The Organizational Dynamics of Administrative Data Production

Tara Slough*

July 23, 2026

Abstract

High-quality administrative data is essential for governance. I argue that collaboration among bureaucrats is an important and under-recognized determinant of the quality of administrative data, and that institutional structure, in turn, shapes incentives for collaboration. Partnering with a government agency in Colombia, I find that local bureaucrats who collaborate with colleagues in their respective mayoralties produce better-quality measures of municipal characteristics—measures that the national government then uses to target social programs. Original quantitative and qualitative data suggest these improvements are attributable to a combination of selection into collaboration and improved information from collaboration. I further show that career bureaucrats and those in more hierarchical mayoralties are more likely to collaborate. These findings suggest that institutional arrangements that foster communication and collaboration among bureaucrats can improve the quality of administrative data.

Word count: 9,916 words

*Associate Professor, New York University, tara.slough@nyu.edu. Thanks to Camilo Rodríguez and Juan Bedoya for expert research assistance and to Dorothy Kronick, Rocío Titiunik, and Elisa Wirsching, and seminar audiences at Berkeley and Stanford for helpful comments.

States are prolific producers of information and data (Scott, 1998). Yet, not all state data are created equal: the quantity and quality of state administrative data are known to vary across countries, subnational units, and indicators (e.g., Martínez, 2021; Jerven, 2013). However, our understanding of the processes that produce variation in data quality is less developed. Prominent explanations for poor or uneven quality data are frequently reduced to variation in state capacity (Lee and Zhang, 2017; Brambor et al., 2020). But most administrative data is produced by actors with varied goals and constraints on reporting behavior. Documenting how data is produced is essential for understanding the consequences of use of these data by governments.

An emerging body of research focuses on the central role of bureaucrats in the production of state data (Eckhouse, 2022; Garbiras-Díaz and Slough, 2026). These empirical settings serve as manifestations of the longstanding idea that bureaucracies are institutions that produce and process large volumes of information (Simon, 1973; Radner, 1992). Existing literature on information transmission and aggregation in bureaucracies uses multiple traditions to understand how information is transmitted "up" through organizations from agents to principals (e.g., Tirole, 1987; Radner, 1992; Garicano, 2000). In contrast, firsthand accounts of bureaucratic data collection often stress the importance of bureaucrat-to-bureaucrat (agent-to-agent) information sharing or collaboration. Understanding this type of agent-to-agent interaction requires a clearer idea of the interactions between bureaucrats who work in the same organization. In this paper, I argue that the organizational structure of bureaucracies shapes incentives to engage in agent-to-agent collaboration, with implications for the quality of administrative data.

This paper considers the role of information sharing between bureaucratic agents in the same organization in the production of administrative data. The central task is to report a (set of) measure(s) to, for example, a national government agency. A bureaucrat tasked with reporting is uncertain about the true value, while their colleagues not tasked with reporting observe it directly (or with less noise). The reporter can choose to collaborate to increase the precision of their belief. However, bureaucrats care about more than getting it right: they may also have preferences over the measure reported insofar as it reflects upon their performance or could be used by the national agency to shape policies that affect their organization (Garbiras-Díaz and Slough, 2026). These motives—and resultant strategic reporting—induce bias in reports, reducing (and possibly eliminating) the informational value of collaboration. Within this framework, bureaucratic collaboration improves the quality (accuracy) of the reported data, but bureaucrats may be less likely to collaborate when collaboration is costly or a reporter's preferences over the use of the data diverges from those of their colleagues.

Incorporating bureaucratic organization into this account helps to understand when bureaucrats choose to collaborate. First, selection into and career progression through bureaucracies endows bureaucrats with different levels of information and different preferences over what is reported to the national agency. In the context I document, political appointees disproportionately oversee individual units and therefore have both (1) substantially less experience in a given organization *and* (2) distinct career concerns that induce different preferences over the reported data. Second, bureaucracies have different pre-existing institutional practices for information sharing across units/offices, which modify the costs of seeking information from other bureaucrats. Both organizational features therefore affect whether a given reporter chooses to collaborate, inducing variation in the quality of the resultant administrative data.

Empirically, I study agent-to-agent collaboration in the production of a new administrative dataset on municipal characteristics for use in the administration of social programs in Colombia, partnering with the national government agency that relies on reports from one bureaucrat in each local mayoralty (*alcaldía*). I embed questions about collaboration in the data collection instrument, generating several original measures of the degree to which the reporter consulted colleagues. By randomly assigning a light-touch encouragement treatment promoting collaboration, I show that all measures of collaboration can be increased on the margin. However, these increases are small and reported collaboration is already widespread in the absence of encouragement, consistent with existing work on the frequency and importance of peer-to-peer interactions in bureaucracies

(Brehm and Gates, 1997; Zacka, 2017).

I show that three pre-specified measures of data quality—congruence with other data sources, (lack of) missingness, and elicited respondent certainty—increase in all measures of collaboration. Additionally, meta-data outcomes provide support for the assumption that collaboration entails some investment of effort: bureaucrats reporting collaboration completed the survey more quickly—indicating greater preparation—and enumerators contracted by the national government to record the data assessed them to be more competent and decisive. Two empirical strategies yield findings that suggest that increases in the quality of data are driven by a combination of selection into collaboration and the increase in data quality stemming from collaboration (a treatment effect).

I then show that bureaucratic hierarchies condition selection into collaboration. I rely on two measures that contextualize the role of bureaucratic hierarchies. First, I use the (reporter) bureaucrat's position within their own secretariat or dependency. Second, using an original collection and coding of *alcaldía* organigrams (organizational charts), I examine the degree to which dependencies are structured vertically (hierarchically) versus horizontally. Quantitative analysis suggests that higher ranking appointed bureaucrats are less likely to collaborate and hierarchical organization of units may facilitate collaboration on the extensive margin.

A collection of thirty semi-structured interviews conducted after the data collection elucidate bureaucrats' incentives for collaboration and their motivations in reporting. The interviews document that data production is a routine task within local bureaucracies and highlight bureaucrats' incentives for collaboration. These interviews suggest that different incentives over the use of data and different institutional practices for collaboration/communication across entities condition the level of agent-to-agent collaboration within an organization. In so doing, they support the link between the theory and quantitative findings.

This paper makes two principal contributions. First, it encourages bureaucratic politics scholars to bring organization back into the study of bureaucracy. While discussions of organization feature prominently in classic work (e.g., Wilson, 1989), recent literature in political science and political economy largely reduces institutional questions to principal-agent (or politician-bureaucrat) relationships (see reviews by Gailmard and Patty, 2012; Grossman and Slough, 2022; Brierley et al., 2023). This paper instead focuses on agent-to-agent communication and learning within bureaucratic hierarchies (Simon, 1962), building on studies of peer influence and accountability including Brehm and Gates (1997) and Zacka (2017). I show that bureaucrats' position within their unit predicts their decision to seek information from colleagues, and that the organization of units or teams covaries with the costs—and therefore the likelihood—of collaboration. The latter finding comes from a novel source of data: organigrams (or organizational charts), which could be used by other bureaucratic politics scholars to further measure bureaucratic organization.

Second, this paper emphasizes the role of bureaucrats who are neither high-level appointees nor front-line service providers in (mostly) small local government organizations. Whereas the American Politics literature emphasizes high-ranking policymaking bureaucrats at the national level, Comparative Politics scholarship increasingly studies service provision by local bureaucrats (Brierley et al., 2023; Grossman and Slough, 2022). Building on Carpenter (2001) and Carpenter and Krause (2012), I show that mid-level bureaucrats have important implications for governance outcomes—here, through an understudied bureaucratic task (data production for the national government) within much smaller organizations. Because these bureaucrats are less visible than high-level national appointees or front-line teachers and healthcare workers, this paper provides novel measurement of the behavior and incentives of a harder-to-measure elite sample.

1 Theory

I describe a simple model of data production here which is formalized in Appendix A1. Consider the behavior of a bureaucrat who is tasked with reporting data to a principal about some state (or aspect) of her jurisdiction. The report of the bureaucrat is used as administrative data by the principal in subsequent policymaking or policy implementation. In order to restrict attention to communication among bureaucrats in the reporting entity, I focus on an interaction between two bureaucrats: one who has been tasked with reporting the state (from hereon, the reporter) and a colleague. Whereas the reporting bureaucrat does not observe the state, her colleague—perhaps a sector expert—does. The model elucidates: (1) the conditions under which the reporter would choose to consult their colleague; and (2) the effect of such collaboration on the report issued by the bureaucrat.

The model builds upon a standard cheap-talk model (Crawford and Sobel, 1982). The central departure is that the reporter (analogous to the receiver) must pay a cost to access the report of her colleague (analogous to the sender). This cost captures the time and effort necessary to consult the colleague, the thought necessary to interpret their report, or an expectation of future requests in the other direction. The imposition of a cost on the reporter—as opposed to the colleague—stands in contrast to some models of knowledge transmission in hierarchical organizations (e.g., Radner, 1993; Garicano, 2000). This departure is purposeful insofar as I seek to model the choice to seek information from a colleague without assuming that the reporter will seek information.[1]

As is standard in cheap talk models, the reporter and colleague may have diverging preferences over the quantity that should be reported to the principal. These diverging preferences may be driven by epistemic disagreement over the attribute being measured. For example, in the context of social policy or services, there may be disagreements over the extent of poverty or the most common problem in a jurisdiction. Disagreements may also be induced by anticipated *use* of the data reported by the reporter (e.g., Camacho and Conover, 2011; Cook and Fortunato, 2022; Garbiras-Díaz and Slough, 2026). For example, higher rates of poverty might reflect failures of an incumbent administration to promote economic development but may draw higher investment from national development or anti-poverty schemes.

Consider first what we would expect in a world in which the reporter and colleague's prefer-

[1] In the empirical setting described below, I document variation in the degree to which the reporter seeks information from a colleague, providing support for this modelling assumption.

ences over the data are perfectly aligned. This means that both bureaucrats are simply trying to report the same quantity. In this scenario, if the reporter consults her colleague, the colleague's information reveals the state, thereby allowing the reporter to issue her optimal report given the state. When choosing whether to consult the colleague, therefore, the reporter would simply weigh this (maximum possible) increase in precision against the cost of collaboration.

Admitting disagreement between the bureaucrats changes this straightforward consideration. Consistent with standard results, the presence of disagreement limits the colleague's (sender's) incentive to communicate the state that they observe. Indeed, when disagreements are present, the colleague reports strategically, which limits or eliminates the reporter's ability to learn from the report. Relative to the case without disagreement, thus, the strategic reporting behavior by the colleague reduces the informational gains from collaboration, thereby dampening incentives to collaborate. For sufficiently low levels of disagreement, the reporter still receives information from consulting their colleague, but must weigh these smaller precision gains against the cost of collaboration.

