 2017 & 2021. Panel A consists of indicators of whether the citizens trust the councilors, feel that they listen to them, approve of their performance, believe that there is corruption in council and government officials and whether the latter go unpunished. Panel B consists of indicators of whether the citizens believes contacted local councilor, govt officials, complained and willing to do so. Panel C consists of indicators of whether the citizens feels that their present living conditions are good and are better compared to others. Panel D & E consist of indicators of whether citizens faced difficulties in accessing (i) identity documents like a birth certificate, driver's license, passport, voter's card or permit from government and (ii) household services (which includes water, electricity and sanitation and had to pay bribes for it respectively). We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). Standard Errors are clustered at the municipality level. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

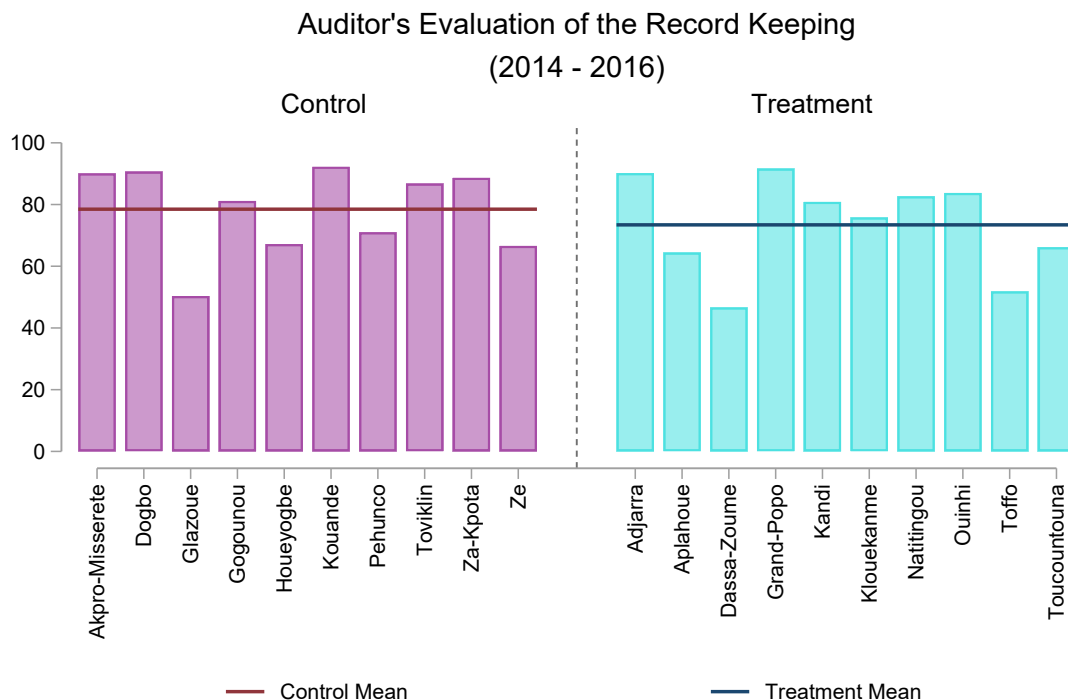
Table A.5: Auditor's Evaluation of Local Government Council

Dependent Variables	Pre-Period Control Mean (1)	DID Estimate (2)	Standard Error (3)	P-Value (4)	Randomization Inference P-Value (5)	Obs. (6)
Overall Performance	67.818	3.2039	2.5163	0.218	0.179	120
Council's Performance	55.774	-2.0356	5.3483	0.708	0.671	112
<b>Sub-Components of Council's Performance</b>						
Local Government Finance	44.231	-3.0958	6.0608	0.615	0.576	112
Rule Following by the Council	96.167	-0.0217	7.2104	0.998	0.995	112

Notes: The unit of observation is municipality-year. Overall Performance is a score assigned by CoNaFiL to a municipality annually based on several criteria including both Bureaucrats' and Council's Performance. Council's Performance is a score constructed from 2 sub-components : (i) Local Government Finance and (ii) Rule Following by the Council. Local Government Finance is in turn constructed from scores given based on increase in own revenue over the year, self-financing of the investment, self-financing capacity, expenditure on maintenance of movable and immovable assets compared to operating expenditure, infrastructure maintenance expenditure, evolution of own revenue, and level of consumption of available resources. Rule Following by the Council is in turn constructed from scores given based on adoption of the initial budget within the legal deadlines, budget vote date, date of voting of the administrative account, administrative account voting date, and availability of the administrative account within the legal deadlines. We report  $p$ -value in Column (4) and  $p$ -value from Randomization Inference with 10,000 repetitions in Column (5). Municipality and year fixed effects are present in the regression specifications. Standard Errors are clustered at the municipality level. Significance levels are denoted as: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

