Bureaucratic Incentives and Data Production: Evidence from Social Registries

Tara Slough*

January 12, 2026

Abstract

One important but often overlooked task of bureaucrats is producing state data. When central governments depend on information from local governments to allocate resources, strategic interactions between local politicians and bureaucrats shape data fed to the central government. This study examines such agency problems in Colombia and Brazil, by studying the social registries used to determine eligibility for means-tested transfers. Using original survey data from Colombian bureaucrats, matched employer-employee records from Brazil, and social registry microdata from both countries, I analyze how mayors' selection and oversight of bureaucrats affect data quality. Findings show that mayors more closely monitor bureaucrats they appoint rather than retain. Appointed bureaucrats, who more frequently share policy goals of the politician, exert more effort and report more poor households, expanding program eligibility. However, mayors must trade off this loyalty with the loss of expertise when making replacing these administrators. While the distortions in registry data are modest, they meaningfully shape the distribution of anti-poverty transfers.

*Associate Professor, New York University, tara.slough@nyu.edu. Thanks to Karen Rojas, Carolina Torreblanca, and Meghna Yadav for expert research assistance. I am grateful to Scott Abramson, Akshay Dixit, Anderson Frey, Dorothy Kronick, Carolina Torreblanca, Clare Tyson, and audiences at the World Bank, the 2nd IE-Rochester Workshop in Political Economy and the Harris School for excellent comments. This research was supported by a grant from the IPA Research Methods Initiative. I thank Anderson Frey for anonymizing and providing access to the CadUnico microdata.

The creation and maintenance of citizen- or household-level administrative data is a comparatively recent phenomenon. State land cadasters—registering the properties of individuals and families—were not widely adopted until the sixteenth through eighteenth centuries (Kain and Baigent, 1992). Systematic national vital statistics—records of individual births, deaths, marriages, and divorces—emerged only in the nineteenth and twentieth centuries in Europe and the United States, respectively Hetzel (1997). The social welfare and safety-net programs adopted in the twentieth and twenty-first centuries necessitated the development of additional types of social registers on individuals and households. Central to the development and collection of each of these forms of administrative microdata is the relationship between the (central) government and agents across the territory, often located within state or local governments.

The creation and maintenance of of social registers that determine access to means-tested social programs is challenging task for national governments. Impoverished populations who are most likely to qualify for and benefit from social programs are often harder for governments to reach (e.g., Scott, 1998; Lee and Zhang, 2016; Bowles, 2020). To address these challenges, register maintenance is generally delegated from the central government to agents of state and/or local governments who have more direct access to potential enrollees (Garbiras-Díaz and Slough, 2025). But given the link between the content of social registers and direct benefits from the central governments, states and localities can draw additional benefits to their localities by expanding rolls (Camacho and Conover, 2011).

Open questions about the resultant quality or legibility of these data are not confined to middling- or low-capacity states. For example, in the US, the federal government does not maintain a register of household-level data on who receives Supplemental Nutrition Assistance Program (SNAP) benefits, a program $122 billion dollar benefit in FY 2024. These records are instead collected in a non-systematic fashion by state governments. Indeed, recent efforts by the US Department of Government Efficiency (DOGE) aimed to pry these rolls—including sensitive information about recipients—from state administrators (Joffe-Block and Fowler, 2025).

I study how a politician's selection of a bureaucrat and their subsequent oversight strategy affect bureaucratic effort and outputs, namely the quality of social registers and consequent access to social programs. I propose a model in which a politician values maximizing enrollment in the social program, regardless of a household's eligibility. There exist two types of bureaucrats: technocrats, who value accurate targeting of the social program, and loyalists who, like the politician, seek to maximize enrollment. The bureaucrat's type is unobserved by the politician, though the politician knows that they are more likely to have a loyalist if they can appoint someone they know (i.e., a campaign worker or donor). The cost of replacing a register administrator is the loss of program-specific knowledge. This tradeoff—more closely aligned preferences at the cost of programmatic expertise—also affect politician's attempt to extract effort from the bureaucrat by choosing an oversight strategy. Increased effort by bureaucrats expands the number of households surveyed and, in the case of loyalists, may inflate the proportion of households that qualify for transfers.