The model yields findings that are relevant for understanding the choice to collaborate and the effects of collaboration. Within the model, the bureaucrat pays for the colleague's report in order to increase the precision of her posterior belief about the state. Disagreement (preference misalignment) shrinks the set of conditions under which the bureaucrat is willing to collaborate because it reduces learning from the colleague's report. This reduction in learning comes from the colleague's strategic efforts to move the reporter's report toward their ideal point.

Collaboration should yield (weakly) more accurate reports. Within cheap talk models, in equilibrium, the colleague's report may or may not allow the reporter to update their belief about the state. However, the reporter would never pay a cost to access their colleague's report if it were not informative. This means that chosen collaboration must have positive informational value because the expected precision gain exceeds its cost. Thus, when examining the relationship between (endogenous) collaboration and accuracy of data outputs, the positive association is driven both by the treatment effect of collaboration and by selection on those treatment effects (bureaucrats opt into collaboration when the informational gains are highest).

1.1 Bureaucratic organization and hierarchy

The account, to this point, lacks consideration of how the bureaucrats fit into a broader institutional structure. Yet, the idea of bureaucrat-to-bureaucrat (agent-to-agent) communication and collaboration makes it important to consider how organizations may facilitate or hinder this collaboration. A defining organizational feature of bureaucracies—and complex organizations generally—is hierarchy (Simon, 1962). Two strands of literature focused on communication within bureaucracies consider different forms of hierarchy.

Organization of units: The first strand of literature considers communication between *units* or *teams* within an organization, focusing on how vertical organization increases the speed at which information moves up through an organization (Simon, 1973; Radner, 1992). Strategic communication introduces frictions through agenda control by lower-level agents (Hammond, 1986) or collusion between agents and supervisors (Tirole, 1987), but the standard rationale for hierarchical design remains the efficient vertical transmission of information (e.g., Garicano, 2000). Information sharing should therefore be less costly when units are vertically oriented, increasing the probability that the reporter seeks to collaborate.

Bureaucratic organizations vary in the degree to which units are organized horizontally versus vertically. Accounts of horizontally differentiated organizations suggest additional (potential) impediments to information sharing in these organizations. Specifically, efforts to maintain autonomy or "turf" can create rivalry between separate (co-equal) units within the same organization (Wilson, 1989). Of course, turf wars do not emerge in all such settings. However, other production strategies such as delegation or referrals across co-equal units do not require collaboration (Herrera, Reuben, and Ting, 2017). This suggests that beyond the informational advantages of vertical hierarchical organizations, greater horizontal differentiation can actively increase costs of information. As a result, (relatively) more vertically organized entities are likely to feature higher levels of collaboration or information sharing.

Hierarchies within units: Beyond the costs of information transmission, the model in Appendix A1 suggests that the reporters will be less likely to collaborate as differences in preferences between individual bureaucrats increase. While there is likely substantial idiosyncratic variation in individual bureaucrats' preferences over reported data, I focus on systematic differences in preferences induced by the type of bureaucratic appointment. Specifically, I consider the different incentives faced by managers versus subordinates within individual units. One feature of the context I study—which is common to many bureaucracies—is that managers are appointed by politicians whereas rank-and-file subordinates are career bureaucrats. A long literature compares the motivations and outputs of appointed versus career bureaucrats (see Hollibaugh (2019) for a review).

Appointed bureaucrats generally have shorter tenures than career bureaucrats, in part because they are easier to remove (Doherty, Lewis, and Limbocker, 2019). This has two implications. First, appointees have less information and expertise—or at least less institutional memory—about the organizations they serve than career bureaucrats (Gailmard and Patty, 2007), suggesting they may have more to gain from collaborating with career colleagues.[2] Second, the transient nature of these appointments induces distinct preferences over the *content* of the data reported. Because an appointee must be routinely prepared to find another job when the elected principal turns over (or before), they face a need to document short-run achievements.[3] This creates an incentive to credit-claim through reported data—an incentive well-documented among politicians (Grimmer, Messing, and Westwood, 2012; Bueno, 2021)—but directed at future employers (e.g., other elected executives or private sector firms) rather than the public.

[2]This is outside of the model since the reporter is always uninformed, but it is straightforward to consider an extension where the reporter observes the state with non-zero probability.

[3]Work on political appointees emphasizes those in the highest tiers of the US federal bureaucracy with lucrative private sector exit options (e.g., Luechinger and Moser, 2014; Kanter and Carpenter, 2023); less attention has been devoted to lower-level appointees with different exit options.

In some domains, claiming credit for recent accomplishments comes at the expense of resources for longer-run deficiencies: claims of alleviated poverty, reduced crime, or improved health outcomes can reduce resources allocated to promote economic development, police services, or healthcare—precisely what career bureaucrats rely on to extend their portfolios. In scenarios—like the empirical application—where claims of short-run progress and prospects for future resources push optimal reports in opposite directions, appointment status induces a wedge between the preferences of appointed and career bureaucrats over the reported data, reducing incentives for collaboration. Further, since better-informed colleagues are more likely to be career bureaucrats than appointees, these biases should disproportionately depress collaboration by reporters who are appointed bureaucrats.[4]

The distinct incentives of appointed versus career bureaucrats thus generate contrasting expectations for collaboration. On one hand, less knowledge or institutional memory about the organizations/jurisdictions they serve should heighten the informational value of collaboration for appointees. However, when appointees' preferences diverge substantially from other bureaucrats with the relevant information, strategic reporting by colleagues renders collaboration less valuable to appointees (relative to career counterparts). These competing expectations posit an empirical question about the relationship between bureaucrat type—appointees versus career civil servants—and their likelihood of seeking information from colleagues.

2 Context

Colombia is a unitary country that has become highly decentralized over the past 35 years (Falleti, 2005). Through the process of decentralization, local governments have increased in both their policy responsibilities and the degree of professionalization or capacity (Fizbein, 1997). These local governments are now responsible for providing the inputs to many national government admin-

[4]If career bureaucrats share similar biases (as a group) for using data to seek resources, a reporter who is a career bureaucrat should be less likely to abstain from seeking information from a career colleague.

istrative data indicators (Camacho and Conover, 2011; Garbiras-Díaz and Slough, 2026; Slough, 2026).

2.1 Local bureaucracies

I measure the behavior of bureaucrats serving in Colombian local governments (*alcaldías*). Local governments are led by an elected mayor. Mayors are elected every four years and cannot run for consecutive re-election. The data collection process that I study took place in late 2022 and 2023, during the final 14 months of the then-mayors' terms. Mayors oversee local bureaucrats. They generally have some latitude to staff local bureaucracies, either by hiring contractors or civil servants, or by transferring existing bureaucrats to other posts within the *alcaldía*. For tasks like data collection that fall outside of the direct purview of a given secretariat or dependency, mayors may have scope to assign the task to a specific bureaucrat.

Local bureaucracies exhibit some important variation in *de-jure* organization across municipalities. In all municipalities, a set of secretariats sits directly under the mayor's dispatch. The specific set of secretariats varies in number and composition across municipalities. Beyond the secretariats, most *alcaldías* have additional dependencies that sit under specific secretariats. Again, the composition of these dependencies varies across municipalities. In some municipalities, there exists a third layer of dependencies. Leaders of dependencies are appointed to these positions by the mayor. These leaders (secretaries or other directors) are often selected from outside the set of existing civil servants and have shorter tenures within these local bureaucracies.

Figure 1 illustrates some features of this variation. In the left panel, I describe the prevalence of the 12 most common secretariats/dependencies across a sample of 832 municipalities that publish organigrams (see Section 3.5 for additional detail). Specifically, the prevalence of these entities ranges from 88.9% (Planning Secretariat) to 27.5% (Education Secretariat). Importantly, this variation is not simply semantic. Examining the correlation between the presence of 62 types of entities (1,891 total bivariate correlations), only two pairs of entities are negatively correlated at

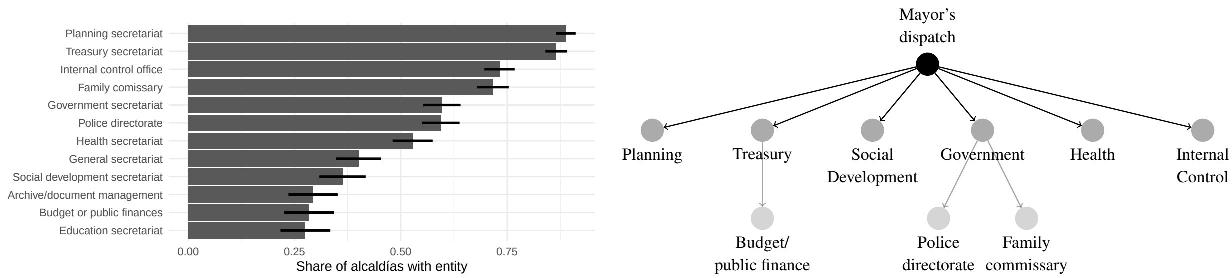

Figure 1: The left panel reports the prevalence of the 12 most common entities in local bureaucracies. The right panel plots the organizational structure for a representative municipality (Morales, Bolívar).

less than -0.3 (Table A6). This suggests that this variation is not simply driven by substitution of one unit name for another. In the right panel, I plot the organizational structure of a representative *alcaldía*. This *alcaldía* was randomly sampled among those comprised of: the median number of secretariats under the mayor's dispatch (6), the median number of entities below the secretariat level (3), and the median number of levels below dispatch (2).

Each of the dependencies in the left panel of Figure 1 (and their analogues in other municipalities) are staffed by bureaucrats of different ranks, whose incentives vary. It is worthwhile to consider two related dimensions along which these individuals vary. First, among civil servants, there are generally five categories of positions: leadership (*director*), professional (*profesional*), technical (*técnico*), support (*asistencial*), and advisory (*asesor*). The first four categories—those that are present in all municipalities—are ranked.[5] Figure A8 presents public sector wage ranges from a stratified random sample of municipalities gathered from the System for the Information and Management of Public Employees (SIGEP). It shows a clear ranking of these categories: leadership, professional, technical, and support, albeit with some wage compression between the lowest two levels (support and technical employees).

Public employees are employed as career bureaucrats (*funcionario de carrera*) or as bureaucrats of free appointment and removal (*funcionario de libre nombramiento y remoción*). Within

[5] See Martínez, Snowberg, and Ting (2024) for further detail on these positions.

municipal governments, employees within the leadership category and advisors (where present) are classified as bureaucrats of free appointment and removal according to national public employment regulation.[6] This means that these leaders (e.g., secretaries of secretariats, or directors of other dependencies) can be appointed and removed freely by mayors. Moreover, if a career bureaucrat (e.g., a professional) is appointed to a leadership position, her form of employment also changes to a bureaucrat of free appointment and removal. This generates changes in incentives between leaders—who are beholden to mayors for continued employment—and career bureaucrats who tend to serve in their specific position and *alcaldías* longer on average (see Figure A9). Moreover, since mayors are elected every four years without consecutive re-election, directors churn through positions and *alcaldías* at a high rate. This feature means that leaders': (1) current positions are more dependent on maintaining the favor of the mayor; and (2) future employment will depend on relationships and achievements within their current position.