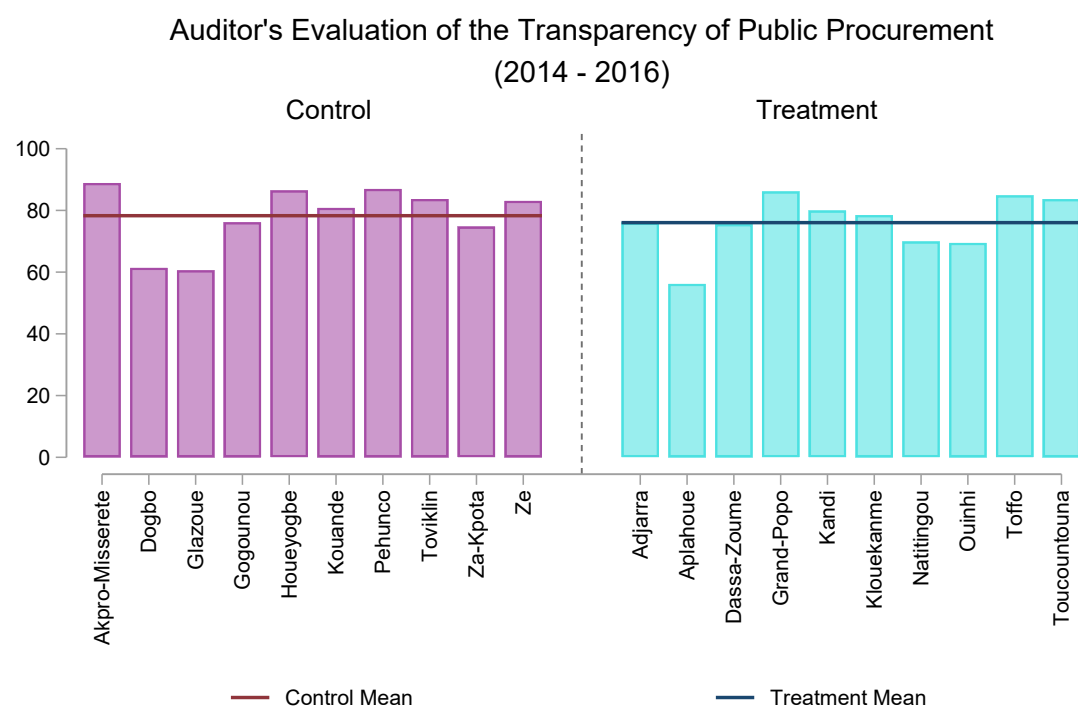
## B Appendix Figures

Figure B.1: Baseline Average of Auditor's Evaluation of Record Keeping by Bureaucrats



Notes. Record Keeping is constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records.

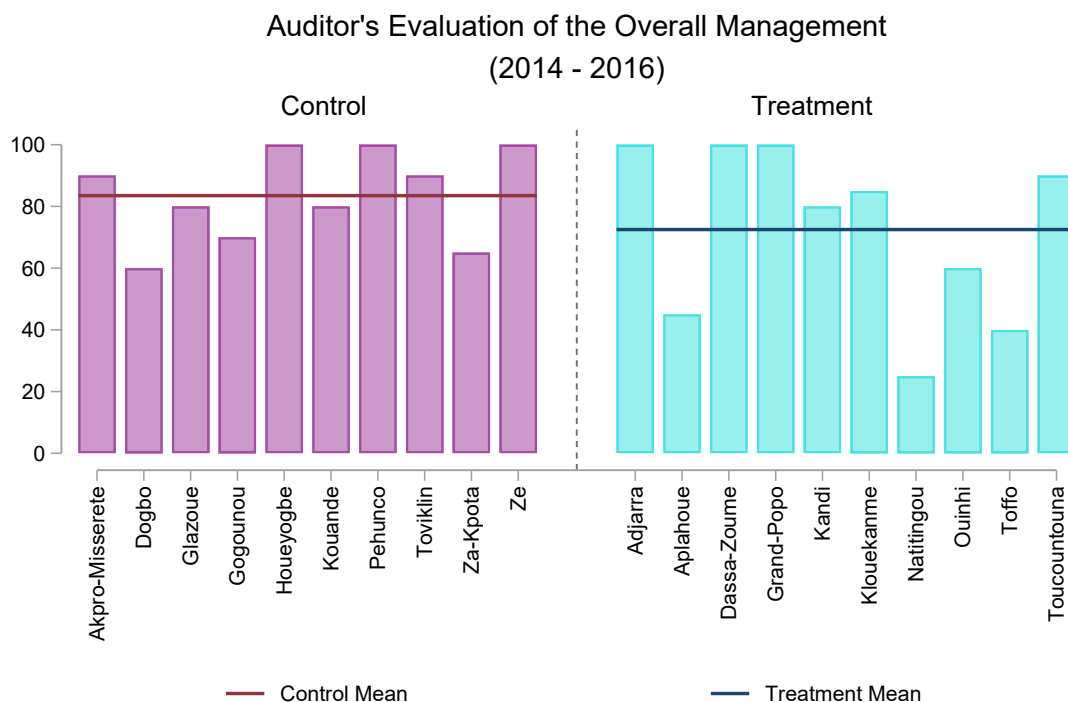
Figure B.2: Baseline Average of Auditor’s Evaluation of Transparency of Public Procurement by Bureaucrats



Notes. Transparency of Public Procurement is constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders.