The model clarifies several features of the environment that inform interpretation of the empirical results. First, selection and monitoring can be complements: politicians monitor bureaucrats that they have appointed more stringently because there are greater gains from a loyalist's effort than a technocrat's effort. Second, while a newly-appointed loyalist will always outwork a new technocrat, the policy (or public service motivation) of technocrats leads to greater gains in effort over the course of their tenure. This means that there exists an empirical question about whether newly appointed bureaucrats or retained bureaucrats (who are marginally less likely to be loyalists) work harder in equilibrium. When newly appointed bureaucrats exert greater effort, we should see new appointees classify more households as poor. Within this environment, a politician's strategic decision whether to retain the bureaucrat will tend to *attenuate* or understate differences in monitoring rates, bureaucrat effort, and bureaucratic classification of households relative to an all-else-equal environment with random retention of bureaucrats.

Empirically, I examine these empirical questions in the context the two largest social registries

in South America: Brazil's Cadastro Único para Programas Sociais (CadUnico) and Colombia's Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (SISBÉN). To examine the relationship between bureaucratic selection and oversight, I rely on an original survey of SISBÉN administrators (one per municipality). Among the $n = 752$ respondents (69% response rate), bureaucrats appointed by the mayor report higher rates of oversight and greater goal setting by their superiors. This provides evidence consistent with a complementarity between bureaucratic selection and oversight effort. With distinct research designs based on the structure of available data from both Colombia and Brazil, I find evidence appointed bureaucrats exert more effort then their retained counterparts. In Colombia, appointees expand the number of households surveyed at a greater rate after the introduction of a new poverty scoring system. In Brazil, individual bureaucrats appointed by the mayor conduct more household interviews in the same municipality and time period, even after adjustment for individual (bureaucrat) characteristics. Remarkably, this increase in effort is quite similar in both contexts: municipalities with appointed SISBÉN administrators increase the rolls by about 12% relative to municipalities with retained administrators. Averaging over annual figures in Brazil suggests that appointed bureaucrats increase the rolls by about 11% relative to their retained counterparts in the same municipality.

Moving from equilibrium strategies to equilibrium outcomes, I examine the scores reported in the social registers. Consistent with the implications of an environment in which new appointees exert greater effort, we observe that new appointees register systematically poorer households than their retained counterparts. In Colombia, this results in systematic shift toward poorer classifications among interviewed households in a given municipality. In Brazil, this corresponds to more households below the conditional income threshold that is a necessary condition for to qualify for the conditional cash transfer. Moreover, within-bureaucrat comparisons over time in the Brazilian data reveal that, consistent with model implications, individual bureaucrats report increasing shares qualifying households as they learn the ropes of the job.

This paper makes three contributions. First, it provides a new perspective on the political conse-

quences of (largely) programmatic means-tested transfers. Existing work examines how *politicians* manipulate program design, rollout, and publicity to gain votes from these programs (de la O, 2013; Zucco, 2013; Manacorda, Miguel, and Vigorito, 2011; Imai, King, and Velasco-Rivera, 2020,?; Camacho and Conover, 2011). This work shows how the implementation of means-testing by bureaucrats affects politicians' ability to use these programs as a vehicle for electoral gain. While local principals have the incentive to stuff the rolls, as documented by Camacho and Conover (2011),[1] their ability to do so is constrained by agency problems with their bureaucrats. In this sense, shirking by bureaucrats insulates the national government (to some degree) from these efforts by local politicians. These efforts limit the ability of local politicians to use these programs to win votes, which may speak to the mixed effects of these programs on support for local mayors (Labonne, 2013).

With respect to the incentives within local governments, I show how selection and oversight can be effectively used in tandem to influence bureaucratic outputs. Most work considers one of these two strategies in isolation. Literature on patronage appointments and bureaucratic transfers suggests that politicians benefit from appointing individuals with aligned preferences or in exchange for past support (Colonnelli, Prem, and TEso, 2020; Akhtari, Moreira, and Trucco, 2022; Toral, 2024*b*). But the question of bureaucratic selection is often divorced from work on bureaucratic oversight or monitoring (Gulzar and Pasquale, 2017; Raffler, 2022; Brierly, 2020; Slough, 2024). As a result, we do not know how use of these two strategies by a principal covary. I show that mayors may oversee appointed bureaucrats more intensively because the marginal return to the bureaucrat's effort is greater when preferences are more closely aligned. This mechanism is distinct from existing arguments about the how patronage facilitates enhanced monitoring through greater information (Toral, 2024*a*). In the context of social registries, thus, stronger elicitation of effort from appointed bureaucrats yields greater transfers from the national government to constituents.