2.2 Prosperidad Social (PS)

A portion of this study was conducted in partnership with Prosperidad Social (PS), a national government department that administers most of Colombia's social welfare programs, including social assistance programs and support for conflict victims. The current collaboration with PS took place in 2022. In the course of the collaboration, Gustavo Petro was elected president and assumed office in August 2022. Petro, Colombia's first leftist president in decades (Freeman, 2023), proposed a number of new social programs and expanded subsidies as part of his platform. It is useful to note that the present data collection exercise (detailed in the next section) was planned prior to Petro's election. However, given the timing of data collection, in interviews, some of the local bureaucrats responsible for collecting and reporting information associated this data collection with Petro's vision for an expanded PS portfolio.

This project was realized in collaboration with the "Targeting Group" within PS. This group

[6] See Article 5 of *Ley 909 de 2004*.

has a scope that cross-cuts all of the program-specific mission directorates within PS. It primarily supports the design, launch, and implementation of new social programs. This support consists of helping each directorate to identify which indicators should be used for planning and targeting of each program. In this sense, the group seeks to provide data or evidence to improve policymaking across the organization. Members of the Targeting Group noted in interviews that the goal of using data or indicators to guide policy across the organization is often aspirational in practice when their recommendations come into tension with the (political) goals of appointed mission directors.

2.3 *Diagnostico Territorial* (DT) Data Collection

This paper studies data collection in the context of an effort by PS to collect wide-ranging municipal-level data for use in targeting social programs. The primary data source used for targeting by PS is the household-level System for the Identification of Potential Beneficiaries for Social Programs (SISBÉN). While these data can be aggregated to the municipal level to understand relative or per-capita levels of poverty and vulnerability, they do not directly measure municipal-level contextual features that are relevant for program administration. In a semi-structured interview, a member of the Targeting Group explained that the goal of DT was to learn "what exists in the territory, what was missing and how we could address it through our programs."

The DT survey instrument included four modules: institutional capacity (up to 214 questions), geographic mobility (up to 15 questions), infrastructure (up to 141 questions), and economic development (up to 57 questions). Note that many questions utilized a conditional structure; these figures therefore represent the maximum possible survey length. The content of the administrative data collected via the survey was developed within the PS' targeting group over a period of nearly five years.

Collecting this type of original administrative data at the municipal level is costly, and the reliance on individual bureaucrats within *alcaldías* was adopted only after two other methods were tested and abandoned. In a first attempt in 2019, PS sought to hire contractors to travel to municipalities to collect the data from multiple informants, but abandoned this strategy after just two trial municipalities. In a second attempt in 2021, PS sought to delegate data collection to regional PS offices, but found these offices were already overburdened by requests from the national government. Responsibility for data collection therefore fell to local bureaucrats, who were tasked with preparing the data and reporting responses by phone to a firm contracted by PS. Bureaucrats were provided with a questionnaire in October 2022 prior to these interviews in November-December 2022.

As in other instances in which national data is collected from local governments, PS sought to minimize the participation of local politicians in the data collection. To do so, they directed the request for data to a contact—always a bureaucrat—within the *alcaldía*. The survey instrument also noted that the desired profile of the bureaucrat would be "an official from the Planning Secretariat, ideally with at least five years of experience working in the *alcaldía*." This strategy aimed to insulate data outputs from distortion by politicians. As I document below, however, not all respondents had this profile.

3 Research Design

To learn about how collaboration and communication within local bureaucracies influences data production, I implemented a research design situated within the context of collection of the DT data. I complement this research design with systematized, semi-structured elite interviews of 30 bureaucrats who reported DT data. I then collect original data on and develop new measures of the institutional structure of local bureaucracies. I detail the design of each attribute of the research design below.

3.1 Measuring Collaboration within Local Bureaucracies

In addition to the DT questions about institutional capacity, mobility, infrastructure, and economic conditions, PS allowed the research team to embed several questions about preparation and reporting into the survey instrument. Unlike the core survey items, these questions about process were not revealed to bureaucrats in advance. In this sense, bureaucrats could not anticipate or prepare optimal responses to these questions. After the mobility, infrastructure, and economic conditions sections, each bureaucrat was asked:

1. "To prepare questions to this section of the survey, did you consult any individual or entity in the *alcaldía* or municipality?" (Yes/No)
2. (If "Yes") "Which dependencies of the *alcaldía* or entities did the officials that you consulted represent?" (Open ended)

For the second question, enumerators reported responses with respect to the institutional functions reported as "present" in the institutional capacity section. These questions allow measurement of the extensive margin of collaboration (through the first question). Moreover, they allow for measurement of collaboration along the intensive margin in two ways. First, by asking these questions repeatedly, we observe roughly the share of topics for which bureaucrats sought information. Second, by measuring the share of dependencies, we observe the set of offices from which advice was solicited.

Using these self-reported measures, I pre-specified four measures of collaboration within *alcaldías*. Consider reported collaboration for survey section $s \in \{\text{Mobility}, \text{Infrastructure}, \text{Economy}\}$ for municipality i. The variable, Collaboration$_{is}$, is a binary indicator coded from question #1 above. The extensive-margin measure of collaboration is a binary indicator for collaboration on any section of the survey:

$$D_i^1 = \mathbb{I}\left[\sum_s \text{Collaboration}_{is} > 0\right].$$

The intensive-margin indicators measure the number of sections s for which the bureaucrat consulted with others:[7]

$$D_i^2 = \frac{\sum_s \text{Collaboration}_{is}}{|s|},$$

and the share of dependencies that the bureaucrat consulted in any section of the survey:

$$D_i^3 = \frac{\sum_{l=1}^{L_i} \mathbb{I}\left[\sum_s \text{Collaborator}_{ils} > 0\right]}{L_i}.$$

In this expression, L_i is the number of dependencies reported in municipality i and Collaborator_{is} is an indicator that takes the value 1 if an official from the dependency was consulted in section s in response to question #2 above. Note that the normalization by L_i deviates from the pre-specified measure but is necessary because L_i is strongly (and positively) correlated with municipal population. Finally, one may be concerned that collaboration involves the mayor, potentially politicizing reporting in line with widely held concerns by national government agencies in Colombia. To this end, I measure whether the bureaucrat mentioned consulting with the mayor, using the following measure:

$$D_i^4 = \mathbb{I}\left[\sum_s \text{Mayor}_{is} > 0\right],$$

where Mayor_{is} is an indicator that takes the value 1 if the bureaucrat mentioned the mayor in section s in response to question #2.

There are two potential concerns with these measures of collaboration. First, they are constructed from self-reports by bureaucrats and may be subject to (experimenter) demand effects. Two observations help to allay this concern. First, by omitting these questions from the instrument that was distributed to bureaucrats in advance, there was no opportunity to plan a response or

[7] The normalization by $|s| = 3$ departs from the pre-specified measure, but is simply a rescaling that ensures that all of the measures of collaboration range from 0 to 1.

consider (in detail) the optimality of different response choices. Second, in qualitative interviews, respondents who did not report consulting others routinely expressed that they had not consulted others, without apparent shame or hesitation. Responses between the self-reports and interviews are positively correlated even though the interviews were conducted three to four months later.

Second, one may be concerned that these measures proxy for the diligence or preparedness of an individual bureaucrat rather than information gathering from other agents. Investment in acquiring information is an important input to preparedness—per the model, it captures costly investment of effort to acquire information to report to the national government—but bureaucrats could also assemble or seek information without consulting colleagues. The research design uses three approaches to isolate the role of collaboration from other sources of information. First, an embedded experiment examines whether collaboration behavior can be changed on the margin. Second, I exploit variation in collaboration on different sections of the survey (s) to examine how select outcomes vary within individual reporting bureaucrats. Finally, qualitative interviews seek to understand how bureaucrats perceive the role of collaboration as it relates to generalized preparedness.

3.2 Embedded Experiment: Encouragement to Collaborate

To generate exogenous variation in collaboration, we embedded an encouragement experiment in the data collection. While this experiment does not manipulate the deeper incentives or institutional constraints that feature in the theory, it serves three purposes. First, it identifies the share of bureaucrats who are nearly-indifferent to collaborating: those for whom a light-touch encouragement could induce collaboration that would not have happened otherwise. Second, depending on the size of this group, it offers identification of causal effects of collaboration among municipalities exogenously induced to collaborate. Finally, it helps to distinguish agent-to-agent collaboration (the focus of this study) from communication between bureaucrats and elected mayors by assessing the degree to which these outcomes move together.

To generate this exogenous variation, in collaboration with PS and IPA, we randomly assigned a message promoting collaboration in the email notifications about the PS survey. Specifically, the treatment group received an additional (bolded and colored) paragraph in the invitation that stated:

> "Due to the thematic breadth of the questionnaire, we suggest that you consult with other officials from the *alcaldía* or the municipality for information that you consider necessary in preparation for the survey."

We employed block random assignment to ensure that municipalities in each administrative category (special and 1-6)—which are based on population size and strategic relevance—were assigned to treatment with probability $1/2$.[8] Since most municipalities are small, blocking ensures that more populous municipalities and those with more resources are evenly allocated to treatment and control.

3.3 Outcome measures

The DT data affords measurement of a number of outcomes that facilitate understanding of data production. I focus on three families of outcomes: the *quality*, *content*, and *metadata* characterizing DT survey data and responses. Table 1 summarizes the outcome measures of interest.

To measure data quality, I consider three classes of measures. First, I evaluate rates of missingness on unconditional questions, those asked of all municipalities regardless of municipal characteristics or responses to other questions. While rates of missingness are generally low, I interpret greater missingness as consistent with lower quality data. Second, I compare reported data from DT to analogous (or closely related) measures obtained from a different source. For example, I compare responses to questions about the presence of five ethnic groups in a municipality to measures of these populations on 2018 census data. While most of the questions on the DT survey lack an external referent (and are therefore not redundant), I identify three subsets of DT questions—on

[8]Note that because public sector wages and staffing vary by municipal category, this is an important blocking variable for a study of bureaucratic organization.

Family	Outcome		Description
Quality	Missingness	$\rightarrow$	Proportion of responses to unconditional questions asked of all municipalities/respondents.
	Accuracy	$\rightarrow$	Where feasible, comparison to different data sources with different data production processes. See Appendix A3.1 for discussion of sources.
	Certainty	$\rightarrow$	Self-reported level of certainty elicited for three responses in the course of the survey.
Content	Reported data	$\rightarrow$	First principal components constructed from all unconditional questions, by section. This results in four indices.
Metadata	Duration	$\rightarrow$	Duration of survey administration (in minutes)
	Assessment	$\rightarrow$	Enumerator assessment of respondent (a) confidence; (b) decisiveness; and (c) competence during survey.