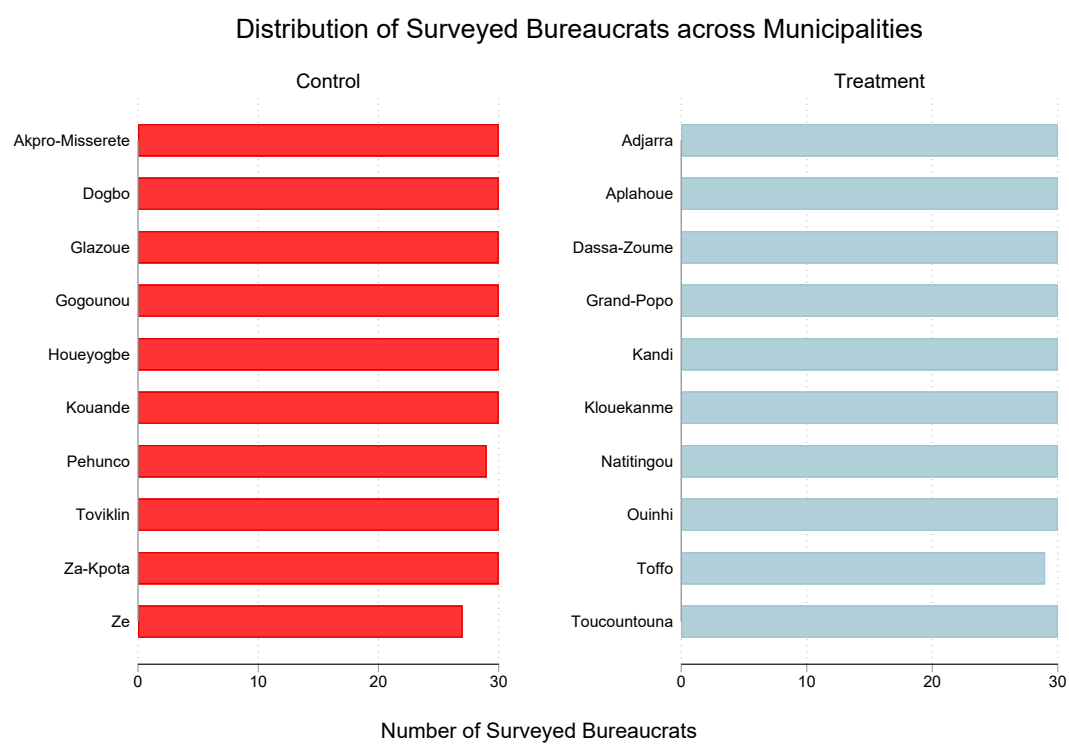


Figure B.3: Baseline Average of Auditor's Evaluation of Overall Management by Bureaucrats



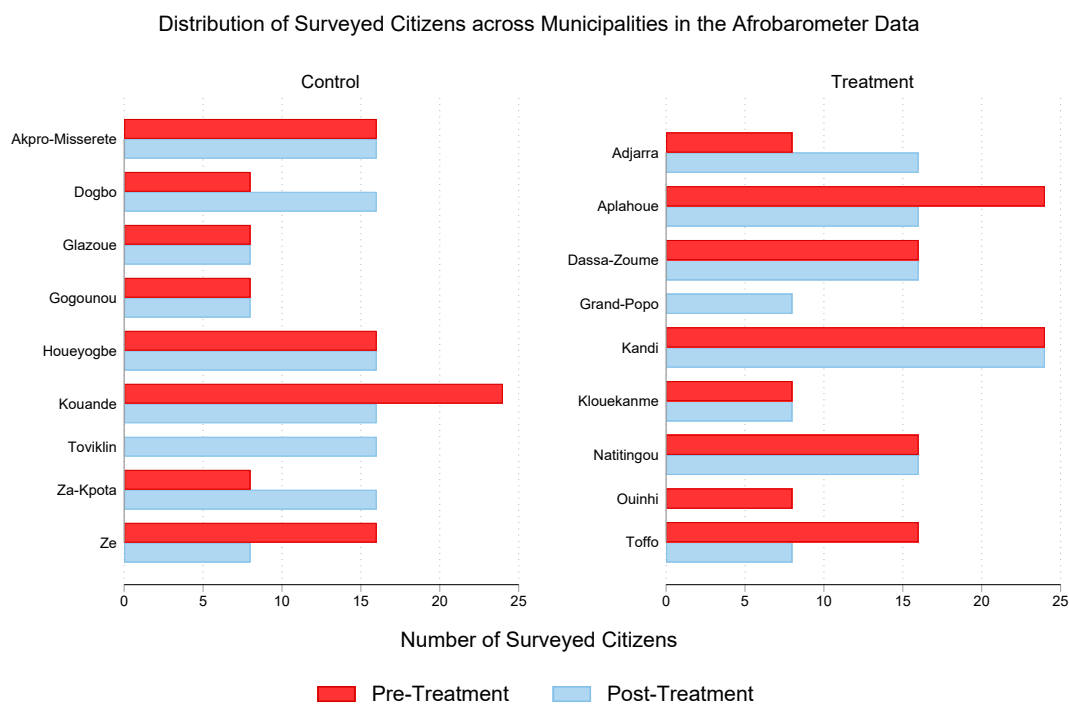
Notes. Overall Management is constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents.