[1] In addition, see Brollo, Kaufmann, and La Ferrara (2020) on how local politicians manipulate enforcement of conditionalities to increase consumption in their municipalities.

This increases the consumption of constituents at the cost of the accuracy of the means-testing scheme and increases costs to the national government.

Finally, this paper proposes a new link between social policy and state information production. I show that emergence of means-tested social programs in Latin America as part of Latin America's "revolution in social policy" placed stringent new demands on local governments de la O (2015); Garay (2017); Niedzwiecki (2018). Specifically, it increased substantially the type and quantity of data that local government were expected to produce for the center (Garbiras-Díaz and Slough, 2023). But this data production is not mechanical. Strategic relationships between local politicians and bureaucrats leave "footprints" in data quality, and hence household eligibility for social programs. This explanation is distinct from existing explanations for variation in the quality of state data that are premised on limited state capacity (Jerven, 2013; Lee and Zhang, 2017; Brambor et al., 2020; Angrist, Goldberg, and Jolliffe, 2021) or incentives for distortion within autocratic regimes (Martínez, 2021; Wallace, 2016; Guriev and Treisman, 2019; Lorentzen, 2014; Edmond, 2013; Trinh, 2021).

1 Theory

I propose a simple model to examine how agency problems within local governments affect the quality of social registries, and hence access to targeted social welfare programs. The model posits a number of empirical implications for data quality, which ultimately translate into variation in the distribution of transfers from the central government to households residing in different constituencies.

Social policy design In a given constituency, e.g., a municipality, there exists a unit mass of households. Households, indexed by i, are characterized by a some measure of income, assets, or consumption, a_i. Denote the cumulative density function of a_i as $F(\cdot)$. The national government exogenously determines the measure of a_i and sets some threshhold in a_i, $\widehat{a} < \max a_i$, which denotes eligibility for a transfer or social program. Specifically, all households for whom $a_i \leq$

	Transfer granted ($T_i = 1$)	Transfer not granted ($T_i = 0$)
Eligible ($a_i \leq \widehat{a}$)	Correct allocation $s_i \leq \widehat{a}$	Type-II error $s_i \in \{(\widehat{a}, 1] \cup \emptyset\}$
Ineligible ($a_i > \widehat{a}$)	Type-I error $s_i \leq \widehat{a}$	Correct denial $s_i \in \{(\widehat{a}, 1] \cup \emptyset\}$

Table 1: Possible relationships between latent eligibility (a_i) and transfer outcomes (T_i). "Correct" is defined relative to the central government's targeting policy, $\widehat{a}$.

$\widehat{a}$, are *eligible* for the program. Within the constituency, the share of the eligible households is therefore $F(\widehat{a}) \in [0, 1)$.

The constituency government is tasked with measuring and reporting a_i to the national government. Measurement consists of a determination of: (1) which households to measure; and (2) a score $s_i \in [0, 1]$ for each measured household. (Let $s_i = \emptyset$ denote the unmeasured score for unmeasured households.) If a measured score is less than or equivalent to $\widehat{a}$, the household is granted the transfer, $T_i = 1$:

$$T_i = \mathbb{I}[s_i \leq \widehat{a}]$$

This yields the combinations of (latent) eligibility and realized transfer allocations reported in Table 1. The table distinguishes between correct determinations—from the perspective of the national government's policy—and two types of errors: Type-I errors of inclusion and Type-II errors of exclusion. Specifically, the table shows that that Type-II errors can be caused by failure to score an eligible household ($s_i = \emptyset$) or inaccurate scoring of the household in which s_i is sufficiently larger than the latent a_i, such that $a_i \leq \widehat{a} < s_i$. In contrast, Type-I errors are caused exclusively by inaccurate scoring in which $s_i \leq \widehat{a} < a_i$.

Ultimately the definition of eligibility and the size/targeting of transfers are policy decisions by the central government. There can be tremendous variation in the design of such policies and their consequences for welfare. In contrast to a large literature on the design of these policies (Coady, Grosh, and Hoddinnott, 204; Alatas et al., 2012; Hanna and Olken, 2018: e.g.,), my focus is their

implementation. My focus is therefore on the goals and actions of local governments who operate within the social policy environment set by the central government.