Table 1: Summary of outcome families and measures from the DT survey.

demographics, natural disasters, and schools—for which relevant data are available. I show that these responses *are* more similar than one would expect by chance alone for all measures (see Figures A5-A6). Finally, to the extent that respondent uncertainty leads to noise in response, several additional questions were embedded in the DT survey to elicit respondent certainty in three responses. These questions were absent from the draft of DT questions that respondents prepared.

Second, the four modules of the DT survey sum to 427 responses per municipality. There is therefore a need to select individual questions or reduce the dimensionality of the data when measuring content. The targeting team at PS did not identify some measures as more important than others, nor did it seem prudent to hand select measures for this analysis. Instead, I conduct principal components analysis on each of the four survey modules. These principal components include only unconditional questions.[9] Appendix A3.2 provides additional information on the resultant principal component outcomes.

Finally, I use two outcome measures beyond responses of individual bureaucrats to the survey; I refer to these measures as forms of metadata. I first measure the duration of responses. The survey was conducted by phone with a fixed set of enumerators. The survey software measured the duration of the survey. Enumerator observation suggested high variation in bureaucratic prepara-

[9]Where responses were missing, I impute the modal response (categorical questions) or the median (open-ended numeric responses).

tion. This preparation facilitated more efficient progress through the survey, so I interpret duration as a measure of (lack of) preparation. I also use three evaluations of bureaucrats by survey enumerators. These questions for enumerators were appended to the DT survey and were filled out at the conclusion of the phone interviews. The questions measure respondent (a) confidence; (b) decisiveness; and (c) competence on five-point scales. Enumerators were directed to evaluate the interviews holistically and were not guided to focus on any specific response. For both metadata outcomes, it is important to compare variation *within enumerator* as some enumerators may have conducted the survey more efficiently or evaluated bureaucrats more harshly than others.

3.4 Qualitative Interviews

In order to understand the processes of collecting and compiling the DT data, I commissioned thirty qualitative interviews of the bureaucrats that prepared and responded to the DT questionnaire. The interviews were conducted in April 2023, four to five months after the DT survey. This slight delay facilitated the use of responses to the DT survey in sampling. Specifically, I employed stratified random sampling on the basis of (1) municipal category; (2) reported collaboration (using the extensive margin indicator D_i^1); and (3) assignment to encouragement. Given the distribution of municipal category and the fairly widespread practice of collaboration, the sampling overweights large municipalities and municipalities without reported collaboration. Of course, not all respondents agreed to participate in an hour-long interview. Unfortunately, as Table 2 reports, among municipalities in which collaboration was not reported, few ($n = 4$) bureaucrats accepted the interview invitation, which means that the share of respondents who reported collaboration (87%) ultimately resembled the share of DT respondents who reported collaboration (88%).

The interviews were conducted by video call by a team of trained Colombian anthropologists. I relied on a team of anthropologists rather than conducting the interviews myself for two reasons. First, past research with local bureaucrats in Colombia suggests that interactions between these bureaucrats and a foreign researcher varies substantially across municipalities, consistent with ob-

Characteristic	Share among municipalities... All ($n = 1,122$)	Completed DT ($n = 906$)	Invited to interview ($n = 124$)	Interviewed ($n = 30$)
Not in smallest mun. category (6)	0.12	0.12	0.48	0.60
Reported collaboration ($D_i^1 = 1$)	–	0.88	0.65	0.87
Assigned to encouragement	0.50	0.50	0.50	0.53

Table 2: Characteristics of municipalities at different stages of selection into interview sample as a function of covariates used in stratification.

servations of variation in positionality in elite interviews (Glas, 2021; Soedirgo and Glas, 2020). While delegation to a team of Colombian anthropologists does not eliminate considerations of positionality, it sought to attenuate this variation across very different municipalities. Second, while the anthropologists were trained on the topics of investigation and the structure of the DT survey, they were blinded to the presence of an experiment and specific hypotheses being tested. This provided a guard against using semi-structured interviews to prime answers or quotes to support my own hypotheses. The interviews generated detailed notes by the anthropologists in addition to recordings (and transcripts) of the conversations.

3.5 Organigram Collection

The third source of data that I employ is an original collection of organigrams from local *alcaldías*. Public entities in Colombia are mandated to post organigrams (organizational charts) for purposes of transparency. I collected all available organigrams ($n = 832$) from the websites of all *alcaldías* that submitted DT data ($n = 902$ of the 1,102 invited municipalities). I used these sources to create a dataset that encodes (1) the set of entities/dependencies reported on the organigram; (2) the location of each unit/dependency in the hierarchy.

The measurement of organizational structure and hierarchy is still rather nascent (Carbonell-Nicolau, 2024); the organigrams depict much more simplistic organizations than many measures are designed to quantify. I therefore rely on three simpler measures. Consistent with Figure 1, all municipal bureaucracies consist of a number of secretariats directly below the mayor's dispatch.

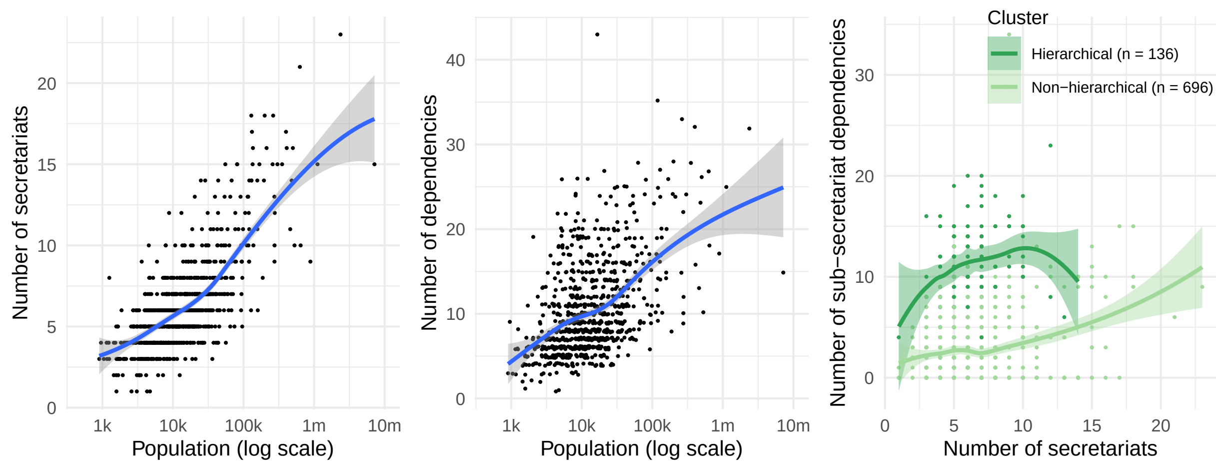

Figure 2: Three measures derived from the coding of organigrams: the number of secretariats (left), the total number of dependencies/units (center), and clusters derived from hierarchical clustering on the set of entities (right).

The number of secretariats generally increases in municipal population ($\rho = 0.72$ with logged population). In contrast, the relationship between population and the total number of dependencies is non-trivially weaker ($\rho = 0.49$) because some smaller municipalities have many more sub-secretariat dependencies. The first two panels of Figure 2 plot the relationship between these measures and municipal population.

The data also permit analysis of the specific set of secretariats and dependencies in each local *alcaldía*. Specifically, I generate a matrix of indicators of the presence of each type of dependency (e.g., planning secretariat or family commissary) for all entities found in at least four municipalities. I then implement hierarchical clustering to classify municipalities in clusters on the basis of the composition of entities. The right panel of Figure 2 depicts the two resultant clusters.[10] One cluster—which I refer to as the "hierarchical" organizations or *alcaldías*—has a much higher number of sub-secretariat dependencies than the "non-hierarchical" organizations, even accounting for municipal population.

Table 3 shows how these two groups of organizations differ. They do not differ in the average

[10]Increasing the number of clusters above two generates small or singleton cluster(s) with only the largest cities.

Characteristics	Counts			Count/Log population		
	Non-hierarchical	Hierarchical	Δ	Non-hierarchical	Hierarchical	Δ
# Secretariats (level-1 span)	6.29 [6.07, 6.51]	6.40 [6.04, 6.75]	0.11 [-0.31, 0.53]	0.65 [0.63, 0.66]	0.66 [0.63, 0.69]	0.01 [-0.02, 0.05]
# of Dependencies/secretariat	0.49 [0.45, 0.53]	1.91 [1.78, 2.04]	1.42 [1.28, 1.55]	0.05 [0.05, 0.06]	0.20 [0.19, 0.22]	0.15 [0.13, 0.17]
# of Hierarchy levels	1.76 [1.72, 1.8]	2.11 [2.06, 2.16]	0.35 [0.28, 0.41]	0.19 [0.18, 0.19]	0.22 [0.22, 0.23]	0.03 [0.02, 0.04]

Table 3: Organizational characteristics of the clusters of *alcaldías*. Each cell reports a mean and 95% confidence interval.

number (span) of the highest level of entities (secretariats) in raw or population-adjusted terms. Note, however, that the non-hierarchical organizations have greater variance in this span, as is evident in the right scatter plot in Figure 2. The organizations differ substantially on both the average number of dependencies per secretariat—hierarchical entities have almost four times more lower-level dependencies. They also differ in the number of levels of the organization—hierarchical entities have 0.35 (≈20%) more levels on average. Note that these are shallow organizations with one to three levels, where a one-level organization has only secretariats. In sum, the clusters classified as "hierarchical" contain more dependencies per each level-one (secretariat) entity as well as somewhat deeper organizations.