Figure B.4: Number of bureaucrats surveyed across treatment and control municipalities in December 2016



Notes. For the endline survey conducted in December 2016, 30 bureaucrats were randomly selected within each of the 20 municipalities. However, there were some exceptions: 27 in Ze and 29 each in Pehunco & Toffo, bringing the total number of observations to 595.

Figure B.5: Number of surveyed citizens across treatment and control municipalities in the Afrobarometer Data



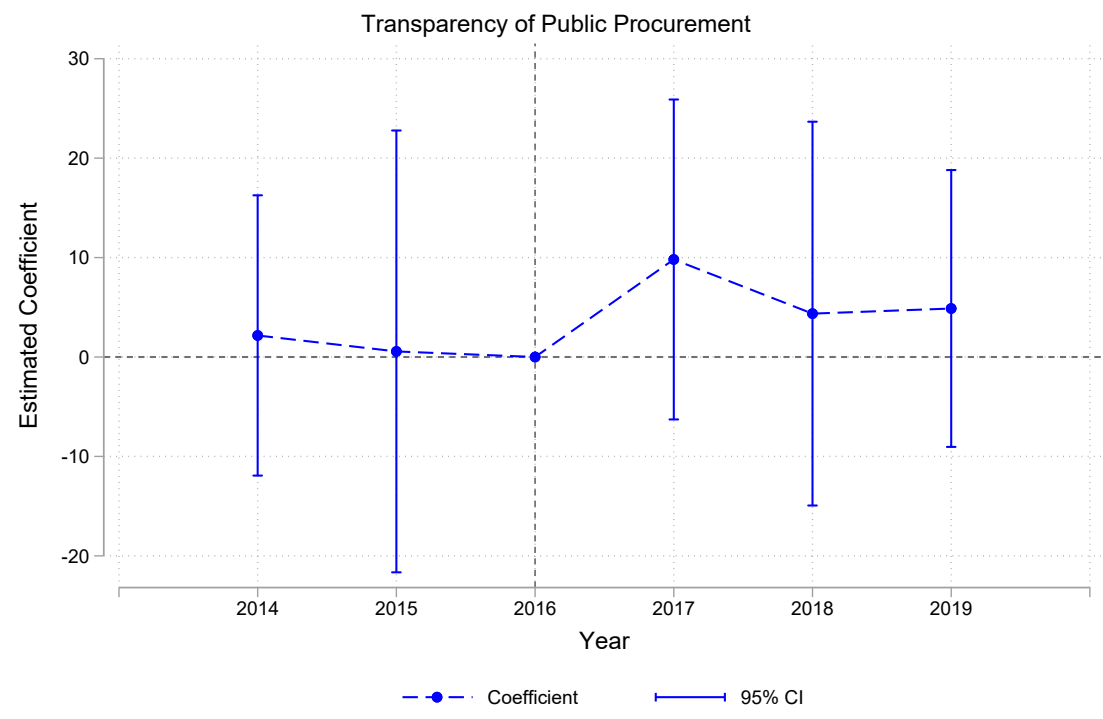
Notes. The sampling frame is a region, constituency and urban-rural location (strata). Within each stratum, clustered random sampling takes place: (i) randomly selecting start points within each PSU, (ii) followed by randomly selecting households from each start point and (iii) finally randomly selecting respondents within a household. There are 8 households surveyed per PSU in each round. We use round 6 (2014) and round 7 (2017) as pre and post treatment periods respectively for our analysis. Note that there are no observations sampled for Pehunco and Toucountouna for either of the rounds. No observations were sampled for round 6 for Toviklin and Grand-Popo.

Figure B.6: Event Study Plot of Auditor's Evaluation of Record Keeping by the Bureaucrats.