Specifically, central governments rely on local governments to score households. This is a time- and effort-intensive task which amounts to gathering data on households and individuals within the municipality. Where large shares of the population are eligible for a transfer (i.e., $F(\widehat{a})$ is large) or benefits are particularly generous, this can amount to meetings with or visits to a large share of households.

The principal-agent problem in local governments: Local governments consist of an elected politician and a bureaucratic agent. Local bureaucrats are tasked with producing and maintaining the data from social registries (Frey and Santarrosa, 2024; Slough, 2022; Camacho, Conover, and Querubín, forthcoming).

A substantial literature posits that local politicians care about the data submitted to social registries. Specifically, by increasing access to social programs funded by the central government, local politicians can direct resources to their constituencies and potentially claim credit for these inflows (Camacho, Conover, and Querubín, forthcoming; Bueno, 2021). Within the classification in Table 1, this implies that politicians value correct allocations and false positives, or errors of inclusion. Frey (2012) further suggests that errors of inclusion may be especially valuable when households know that they should not have been scored as eligible but receive the transfer, they become dependent on the local government to maintain this misclassification to continue receiving benefits. Both true and false positives draw increased resources to the municipality. The implication here, is that there is a preference for (1) inducing bureaucrats to score more households; and (2) reporting of lower scores, holding fixed the given attributes of a household.

In pursuit of these policy goals, a politician makes two choices with respect to the bureaucrat. They *select* a bureaucrat to maintain the register by choosing whether to retain the "incumbent" bureaucrat or appoint someone new. I denote retention of the incumbent bureaucrat as $r \in \{0, 1\}$, where $r = 1$ signifies a decision to retain the incumbent. They then commit to an *oversight strategy*

to be imposed upon the bureaucrat. This oversight strategy can be interpreted as a product of the (a) rate of monitoring and (b) the severity of punishments imposed when the bureaucrat shirks. Civil service systems or public employment regulation constrain both bureaucratic selection and oversight to some degree. However, I contend that both decisions are choices of most politicians. Even if a politician cannot remove an incumbent register adminstrator from their bureaucracy, they are likely able to transfer that individual to another job. Similarly, even in the absence of an ability to fire a wayward administrator, they can engage in more or less oversight of the employee.

Bureaucrats value the social policy outcomes, but do so in different ways. Specifically, consider two types of bureaucrats: a technocrat (T) and a loyalist (L). The technocrat values maximizing correct allocations and denials. In this sense, her preferences over the transfer mirror those of the national government or the social policy designer.[2] The loyalist's preferences over the policy mirror those of the local politician: they value expanding the rolls through some combination of correct allocations and false positives. A bureaucrat knows her type but the principal does not. The principal knows only the share of loyalists, $\pi_t \in [0, 1]$, and thus the share of technocrats, $1 - \pi_t$.

In addition to preferences over policy outcomes, bureaucrats must exert effort to locate, survey, and report the assets of citizens. Bureaucrats choose to exert effort $e \in [0, 1]$ at cost $ce^2/2$, where $c > \frac{1+\sqrt{5}}{2}$.[3] One can interpret e as the share of citizens surveyed. To induce the bureaucrat to exert effort, the local politician monitors and punishes a lack of effort with an expected penalty $m > 0$. By expected penalty, m captures product of the rate of monitoring and the extent of penalties imposed.

While effort is required to enter a household in the register, registration is not alone sufficient to ensure that a household qualifies for the transfer. The bureaucrat must score a household or

[2]By referring to the bureaucrat as a "technocrat," I do not imply that the national government's targeting policy is "technocratic" or non-discretionary. I simply mean that the bureaucrat seeks to implement the national government's chosen policy.

[3]This lower bound on c is sufficient to ensure that, consistent with empirical observation, no bureaucrat scores every household in their jurisdiction.

provide a battery of assets. This generally requires some expertise, both in terms of knowledge of the population and knowledge of the instrument. This is represented the parameters $\sigma_t^{\mathcal{E}} \in [1/2, 1]$ and $\sigma_t^{\mathcal{I}} \in [1/2, 1]$, which are defined as follows:

$$\sigma_t^{\mathcal{E}} = \Pr(s_i < \widehat{a} \mid a_i < \widehat{a})$$
$$\sigma_t^{\mathcal{I}} = \Pr(s_i > \widehat{a} \mid a_i > \widehat{a})$$

Both parameters are subscripted by t, which indexes the bureaucrat's tenure. A bureaucrat who is retained by the politician ($t = 2$) has greater knowledge of the program such that $\sigma_2^{\mathcal{E}} = \sigma_1^{\mathcal{E}} + \delta$ and $\sigma_2^{\mathcal{I}} = \sigma_1^{\mathcal{I}} + \delta$, where $\delta \in (0, 1 - \max\{\sigma_1^{\mathcal{E}}, \sigma_1^{\mathcal{I}}\})$. This parametric assumption captures (in reduced form) that bureaucrats learn from experience over the course of their tenure in a job.