3.6 Quantities of Interest and Estimation

The encouragement design facilitates our understanding of collaboration within *alcaldías* in the process of data production. I begin by reporting the average treatment effect (ATE) of the encouragement to collaborate (Z_i) on the four measures of collaboration, D_i^j for $j \in \{1, ..., 4\}$ by estimating (1) using OLS. In this equation, δ is the estimator of the ATE. $\boldsymbol{\gamma}_c$ denotes a vector of municipal category fixed effects. Recall that the treatment assignment was blocked on municipal category, so these fixed effects are equivalent to block fixed effects.

$$D_i^j = \delta Z_i + \boldsymbol{\gamma}_c + \varepsilon_i \tag{1}$$

This expression serves as the first-stage of a two-stage least squares estimator of the complier average causal effect (CACE). Note, however, that when the first-stage relationship between Z_i and D_i^j is weak, the two-stage least squares is a biased and inconsistent estimator of the CACE. I evaluate the first-stage relationship using the F-statistic and follow the rule-of-thumb proposed by Stock, Wright, and Yogo (2002).[11]

I also consider the relationship between the endogenous collaboration measures and outcomes of interest. In the OLS regression in (2), β_1 describes the expected change in Y_i when moving from no collaboration ($D_i^j = 0$) to the maximum possible collaboration given measure j ($D_i^j = 1$):

$$Y_i = \alpha + \beta_1 D_i^j + \varepsilon_i \tag{2}$$

Given the importance of state capacity as the dominant existing explanation for the quality of state data, I also employ a set of covariates that seek to capture variation in the capacity of municipal governments. First, I include municipal category fixed effects. A second measure developed by the National Planning Department, the municipal performance index (MDM), seeks to summarize public management and achievement of development outcomes at the municipal level. This is principally a measure of municipal government outputs. Finally, I include an indicator for departmental capitals, since the national government generally has stronger institutional presence in these municipalities. This presence may have spillovers for local governments in terms of the labor market for public employees and greater oversight of local institutions (Toral, 2024). I refer to this set of pre-treatment controls as a benchmark.

$$Y_{ic} = \beta_1 D_i^j + \underbrace{\beta_2 \text{Capital}_i + \beta_3 \text{MDM}_i + \boldsymbol{\gamma}_c}_{\text{Benchmark covariates}} + \varepsilon_i \tag{3}$$

Per the theory, β_1 should be a biased estimator of the average treatment effect of collaboration because it incorporates both selection into collaboration (confounding) and the treatment effect of collaboration on data outcomes. I utilize two strategies to document that both selection and treatment effects drive the association between collaboration and data outputs. First, for outcomes

[11] I report the reduced-form intent-to-treat effects of encouragement on outcomes of interest in Table A7.

that can be measured at the level of the DT module (or section), I estimate regressions of the form:

$$Y_{is} = \eta \text{Collaboration}_{is} + \boldsymbol{\delta}_i + \boldsymbol{\psi}_s + \varepsilon_{is} \tag{4}$$

$\text{Collaboration}_{is}$ is an indicator for reported collaboration on a given module (as above). The municipality fixed effects, $\boldsymbol{\delta}_i$, aim to purge selection that occurs at the level of the municipality or bureaucrat. I also conduct a sensitivity analysis following the diagnostics of Cinelli and Hazlett (2020). This exercise seeks to quantify how large selection would need to be to explain away treatment effects as an explanation for the observed β_1 estimates (or inferences).

Finally, I pursue a number of analyses to describe the correlates of selection into collaboration. For these analyses, I (largely) estimate OLS specifications of the form:

$$D_i^j = \boldsymbol{\delta}\mathbf{G}_i + \beta_1 Z_i + \underbrace{\beta_2 \text{Capital}_i + \beta_3 \text{MDM}_i + \boldsymbol{\gamma}_c}_{\text{Benchmark covariates}} + \varepsilon_i, \tag{5}$$

where $\mathbf{G}_i$ includes a set of one or more organizational or institutional variables, measured prior to the DT data collection exercise.

4 Results

4.1 Collaboration in the process of data production

I document the extent of collaboration in the process of data collection and the effects of the encouragement treatment in Table 4. Collaboration within the *alcaldía* is already widespread. In control: 86.3% of bureaucrats consulted at least one other bureaucrat and the average bureaucrat consulted another bureaucrat for 2.3 (0.765×3) of 3 sections and reported consulting 25.7% of the common set of entities in the *alcaldía*. In contrast, just 4% of bureaucrats report consulting the mayor when preparing responses. This widespread peer-to-peer communication is consistent with the importance of peer-to-peer interactions posited by Brehm and Gates (1997) and Zacka (2017).

	D_1 : Extensive margin (1)	D_2 : Intensive margin (sections) (2)	D_3 : Intensive margin (entities) (3)	D_4 : Mayor consulted (4)
Encouragement	0.043*	0.059**	0.032*	0.011
	(0.021)	(0.023)	(0.014)	(0.014)
Mun. category indicator	✓	✓	✓	✓
Number of obs.	902	902	902	902
Control mean	0.863	0.765	0.257	0.04
Control std. dev.	0.344	0.366	0.203	0.196
F-statistic	4.796	6.996	5.311	0.664

$^{+}p < 0.1; ^{*}p < 0.05; ^{**}p < 0.01$

Table 4: Estimates of the ATE of the encouragement treatment on different measures of collaboration within the *alcaldía*. Heteroskedasticity-robust standard errors in parentheses.

Further, the encouragement treatment increased reported collaboration with other bureaucrats on both the extensive and intensive margins. It induced bureaucrats in an additional 4.3 percent of municipalities to collaborate with at least one other bureaucrat. It also increased the extent of collaboration across survey sections by 0.16 control group standard deviations and increased the share of entities consulted by a similar 0.16 control group standard deviations. There is no detectable increase in bureaucrats consulting the mayor directly. This provides suggestive evidence that different processes are at work: agent-to-agent communication or collaboration follows a distinct logic from better-studied agent-to-principal communication.

Whereas these treatment effects on collaboration within the bureaucracy are distinguishable from zero, they are modest in magnitude. Indeed, the F-statistics of these first-stage regressions are substantially smaller than existing rules-of-thumb (e.g., an F-statistic of 10 in Stock, Wright, and Yogo, 2002). This may be because collaboration was already fairly widespread in the absence of treatment (limiting the scope for increasing collaboration). It may also be a function of the light-touch treatment itself. The estimator of the CACE may be biased when the first-stage is weak. As a consequence, I do not estimate the CACE of the encouragement treatment.

The interviews suggest that collaboration within many *alcaldías* is a routine part of the data collection process. This collaboration takes different forms in different municipalities. Some respondents mentioned informal communications. For example, the head of a research unit in a large departmental capital mentioned contacting secretaries of Planning, Government, and Education through WhatsApp, obtaining a response in 1-2 days. The respondent credited the current administration with reducing the administrative burden of collaboration across the *alcaldía*. In other municipalities, however, the process was substantially more "bureaucratic." In a Category 2 city, in contrast, a career civil servant reported seeking signatures from other secretaries before seeking information. He reported that his colleagues "do not feel as committed to delivering the information" in response to personal email or WhatsApp messages.

Not all bureaucrats reported consulting others in the municipal administration, however. The bureaucrats generally provided one of two rationales for not consulting others. First, some bureaucrats freely shared that they failed to prepare responses to the survey. In these cases, lack of collaboration was simply a consequence of a lack of preparation. Second, others mentioned a lack of collaboration within the *alcaldía* more generally. For example, the director of the development plan in a Category 2 city reported:

> "Unfortunately, right now dependency and each Secretariat of the administration is trying to focus on a successful closure to the administration. I have seen some of my colleagues more concerned about protecting their image as professionals and preparing their resumé for the next administration . . . Each one focuses individually on what they have to do instead of supporting and cooperating with us."

This statement and several others raise the possibility of tension between career incentives of individual bureaucracies and willingness to share information across the *alcaldía*. Moreover, it notes segmentation of secretariats within the organization of the *alcaldía*. I return to these themes in Section 5.

The interviews shed some light on the role of the mayor in the collection of DT information. While few bureaucrats reported consulting the mayor (Table 4), some reported the direct or indirect influence of the mayor in reporting. Several bureaucrats reported that the mayor had specifically delegated the task to them.[12] In other cases, the mayor may have encouraged reporting of certain features. For example, a civil servant in the Secretariat of Social Development of a Category 6 municipality believed that the mayor asked him to complete the survey because of his nearly 30-year tenure in the *alcaldía*. The bureaucrat recalls that the mayor told him: "if you are asked about this, look at the infrastructure, the roads, which is what is most needed so that farmers can harvest their crops" in an effort to draw investment from PS.

4.2 Collaboration and data outputs

I now consider the relationship between collaboration and the data submitted by each bureaucrat. I proceed by reporting the associations between the three measures of collaboration (D_i^1, D_i^2, and D_i^3) and the outcomes in Table 1. Figure 3 reports estimates of β_1 following estimation equations (2) and (3). Collaboration—on the intensive and extensive margins—is associated with higher quality responses. Bureaucrats who report (more) collaboration produce data with less missingness, greater certainty in select responses, and a closer correspondence to existing administrative measures gathered through alternate means. These bureaucrats also completed the survey more quickly and were assessed to be more confident, decisive, and competent by survey enumerators. While the coefficient magnitudes are substantially larger for the second intensive margin measure (right panel), recall that the standard deviation for this collaboration measure is markedly lower (Table 4). This means that a one unit change (from zero to one) is larger relative to the variation in the sample than for the other measures of collaboration, so the larger coefficient estimates should not be surprising.

The relationship between collaboration and the content of responses (as measured by principal components indices) is weaker. Here, the principal components are oriented such that the first

[12]This contrasts with PS's goal of avoiding political interference by directing surveys directly to bureaucrats. It is unclear whether bureaucrats believed they had received the survey at the request of the mayor, or whether the mayor reassigned the survey upon learning about its existence.

principal components from each section (reported in Figure 3) are positively correlated. The discussion in Appendix A3.2 suggests that these principal components (in their current orientation) are negatively correlated with measures of economic or institutional development. In the bivariate regressions, both intensive margin collaboration indicators produce statistically significant, negative estimates of β_1, suggesting collaboration is associated with municipalities that appear poorer or more deprived. However, the capacity controls substantially attenuate these estimates. There is therefore not strong evidence that collaboration induces systematic biases in the content of responses in the one direction. Of course, if collaboration generates biases in both directions, the estimates in Figure 3 would mask this heterogeneity.

While these estimates suggest that collaboration is associated with improved response quality, there are two broad classes of explanations for these findings. First, it may be the case that the types of municipalities that provide higher quality responses are also those that are more likely to collaborate. For example, the qualitative evidence that some bureaucrats simply did not prepare for the survey, likely yielding worse response quality, may provide one such logic for *selection*. Second, the act of collaborating may have *treatment effects* on response quality by leveraging more information. To this end, Table A7 reports intent-to-treat effects from the experiment; while low levels of compliance render these estimates relatively noisy, the sign of estimated ITT effects matches the OLS estimates in Figure 3 for five of six indices. In interviews, many respondents believed that collaboration improved their ability to accurately respond to the questions on the DT survey. For example, a civil servant in the Planning Secretariat from a Category 5 municipality suggested that collaboration yielded more accurate data by leveraging expertise across the institution. He stated: "we work as a team with the directors of the administration ... here, each of the departments handles specific issues, so we send or share the questionnaire and each one gives the answers according to their specialty."