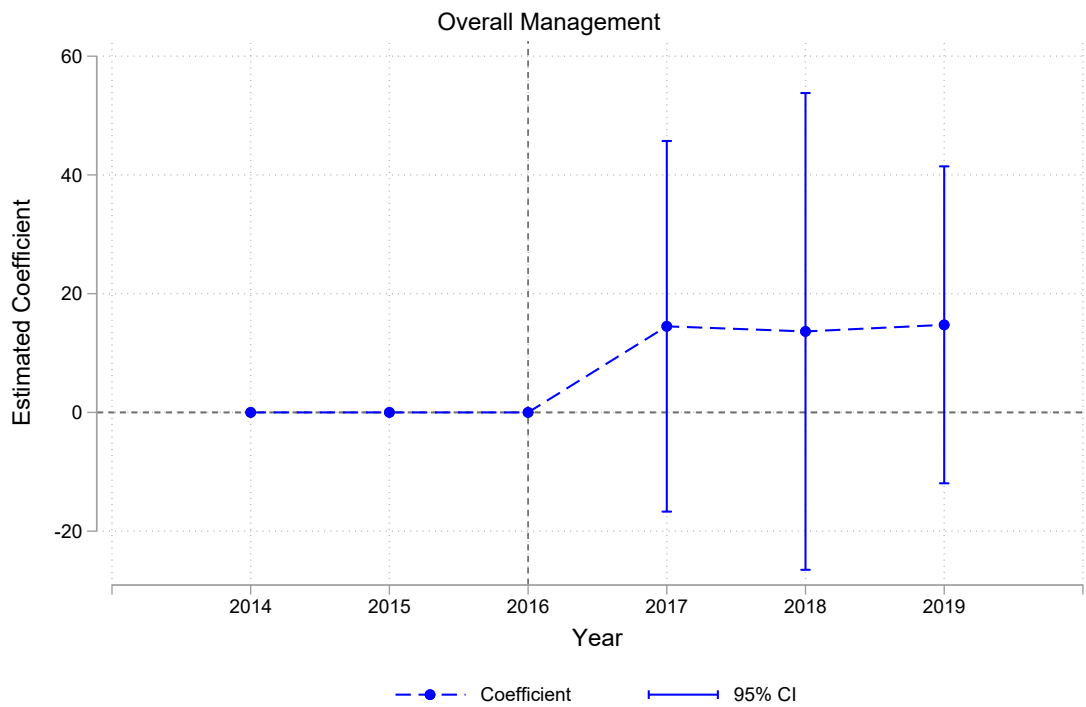
Notes. Record Keeping is constructed from scores given based on availability of quarterly and annual execution statements of transferred resources, functionality of an archiving and documentation device, updating of physical records and existence of stock records, updating of physical registers by the authorizing officer, and proper maintenance of accounting records.

Figure B.7: Event Study Plot of Auditor’s Evaluation of Transparency of Public Procurement by the Bureaucrats.



Notes. Transparency of Public Procurement is constructed from scores given based on legal existence and functionality of the main PM bodies and functionality of the S/PRMP, compliance with public procurement procedures, functionality of public procurement award and control bodies, availability of the FADeC enforcement point, quality of the administrative account and traceability of transfers, and execution of public orders.

Figure B.8: Event Study Plot of Auditor's Evaluation of Overall Management of the Bureaucracy.



Notes. Overall Management is constructed from scores given based on effectiveness of the role of coordinator of local services by the Secretary General, and management of civil status documents. No data exists prior to 2016 for this score.

## C Appendix: For Online Publication

### C.1 Details of the criteria used by auditors to audit the local government

Table C.1: Details of the criteria used by auditors to audit the local government's record keeping

Year	Record Keeping Criterion	Maximum Score	Total Score
2014	Proper maintenance of accounting records	5	5
2015	Proper maintenance of accounting records	5	5
2016	Features of the archiving and documentation system*	5	17 <sup>‡</sup>
	Availability of quarterly and annual execution statements (financial and physical) of transferred resources	7 <sup>†</sup>	
	Updation of physical registers by the authorizing officer	5	
2017	Features of the archiving and documentation system	5	16
	Availability of quarterly and annual execution statements (financial and physical) of transferred resources	6	
	Updation of physical registers by the authorizing officer	5	
2018	Features of the archiving and documentation system	5	16
	Availability of quarterly and annual execution statements (financial and physical) of transferred resources	6	
	Updation of physical records and existence of stock records	5	
2019	Features of the archiving and documentation system	5	16
	Availability of quarterly and annual execution statements (financial and physical) of transferred resources	6	
	Updation of physical records and existence of stock records	5	

\*It involves existence of documentation room accessible to the public and equipped with storage furniture

<sup>†</sup>Exceptions: 10 for Kouande, Pehunco and Ze

<sup>‡</sup>Exceptions: 20 for Kouande, Pehunco and Ze

Table C.2: Details of the criteria used by auditors to audit the local government's transparency of public procurement