Because the loyalist type also values errors that grant the transfer, consider one additional measure of expertise: the ability to manipulate or "fudge" data such that ineligible households might be classified as transfer-eligible. Let $\sigma_t^{\mathcal{M}} \in [0, \sigma_t^{\mathcal{E}} + \sigma_t^{\mathcal{I}} - 1]$ represent a newly-appointed loyalist's ability to manipulate reported eligibility.[4] As above, I will assume that $\sigma_2^{\mathcal{M}} = \sigma_1^{\mathcal{M}} + \delta$, since the bureaucrat can also learn how to misclassify households to increase the share of households that access benefits. Thus, the objective of each type of bureaucrat is given by:

$$E[U_T(e; \sigma_t^{\mathcal{E}}, \sigma_t^{\mathcal{I}})] = F(\widehat{a}) \underbrace{e\sigma_t^{\mathcal{E}}}_{\substack{\text{Correct} \\ \text{allocation}}} + (1 - \widehat{F}(a)) \underbrace{(1 - e + e\sigma_t^{\mathcal{I}})}_{\text{Correct denial}} - (1 - e)m - \tfrac{ce^2}{2} \quad (1)$$

$$E[U_L(e; \sigma_t^{\mathcal{E}}, \sigma_t^{\mathcal{I}}, \sigma_t^{\mathcal{M}})] = F(\widehat{a}) \underbrace{e\sigma_t^{\mathcal{E}}}_{\substack{\text{Correct} \\ \text{allocation}}} + (1 - \widehat{F}(a)) \underbrace{e(1 - \sigma_t^{\mathcal{I}} + \sigma_t^{\mathcal{M}})}_{\text{False positive}} - (1 - e)m - \tfrac{ce^2}{2} \quad (2)$$

These objectives encode a number of assumptions. First, the technocrat and loyalist vary only in their preferences over policy outcomes, not in their knowledge/competence or their cost of effort. This is desirable insofar as it does not "bake in" additional differences between the two types of

[4] By constraining $\sigma_t^{\mathcal{M}} \leq \sigma_t^{\mathcal{E}} + \sigma_t^{\mathcal{I}} - 1$, I ensure that eligible households are at least as likely to qualify for the program as ineligible households, since $\sigma_t^{\mathcal{E}} \geq 1 - \sigma_t^{\mathcal{I}} + \sigma_t^{\mathcal{M}}$.

bureaucrat. Second, I assume that monitoring simply incentivizes effort, rather than attention to specific types of households (i.e., eligible or ineligible). This modeling choice has two benefits. Practically, it reflects the fact that qualification for social programs is often rendered by a higher level (national) authority and this process takes time. Thus, micromanaging a bureaucrat's allocation of effort or the designation of individual households is difficult. Intellectually, it helps to clarify which strategic forces generate variation in access to social programs.

In addition to changes in bureaucratic knowledge, politicians face different pools of bureaucrats. Specifically, I assume that the pool of new appointees is weakly more likely to be loyalists (to the current politician) than the pool of incumbents willing to continue to serve, i.e., $\pi_2 = \pi_1 - \rho$, where $\rho \in (0, \pi_1)$. Studies of patronage routinely find that politicians appoint individuals who have previously shown support (Colonnelli, Prem, and TEso, 2020). Presumably one source of support is aligned preferences. Further, to the extent that incumbent registrar administrators are loyalists to a politician's (possibly unaligned) predecessor, they are unlikely to function as loyalists to an unaligned successor. Finally, efforts by the national government to train local registrars aim to increase their knowledge while inculcating preferences for accuracy. One can interpret ρ as a measure of the level of patronage in bureaucratic staffing, i.e., the degree to which the pool of bureaucrats changes from administration to administration.