Ultimately, the theory predicts that both phenomena—selection into and effects of collaboration—should be present. Consistent with theory, the quantitative data provide suggestive evidence that

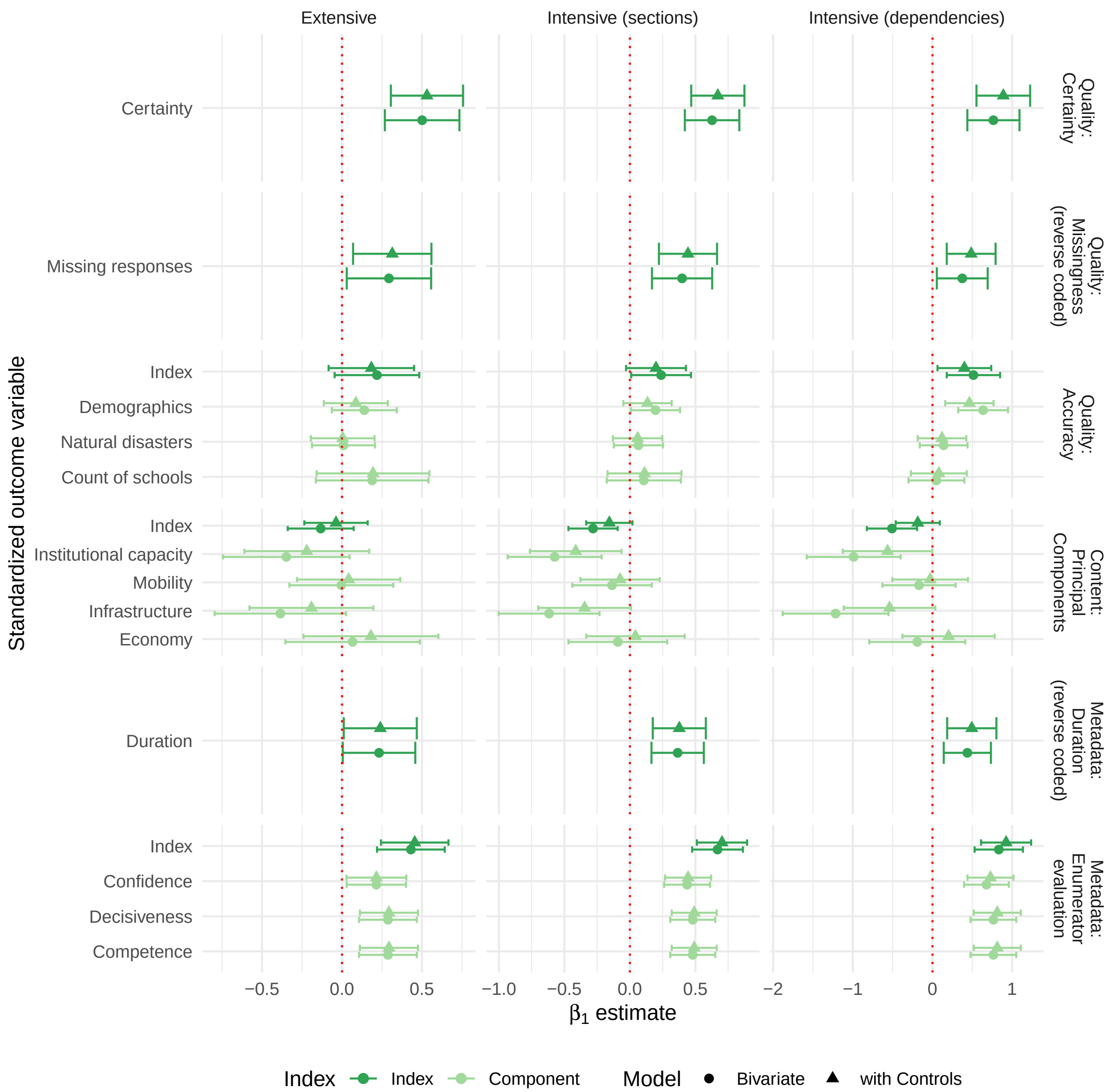

Figure 3: This figure reports estimates of β_1 following (2) and (3). All outcome variables are standardized, so effect sizes can be interpreted in standard deviation units. 95% confidence intervals are calculated using heteroskedasticity-robust standard errors.

both are operative. First, I leverage the fact that collaboration was measured separately for three sections of the survey. Two of the six outcomes—missingness and certainty—were measured using a parallel design within each section. Table 5 measures how the marginal effect of collaboration (in a section) changes when municipality fixed effects are included. If all selection were occurring at the level of the municipality—e.g., because a bureaucrat was insufficiently diligent or simply did not prepare for the survey—we would expect the effect of collaboration to disappear with the inclusion of municipal fixed effects.

Table 5 suggests that the associations between collaboration and data quality are not entirely driven by selection at the level of the bureaucrat/municipality. When considering missingness, the inclusion of municipal fixed effects in Column (3) attenuates the estimate in Column (2) by 28% (from 0.222 to 0.159). Under the assumption that bureaucrats do not purposely select which sections on which to consult a colleague, this finding suggests that 72% of the reduction in missingness on sections where bureaucrats consulted others is a treatment effect of this collaboration. Comparing Column (5) to (6), the same exercise suggests that municipal fixed effects attenuate the association in Column (5) by a more substantial 68% (from 0.318 to 0.103). While this relative reduction is larger in magnitude, the estimate is less precisely estimated ($p = 0.19$). This constitutes a relatively hard test for treatment effects. If a bureaucrat consulted on specific topics on which they have less knowledge, for example, this interpretation should underestimate the effect of collaboration on data quality. In sum, this analysis supports the argument that both treatment effects *and* selection play a non-trivial role in generating the positive association between bureaucratic collaboration and data quality.

Sensitivity analysis provides an alternative approach to considering the role of selection versus treatment effects when thinking about the relationship between bureaucratic collaboration and data quality. Table 6 reports the diagnostics proposed by Cinelli and Hazlett (2020) along with a benchmarking exercise inspired by the substantive literature on state information. For each outcome (index, where relevant) and collaboration measure in Figure 3, I report three diagnostics. The first,

	Missingness (standardized, reverse coded)			Certainty (standardized)		
	(1)	(2)	(3)	(4)	(5)	(6)
Collaborated on section	0.221^{**}	0.222^{**}	0.159^{+}	0.299^{**}	0.318^{**}	0.103
	(0.061)	(0.062)	(0.090)	(0.056)	(0.054)	(0.078)
Encouragement	✓	✓		✓	✓	
Section		✓	✓		✓	✓
Municipality			✓			✓
Question					✓	✓
Number of obs.	2706	2706	2706	2706	2706	2706

$^{+}p < 0.1; ^{*}p < 0.05; ^{**}p < 0.01$

Table 5: The collaboration measure is an indicator that takes the value 1 if the bureaucrat reported collaboration on a given section. Both outcomes are standardized so coefficient magnitudes can be interpreted in standard deviation units. The encouragement treatment was assigned at the municipal level so the treatment indicator is subsumed in the municipal fixed effects in Columns (3) and (6). Respondent certainty was elicited for a randomly selected open-ended question in each section so Columns (5)-(6) include question fixed effects. Standard errors are clustered at the municipal level.

$R^2_{Y\sim D|\mathbf{X}}$, asks how much variation in collaboration (D) would an omitted confounder that explains *all* residual variation in each outcome (Y) require in order to attribute the estimate of β_1 to selection alone? While this diagnostic is generally small (ranging from <0.01 to 0.07, depending on collaboration measure and outcome), it is useful to note that among the set of covariates describing bureaucrats and municipalities (see Figure A10), the strongest single predictor of collaboration had an R^2 of 0.023 (for D_i^1), 0.05 (for D_i^2), and 0.04 (for D_i^3). The second and third diagnostics quantify the minimum proportion of residual variance that an omitted confounder would need to explain in both the collaboration measure and the outcome of interest to explain away the point estimate ($RV_{q=1}$) and the statistical significance ($RV_{q=1,\alpha=0.05}$) of the covariate-adjusted estimate of β_1 in Figure 3. With the exception of the content outcome, these robustness values are much larger than the variation in the treatment and/or outcome that can be explained using the available covariates. All three diagnostics suggest that selection is unlikely to be the only explanation for the positive association between collaboration and properties of the data.

The benchmarking exercise builds upon the dominant explanation in the extant literature on

Outcome (Index)	Collaboration measure	$R^2_{Y\sim D\vert\mathbf{X}}$	$RV_{q=1}$	$RV_{q=1,\alpha=0.05}$	Benchmark
Certainty	Extensive margin	0.03	0.16	0.10	5.77×
	Intensive margin (sections)	0.06	0.22	0.16	7.66×
	Intensive margin (entities)	0.04	0.18	0.12	7.57×
Missingness	Extensive margin	0.01	0.10	0.04	5.51×
	Intensive margin (sections)	0.03	0.15	0.09	6.64×
	Intensive margin (entities)	0.01	0.10	0.04	7.65×
Accuracy	Extensive margin	0.00	0.06	0.00	–
	Intensive margin (sections)	0.00	0.07	0.00	4.58×
	Intensive margin (entities)	0.01	0.08	0.02	6.95×
Principal components	Extensive margin	0.00	0.01	0.00	–
	Intensive margin (sections)	0.00	0.06	0.00	–
	Intensive margin (entities)	0.00	0.04	0.00	–
Duration	Extensive margin	0.01	0.08	0.01	1.82×
	Intensive margin (sections)	0.02	0.13	0.07	8.13×
	Intensive margin (entities)	0.01	0.10	0.04	3.29×
Enumerator evaluations	Extensive margin	0.02	0.14	0.09	6.15×
	Intensive margin (sections)	0.07	0.23	0.18	10.75×
	Intensive margin (entities)	0.05	0.20	0.15	7.13×

Table 6: Sensitivity diagnostics proposed by Cinelli and Hazlett (2020) for each outcome. The benchmark employs: a matrix of indicators for municipal category, the DNP's municipal performance index, and an indicator for departmental capitals.

the quality of state information: state capacity. I measure state capacity using three measures: municipal capacity (the blocking variable in the experiment), an indicator for status as a departmental capital, and the National Department of Planning's municipal performance index (as of 2021). The benchmark asks: how much stronger than these capacity measures would an omitted confounder need to be to attribute the covariate-adjusted estimates in Figure 3 to selection alone? The right column of Table 6 suggests that such a confounder (or set of confounders) would need to be substantially stronger than capacity. For all measures except content (the principal components indices) and duration, confounders would need to explain $\geq 4.5\times$ the amount of variation in *both* collaboration and the outcome of interest.