Year	Transparency of Public Procurement Criterion	Maximum Score	Total Score
2014	Existence & functioning of the main bodies as well as the Secretariat responsible for public procurement <sup>§</sup>	9	15
	Availability of details regarding the spending of the FaDeC funds for the auditors	6	
2015	Existence & functioning of the main bodies as well as the Secretariat responsible for public procurement	9	15
	Availability of details regarding the spending of the FaDeC funds for the auditors	6	
2016	Functioning of the bodies responsible for public procurement and control of public contracts <sup>¶</sup>	10	30
	Compliance with public procurement procedures	20	
2017	Functioning of the bodies responsible for public procurement and control of public contracts	6	35
	Compliance with public procurement procedures	10	
	Execution of public orders	15	
	Quality of the administrative account and traceability of transfers	4	
2018	Functioning of the bodies responsible for public procurement and control of public contracts	6	35
	Compliance with public procurement procedures	10	
	Execution of public orders	15	
	Quality of the administrative account and traceability of transfers	4	
2019	Functionality of public procurement award and control bodies	5	34
	Compliance with public procurement procedures	10	
	Execution of public orders	15	
	Quality of the administrative account and traceability of transfers	4	

<sup>§</sup>3 points each for the functioning of each body responsible for public procurement.

<sup>¶</sup>Points are assigned on the number and exhaustiveness of the reports published by the bodies responsible for public procurement.



Table C.3: Details of the criteria used by auditors to audit the local government's overall management

Year	Overall Management Criterion	Maximum Score	Total Score
2014	N/A	N/A	N/A
2015	N/A	N/A	N/A
2016	Effectiveness of the role of coordinator of local services by the Secretary General	5	5
2017	Effectiveness of the role of coordinator of local services by the Secretary General	5	5
2018	Effectiveness of the role of coordinator of local services by the Secretary General	5	5
2019	Effectiveness of the role of coordinator of local services by the Secretary General	5	10
	Management of Civil Status documents*	5	

---

\*These are vital records of life events kept under governmental authority, including birth certificates, marriage licenses, divorce certificates and death certificates.

Table C.4: Summary of the criterion used by auditors across years

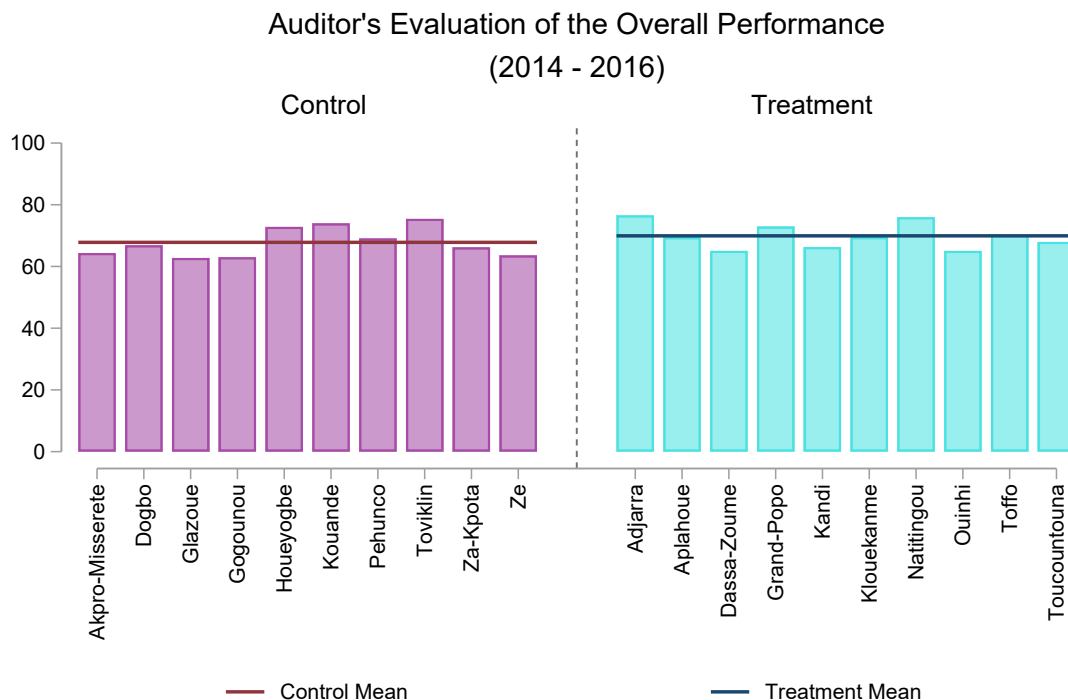
Indices / Sub-Indices		Year					
		2014	2015	2016	2017	2018	2019
Record Keeping	Number of Criteria	1	1	3	3	3	3
	Maximum Score	5	5	17 <sup>†</sup>	16	16	16
Transparency of Public Procurement	Number of Criteria	2	2	2	4	4	4
	Maximum Score	15	15	30	35	35	34
Overall Management	Number of Criteria	0	0	1	1	1	2
	Maximum Score	N/A	N/A	5	5	5	10
Bureaucrats' Audit-Based Performance	Number of Criteria	3	3	6	8	8	9
	Maximum Score	20	20	52 <sup>**</sup>	56	56	60