With regard to the policy, the politician seeks to maximize the number of registrants that qualify for the transfer, i.e., correct allocations and false positives. They monitor to incentivize bureaucratic effort, but monitoring is costly.

$$\begin{aligned} E[U_P(m_t; \pi_t, \sigma_t^{\mathcal{E}}, \sigma_t^{\mathcal{I}}, \sigma_t^{\mathcal{M}})] = \pi_t \underbrace{e_L \left[F(\widehat{a})\sigma^{\mathcal{E}} + (1 - F(\widehat{a}))(1 - \sigma^{\mathcal{I}} + \sigma^{\mathcal{M}})\right]}_{\text{Loyalist}} + \\ (1 - \pi_t) \underbrace{e_T \left[F(\widehat{a})\sigma^{\mathcal{E}} + (1 - F(\widehat{a}))(1 - \sigma^{\mathcal{I}})\right]}_{\text{Technocrat}} - \frac{m^2}{2} \end{aligned} \tag{3}$$

The sequence of the game is as follows:

1. The politician decides whether to retain the past bureaucrat or appoint a new bureaucrat.

2. The politician commits to monitoring rate, m.

3. The bureaucrat exerts effort, e, to score households.

4. The transfer is allocated on the basis of the bureaucrat's scores.

Equilibrium behavior: I characterize the subgame perfect Nash Equilibrium of the game. This equilibrium is a mapping of the politician's retention decision, $r \in \{0, 1\}$, to their monitoring strategy $m : \{0, 1\} \to \mathbb{R}_+$ to the bureaucrat's effort $e : \{0, 1\} \times \mathbb{R}_+ \to [0, 1]$. The model is solved by backward induction.

Consider first the bureaucrat's equilibrium effort allocation. Following (2) the technocrat's optimal effort is:

$$e_t^{T*} = \max\{0, \frac{F(\widehat{a})(1 + \sigma_t^{\mathcal{E}} - \sigma_t^{\mathcal{I}}) + \sigma_t^{\mathcal{I}} + m - 1}{c}\}$$

The loyalist's optimal effort is:

$$e_t^{L*} = \frac{F(\widehat{a})(\sigma_t^{\mathcal{E}} + \sigma_t^{\mathcal{I}} - \sigma_t^{\mathcal{M}} - 1) + 1 - \sigma_t^{\mathcal{I}} + \sigma_t^{\mathcal{M}} + m}{c}$$

Comparing these quantities, it is straightforward to see that $e_t^{L*} - e_t^{T*} > 0$, which simplifies to:

$$(1 - F(\widehat{a}))(2(1 - \sigma_t^{\mathcal{I}}) + \sigma_t^{\mathcal{M}}) > 0,$$

when the technocrat exerts effort (and is straightforward otherwise). This shows that all else equal, the loyalist exerts more effort than the technocrat. False positives—which are valued by the loyalist but not the technocrat—require effort, whereas correct denials—which are valued by the technocrat but not the loyalist—do not necessarily require effort, since a correct denial could stem from not being scored. As a consequence, the loyalist is more motivated to exert effort. Note further that

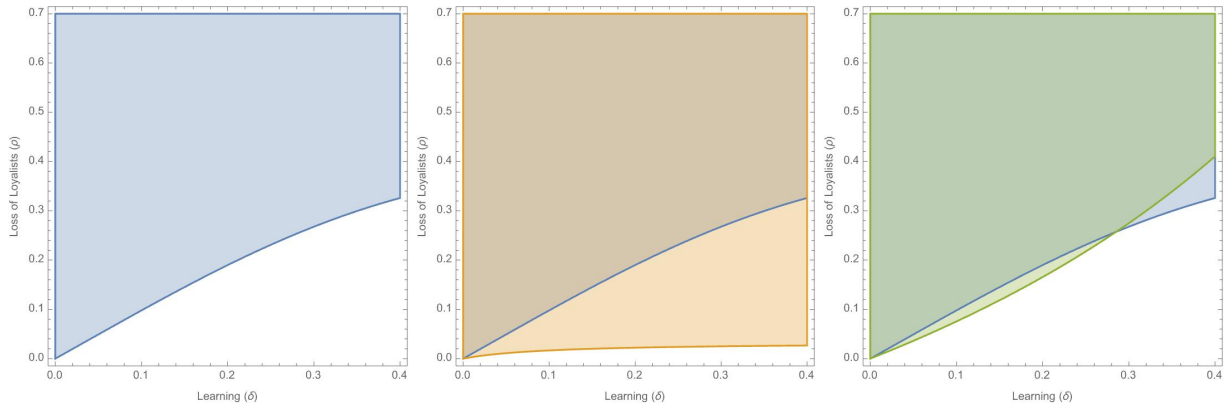
the effort of both types of bureaucrats is increasing in their ability to accurately classify eligible households ($\sigma_t^{\mathcal{E}}$). The technocrat's effort is also increasing in her knowledge of how to classify ineligible households ($\sigma_t^{\mathcal{I}}$), whereas the loyalist's effort is decreasing in this type of knowledge. This implies that learning induces different incentives for each type to exert effort.