In sum, these analyses suggest that data quality increases in three measures of collaboration between bureaucrats within *alcaldías*. Qualitative interviews support the idea that this positive correlation is driven by both positive selection into collaboration and increases in information that

Covariate	Secretaries No ($n = 445$)	Secretaries Yes ($n = 456$)	Secretaries Difference	Entity leaders No ($n = 380$)	Entity leaders Yes ($n = 522$)	Entity leaders Difference
Male	0.57 (0.02)	0.7 (0.02)	0.12 (0.03)	0.55 (0.03)	0.69 (0.02)	0.14 (0.03)
Age	38.96 (0.52)	36.87 (0.44)	-2.08 (0.68)	39.31 (0.58)	36.9 (0.4)	-2.41 (0.71)
Native to muni.	0.38 (0.02)	0.29 (0.02)	-0.08 (0.03)	0.38 (0.03)	0.3 (0.02)	-0.09 (0.03)
Served $\leq$3 years	0.45 (0.02)	0.71 (0.02)	0.26 (0.03)	0.43 (0.03)	0.69 (0.02)	0.26 (0.03)

Table 7: Characteristics of secretaries versus non-secretaries (left) and unit leaders versus non-leaders (right). Robust standard errors are in parentheses

come from collaboration. While the research design cannot quantitatively decompose the correlation into selection versus treatment effects, exploiting variation in collaboration across sections of the survey (within municipality) and bounding the possible extent of selection suggest that both processes are at work.

5 The role of bureaucratic organization

The evidence to this point suggests that the DT data production involves frequent interactions between multiple bureaucrats and that these interactions improve data quality. How does the variation in the organizational structure of the *alcaldías* shape these interactions?

5.1 Hierarchy within units

I first consider the rank of bureaucratic reporters within their units. Table 7 compares several characteristics of bureaucrats who lead their units, as secretaries in secretariats (left) or secretaries or directors of any unit (right). Leaders are more likely to be male, slightly younger (on average), and less likely to be a native of the municipality they serve in. However, the largest difference is in the length of their tenure. Given that the DT data was collected at the end of the third year of all mayors' terms, Table 7 suggests that 71% of secretaries (resp. 69% of leaders) were appointed to the *alcaldía* by the current mayor, in contrast to 45% of non-secretaries (resp. 43% of non-leaders). It is useful to note that despite PS' request for bureaucrats with at least five years of experience in a given *alcaldía*, just 42% of respondents had this much experience.

Relative to "non-leaders" within bureaucratic agencies, leaders are disproportionately appointed by mayors. This churn from administration to administration means that leaders: (a) have less knowledge about/experience in the governments that they work in; and (b) must regularly be preparing for the next appointment. These sentiments were expressed regularly in the interviews. A civil servant in the Planning Secretariat of a Category 6 municipality decried the loss of information that accompanies mayoral transitions, stating that effective transitions require "specialized and long-standing personnel who know the municipality. [The continuity of these employees] guarantees that the transition process takes place more effectively thanks to the wealth of experience they bring." Recall also the bureaucrat quoted above who described secretaries needing to brush up their resumés at the end of the mayoral term.

This discussion suggests that on one hand, secretaries may know less about conditions of a municipality, increasing the possibility of learning from collaboration. On the other, it highlights the misalignment between their incentives and other (more numerous) bureaucrats, given that most entities have one leader and multiple staff members. This misalignment yields potential differences in bureaucrats' preferred reports, reducing gains from communication. Which (directional) effect dominates speaks to whether secretaries/leaders are more or less prone to collaboration when tasked with data collection.

Interviews suggest that distinct biases are frequently linked to the respondent's position within their unit. For example, a leader in a Category 2 municipality reported "we try to emphasize the positive things we have achieved and leave other things aside and answer in a more general way, because the survey allowed us to give less precise, more general, answers on those issues in which we have a high [political] vulnerability." In contrast, a civil servant with 28 years of experience in a Category 4 municipality indicated that "the local administration of [municipality] is interested in using the information reported to serve to solve 'deficiencies' and needs that affect the quality of life of its inhabitants. Therefore, [we] hope that the national government will find ways to solve the problems and allocate more resources." In the case of social development, these goals are often at odds: poor development outcomes may be costly to the reputations of incumbents but useful for procuring more resources. While explicit motives over the use of the data did not emerge in every interview, these discussions suggest varying motivations that largely correspond to individuals' ranks.

Figure 4 provides compelling evidence that secretaries (and leaders in Appendix A11) request less information than rank-and-file bureaucrats.[13] Across all measures, we observe that the entire distribution of collaboration for secretaries is shifted toward lower collaboration. This is evident from the empirical CDFs because the line for secretaries is consistently above that of non-secretaries. Indeed, across a battery of municipal-level covariates as well as bureaucrat-level characteristics, this rank-based distinction is the single most prognostic predictor of each of the three collaboration measures. In sum, there is strong evidence that secretaries or leaders engage in less collaboration than their subordinate counterparts. This is despite the fact that they have less experience in the municipalities and *alcaldías* that they serve.

One concern with this analysis is how bureaucrats were selected to report the data to PS. Assignment of the DT survey to a secretary occurred as a function of PS' identification of an initial respondent in each municipality. This process relied on existing contacts within *alcaldías* and input from regional PS offices. Comparison of the name of the ultimate respondent to the name of the invitee reveals that the initial invitee served as the reporter in 93.6% of municipalities (that completed DT data collection). Where the initial invitee did not respond, the survey was delegated to a lower-level bureaucrat in 57% of cases.[14] Importantly, the difference between secretaries and non-secretaries is very similar in subsamples of original respondents and replacement respondents (Figure A12). This suggests that the method through which PS selected reporters in each munici-

[13]The rest of this section compares secretaries to non-secretaries. Results are very similar for the slightly more general leader vs. non-leader categorization, which is more encompassing but less precise.

[14]This is substantially higher than the title/role misclassification rate of 5.2% of units among invited bureaucrats who answered the survey.

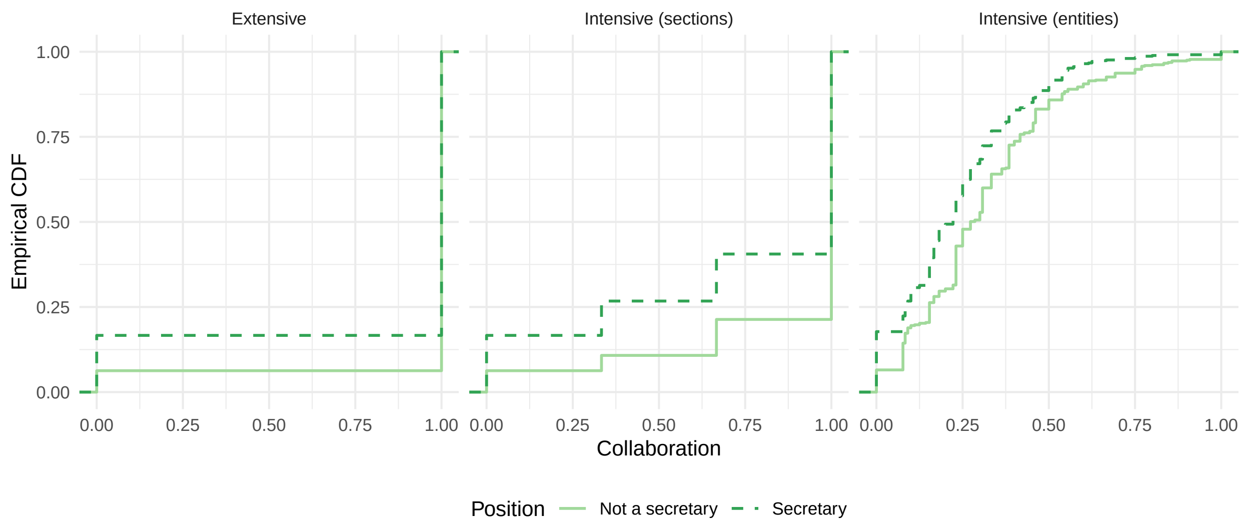

Figure 4: Secretaries (and leaders) exhibited lower levels of collaboration, as measured with all three measures. Each line depicts the empirical cumulative distribution function (ECDF) of a collaboration measure among a subgroup of respondents.

pality is unlikely to be the source of this difference between secretaries and non-secretaries.

To what extent do these different propensities to collaborate affect data quality? This question depends on whether collaboration produces the same effects among both secretaries and non-secretaries. Figure 5 reports subgroup-specific estimates of β_1 (in (2) and (3)) for each outcome index. It shows that, with one exception, the marginal effects of collaboration share the same sign in both subgroups. Moreover, in most cases, it is not possible to reject a null that the estimated coefficients are the same. These estimates support the idea that collaboration enhances accuracy and thus the quality of data outputs.[15] While collaboration should increase quality measures for both subgroups, it should not necessarily move the content of reported data in the same direction. Given the orientation of the principal components, higher values of β_1 correlated with lower reported development levels and lower values of β_1 correlate with higher reported levels of development. The pink points—for the principal components index—are consistently above the 45-degree line and take different signs with the extensive margin indicator. This is consistent with qualitative anec-

[15]The theoretical model does not make further predictions about the relative magnitudes of these increases in accuracy.

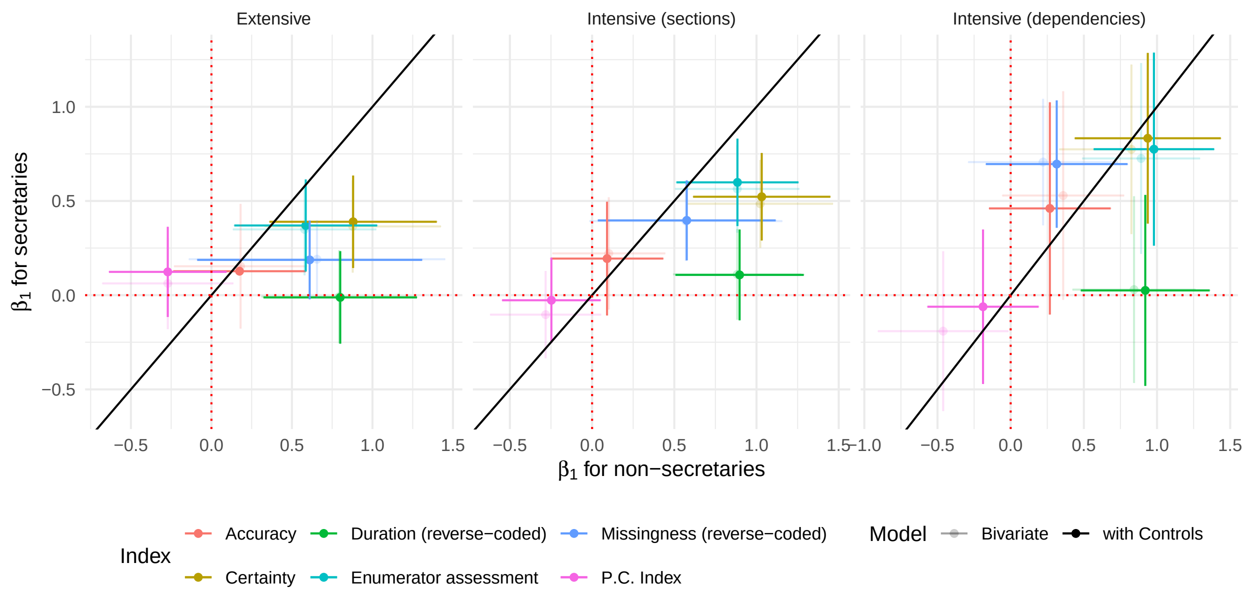

Figure 5: Estimates of β_1 (2) (transparent) and (3) (solid) among non-secretaries (x-axis) and secretaries (y-axis). The 45-degree line indicates that the estimates among each subgroup are equivalent. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors

dotes of non-secretaries (civil servants) trying to draw investment while secretaries use information they collect to tout achievements (or, at least not exaggerate need). However, these estimates are quite noisy and are not consistently significant at conventional levels. As in the previous section, these estimates reflect some combination of selection and treatment effects. Nevertheless, these figures support the idea that selection into collaboration that varies with *rank within a unit* yields uneven impacts on data quality.