<sup>†</sup>Exceptions: 20 for Kouande, Pehunco and Ze

<sup>\*\*</sup>Exceptions: 55 for Kouande, Pehunco and Ze

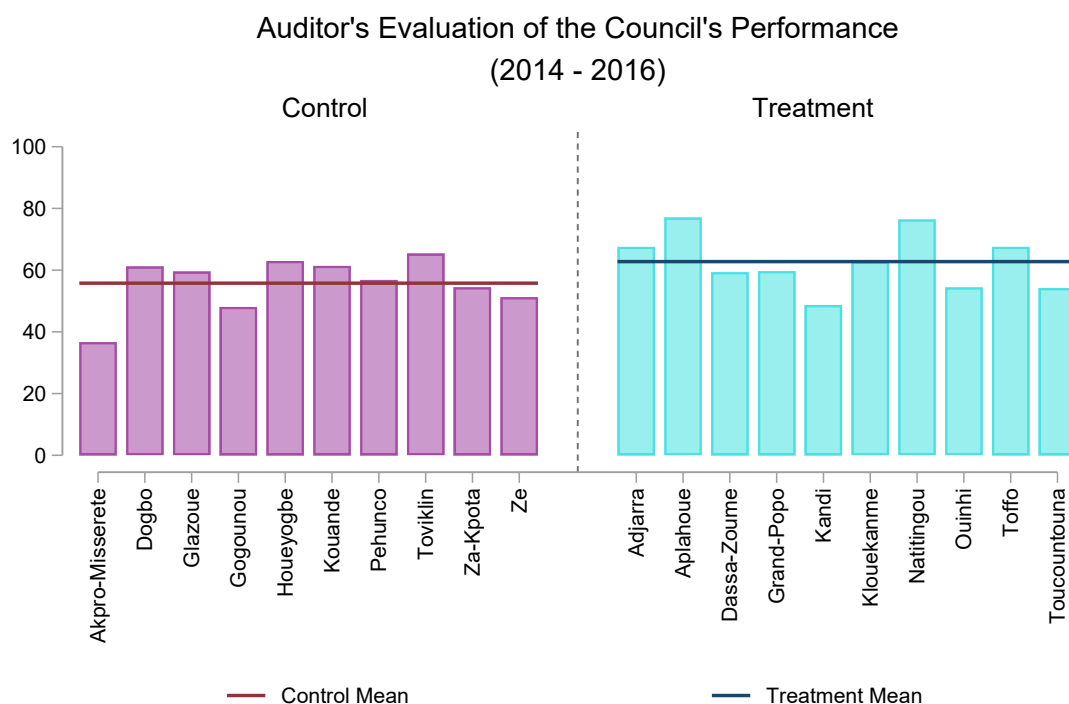
## C.2 Auditors' baseline evaluation of municipal governments based on other criteria

Figure C.1: Baseline Average of Auditor's Evaluation of Overall Performance of the Local Government



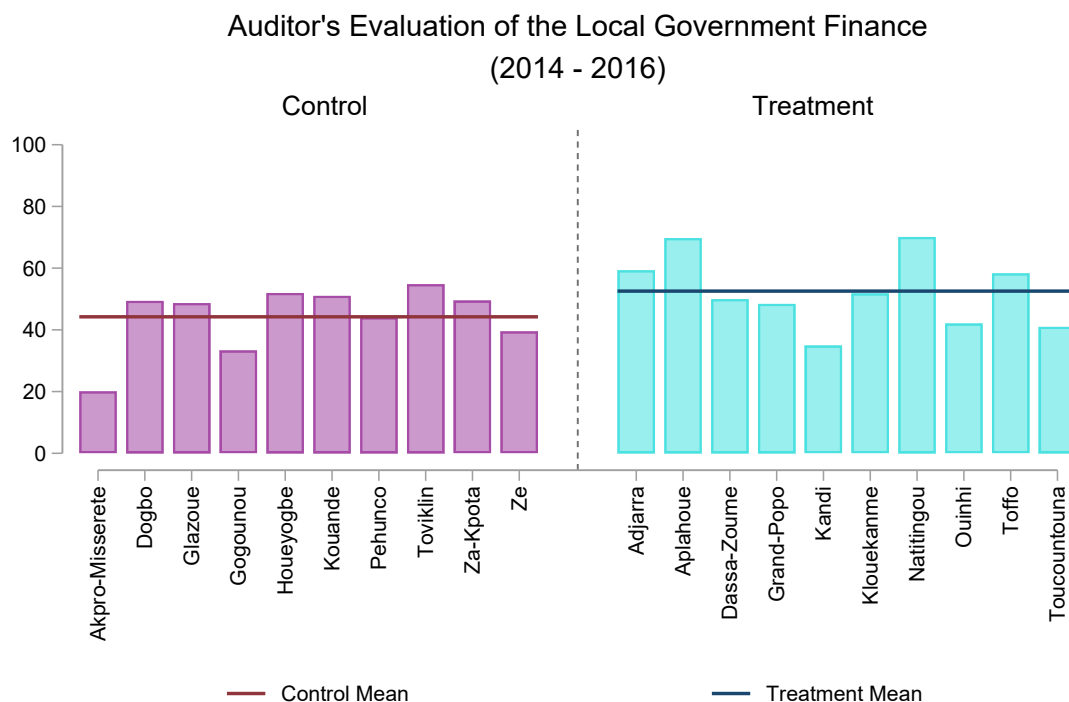
Notes. Overall Performance is a score assigned by FADeC to a municipality annually based on several criteria including both Bureaucrats' and Council's Performance.

Figure C.2: Baseline Average of Auditor's Evaluation of the Council's Performance.



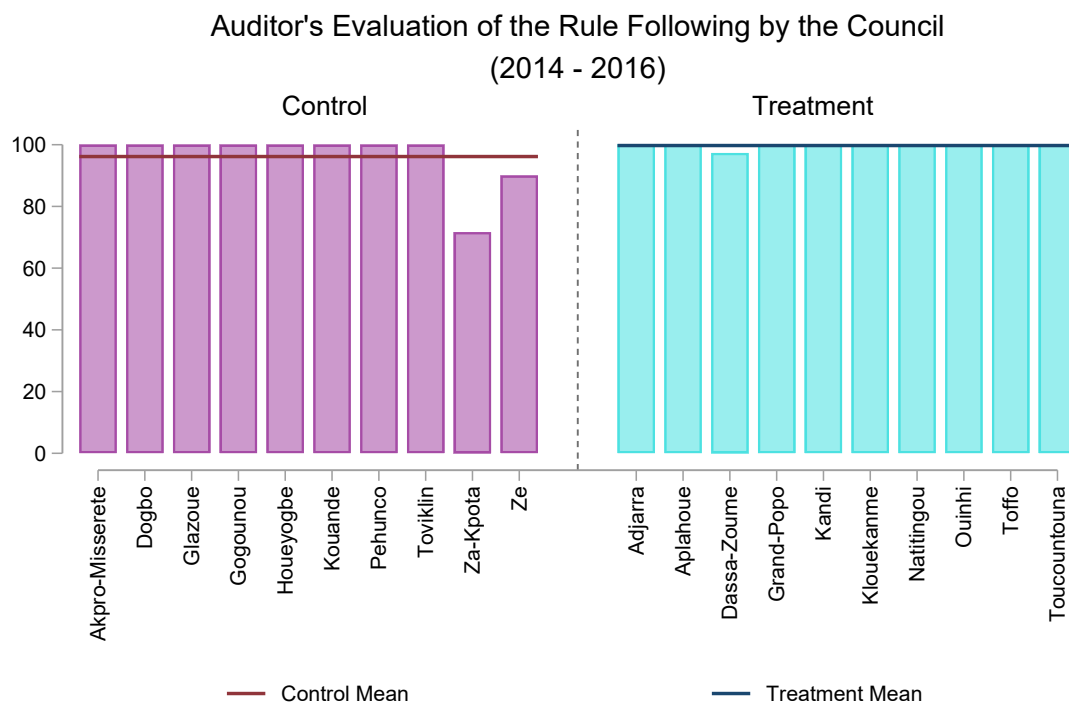
Notes. Council's Performance is a score constructed from 2 sub-components : (i) Local Government Finance and (ii) Rule Following by the Council. Local Government Finance is constructed from scores given based on increase in own revenue over the year, self-financing of the investment, self-financing capacity, expenditure on maintenance of movable and immovable assets compared to operating expenditure, infrastructure maintenance expenditure, evolution of own revenue, and level of consumption of available resources. Rule Following by the Council is in turn constructed from scores given based on adoption of the initial budget within the legal deadlines, budget vote date, date of voting of the administrative account, administrative account voting date, and availability of the administrative account within the legal deadlines.

Figure C.3: Baseline Average of Auditor's Evaluation of the Local Government's Finances



Notes. Local Government Finance is constructed from scores given based on increase in own revenue over the year, self-financing of the investment, self-financing capacity, expenditure on maintenance of movable and immovable assets compared to operating expenditure, infrastructure maintenance expenditure, evolution of own revenue, and level of consumption of available resources.

Figure C.4: Baseline Average of Auditor's Evaluation of Rule Following by Council



Notes. Rule Following by the Council is in turn constructed from scores given based on adoption of the initial budget within the legal deadlines, budget vote date, date of voting of the administrative account, administrative account voting date, and availability of the administrative account within the legal deadlines.