5.2 Organization of units

The organization of units in the same *alcaldía* may also affect bureaucrats' choice of whether to collaborate. I argued that within hierarchical organizations, communication between dependencies is likely more routinized as part of vertical information flows up the organization. This practice and experience implies that collaboration should impose lower cost on bureaucrats. Table 8 reveals that in more hierarchically-organized *alcaldías*, collaboration is more common, at least along the extensive margin. Specifically, hierarchical organization corresponds to a 6.8 percentage point increase

	Extensive margin			Intensive (sections)			Intensive (dependencies)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hierarchical alcaldía	0.061*	0.065**	0.068**	0.047^{+}	0.052^{+}	0.055^{+}	0.027	0.030	0.029
	(0.024)	(0.024)	(0.024)	(0.028)	(0.028)	(0.028)	(0.019)	(0.019)	(0.019)
Number of secretariats (std.)	0.010	-0.002	-0.010	0.017	0.002	-0.010	0.029***	0.022**	0.019^{+}
	(0.011)	(0.011)	(0.015)	(0.011)	(0.011)	(0.016)	(0.008)	(0.008)	(0.011)
Secretary indicator		✓	✓		✓	✓		✓	✓
Benchmark covariates			✓			✓			✓
Outcome mean	0.889	0.889	0.889	0.801	0.801	0.801	0.277	0.277	0.277
Outcome standard dev.	0.314	0.314	0.314	0.340	0.340	0.340	0.208	0.208	0.208
Outcome range	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
Number of Obs.	828	827	827	828	827	827	828	827	827
R^2	0.006	0.030	0.035	0.005	0.047	0.050	0.022	0.045	0.055

Table 8: Association between two measures of the organization of units/dependencies within local *alcaldías* and selection into collaboration. Heteroskedasticity-robust standard errors in parentheses.

in the probability of any collaboration (Column 3). This is equivalent to 0.22 standard deviations of the outcome variable. These associations weaken for the intensive margin indicators, especially when considering the share of dependencies consulted. One hypothesis for this attenuation is that hierarchy promotes collaboration within a branch of the hierarchy but not more widely across the organization. As such, if collaboration remains siloed within specific units, we would not expect as much movement on this intensive margin indicator. In sum, these findings offer qualified support for the idea that hierarchy across dependencies facilitates agent-to-agent communication.

The number of secretariats, a second measure derived from the organigram data, exhibits a near-zero conditional association with the extensive or first intensive margin indicator of collaboration. However, a one standard deviation increase in the number of secretariats (2.8 secretariats) is associated with a smaller (if more precisely estimated) increase in the share of dependencies consulted.

6 Conclusion

In this paper, I argue that agent-to-agent collaboration between bureaucrats shapes the production and quality of state data. Moreover, the organizational structure of bureaucracies predicts selection into collaboration. Two implications of these findings are important. First, facilitating collaboration and communication between bureaucrats may lead to improvements in the administrative data that states use to govern. Second, more targeted delegation of information collection to longer-serving bureaucrats in the middle of bureaucratic hierarchies may be one approach that states could take to achieve this end.

This paper studies agent-to-agent collaboration in the context of a data production process in which such collaboration should be especially valuable: a long, cross-sector survey assigned to a single bureaucrat in each *alcaldía*. Such surveys from national to local governments are common—measuring public finances, governance, infrastructure, and/or capacity (e.g., the US Census of Governments, Mexico's National Census of Local Governments, or Brazil's Survey of Basic Municipal Information). Within-sector data requests to local governments, schools, or health departments are similarly lengthy and technical, which may well induce respondents to draw on colleagues' expertise. The present case is thus useful for directly measuring data production behavior because of the partnership with PS, but the collaboration dynamics I characterize are likely present in other contexts.

This paper opens a number of avenues for future investigation. Specifically, it shows a need for a renewed focus on bureaucratic organization as a determinant of bureaucratic outputs. The collection and coding of organigram data from Colombian *alcaldías* proposes one way to systematically measure and code these types of organizations. Similar source materials are available for many other bureaucratic organizations globally. It further directs our attention to the behavior and outputs of mid-level bureaucrats who have largely been ignored in favor of political appointees at the highest levels of US politics and lower-level frontline service providers in comparative politics.

References

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2020. "The lay of the land: Information capacity and the modern state." *Comparative Political Studies* 53 (2): 175–213.

Brehm, John, and Scott Gates. 1997. *Working, Shirking, and Sabotage: Bureaucratic Response to a Public Debate*. Ann Arbor, MI: University of Michigan Press.

Brierley, Sarah, Kenneth Lowande, Rachel Augustine Potter, and Guillermo Toral. 2023. "Bureaucratic Politics: Blind Spots and Opportunities in Political Science." *Annual Review of Political Science* 26: 271–290.

Bueno, Natália S. 2021. "The Timing of Public Policies: Political Budget Cycles and Credit Claiming." *American Journal of Political Science* 67 (4): 996–1011.

Camacho, Adriana, and Emily Conover. 2011. "Manipulation of Social Program Eligibility." *American Economic Journal: Economic Policy* 3 (2): 41–65.

Carbonell-Nicolau, Oriol. 2024. "Measuring Hierarchy." Working paper, Rutgers University.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton, NJ: Princeton University Press.

Carpenter, Daniel P., and George A. Krause. 2012. "Reputation and Public Administration." *Public Administration Review* 72 (1): 26–32.

Cinelli, Carlos, and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society B* 82 (1): 39–67.

Cook, Scott J, and David Fortunato. 2022. "The Politics of Police Data: State Legislative Capacity and the Transparency of State and Substate Agencies." *American Political Science Review* pp. 1–16.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–1451.

Doherty, Kathleen M., David E. Lewis, and Scott Limbocker. 2019. "Executive Control and Turnover in the Senior Executive Service." *Journal of Public Administration Research and Theory* 29 (2): 159–174.

Eckhouse, Laurel. 2022. "Metrics Management and Bureaucratic Accountability: Evidence from Policing." *American Journal of Political Science* 66 (2): 385–401.

Falleti, Tulia G. 2005. "A Sequential Theory of Decentralization: Latin American Cases in Comparative Perspective." *American Political Science Review* 99 (3): 327–346.

Fizbein, Ariel. 1997. "The Emergence of Local Capacity: Lessons From Colombia." *World Development* 25 (7): 1029–1043.

Freeman, Will. 2023. "Colombia Tries a Transformative Left Turn." *Current History* 122 (841): 69–74.

Gailmard, Sean, and John W. Patty. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science* 51 (4): 873–889.

Gailmard, Sean, and John W Patty. 2012. "Formal models of bureaucracy." *Annual Review of Political Science* 15: 353–377.

Garbiras-Díaz, Natalia, and Tara Slough. 2026. "Government Oversight and Inter-Institutional Legibility: Evidence from Colombia,." *Journal of Politics* Forthcoming.

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

Glas, Aarie. 2021. "Positionality, Power, and Positions of Power: Reflexivity in Elite Interviewing." *PS: Political Science & Politics* 54 (3): 438–442.

Grimmer, Justin, Solomon Messing, and Sean J. Westwood. 2012. "How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation." *How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation* 106 (4): 703–719.

Grossman, Guy, and Tara Slough. 2022. "Government Responsiveness in Developing Countries." *Annual Review of Political Science* Forthcoming.

Hammond, Thomas H. 1986. "Agenda Control, Organizational Structure, and Bureaucratic Politics." *American Journal of Political Science* 29 (1): 1–28.

Herrera, Helios, Ernesto Reuben, and Michael M. Ting. 2017. "Turf Wars." *Journal of Public Economics* 152: 143–153.

Hollibaugh, Garry E. 2019. *Oxford Research Enclopedia of Politics*. New York, NY: Oxford University Press chapter How Effecive are Political Appointees?

Jerven, Morten. 2013. *Poor Numbers: How We are Misled by African Development Statistics and What to Do About It*. Ithaca, NY: Cornell University Press.

Kanter, Genevieve P., and Daniel Carpenter. 2023. "The Revolving Door in Health Care Regulation." *Health Affairs* 42 (9): 1298–1303.

Lee, Melissa M, and Nan Zhang. 2017. "Legibility and the informational foundations of state capacity." *The Journal of Politics* 79 (1): 118–132.

Luechinger, Simon, and Christoph Moser. 2014. "The Value of the Revolving Door: Political Appointees and the Stock Market." *Journal of Public Economics* 119: 93–107.

Martínez, Luis R. 2021. "How Much Should We Trust the Dictator's GDP Growth Estimates?" *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-78).

Martínez, Luis R., Erik Snowberg, and Michael M. Ting. 2024. "The Anatomy of Colombian Bureaucratic Hierarchies." CAF Working Paper #2024/05.

Radner, Roy. 1992. "Hierarchy: The Economics of Managing." *Journal of Economic Literature* (1382-1415).

Radner, Roy. 1993. "The Organization of Decentralized Information Processing." *Econometrica* 61 (5): 1109–1146.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106 (6): 467–482.

Simon, Herbert A. 1973. "Applying Information Technology to Organizational Design." *Public Administration Review* 33 (3): 268–278.

Slough, Tara. 2026. "Bureaucratic Incentives and Data Production: Evidence from Social Registries." Working paper, New York University.

Soedirgo, Jessica, and Aarie Glas. 2020. "Toward Active Reflexivity: Positionality and Practice in the Production of Knowledge." *PS: Political Science & Politics* 53 (3): 527–531.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20 (4): 518–529.

Tirole, Jean. 1987. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, and Organization* 2 (2): 181–214.

Toral, Guillermo. 2024. "Street-Level Rule of Law: Prosecutor Presence and the Fight against Corruption." Working paper available at `https://www.guillermotoral.com/prosecutor_presence.pdf`.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.

Zacka, Bernardo. 2017. *When the State Meets the Street: Public Service and Moral Agency*. Cambridge, MA: Harvard University Press.