Moving to the politician's monitoring strategy, plugging e_t^{T*} and e_t^{L*} into (3) and maximizing yields:

$$m_t^* = \begin{cases} \frac{1-\sigma_t^{\mathcal{I}}+\sigma_t^{\mathcal{M}}\pi_t+F(\widehat{a})(-1+\sigma_t^{\mathcal{E}}+\sigma_t^{\mathcal{I}}-\sigma_t^{\mathcal{M}}\pi_t)}{c} & \text{if } F(\widehat{a}) \geq \frac{c(1-\sigma_t^{\mathcal{I}})+\sigma_t^{\mathcal{I}}-\sigma_t^{\mathcal{M}}\pi_t-1}{c(1+\sigma_t^{\mathcal{E}}-\sigma_t^{\mathcal{I}})+\sigma_t^{\mathcal{E}}+\sigma_t^{\mathcal{I}}-\sigma_t^{\mathcal{M}}\pi_t-1} \\ \frac{\pi_t(1-\sigma_t^{\mathcal{I}}+\sigma_t^{\mathcal{M}}+F(\widehat{a})(-1+\sigma_t^{\mathcal{E}}+\sigma_t^{\mathcal{I}}-\sigma_t^{\mathcal{M}}))}{c} & \text{else} \end{cases}$$

The first expression gives the optimal oversight strategy when the technocrat exerts some effort, i.e., $e_t^{T*} > 0$. The latter expression gives the optimal oversight strategy when the technocrat exert no effort ($e_t^{T*} = 0$). By straightforward examination of these expressions, we can see that direct implication of m_t^* is that the monitoring rate is increasing in the share of loyalists in the pool of bureaucrats ($\frac{\partial m_t^*}{\partial \pi_t} > 0$). At first glance, this is counterintuitive because we might expect more monitoring where preference misalignment between the principal and agent is larger, which occurs when there are more technocrats in the pool of bureaucrats. However, because loyalists share policy preferences with the politician, the marginal return to their effort is greater, thereby increasing the politician's willingness to engage in costly monitoring to induce the bureaucrat to register households.

Finally, consider the politician's retention decision. It is straightforward to see that the politician will retain the bureaucrat if their expected utility from retention exceeds that of contracting a new bureaucrat. The tradeoff here is straightforward: retaining a bureaucrat increases the bureaucrat's ability to accurately allocate the service correctly and, in the case of a loyalist, pad the rolls with ineligible households. However, it reduces the share of loyalists in the pool. But beyond this (assumed) tradeoff, the politician's monitoring strategy and bureaucrat's effort allocation influence



(a) The politician appoints a new bureaucrat in the blue region and retains the incumbent bureaucrat in the white region. This depicts selection into treatment.

(b) The politician monitors a new bureaucrat more intensively in the gold region ($m_1^* > m_2^*$). A new bureaucrat is appointed in the blue region.

(c) A newly appointed bureaucrat exerts more effort in expectation in the green region ($E[e_1^*] > E[e_2^*]$). A new bureaucrat is appointed in the is the blue region.

Figure 1: All plots set $F(\widehat{a}) = 0.25$, $c = 4$, $\sigma_1^{\mathcal{E}} = .6$, $\sigma_1^{\mathcal{I}} = .6$, $\sigma_1^{\mathcal{M}} = .1$, $\pi_1 = 0.7$.

the politician's decision of whether to keep the bureaucrat or hire someone new.

Proposition 1 (Equilibrium)**.** *If* $E[U_P(m_2^*; \pi_1 - \rho, \sigma_1^{\mathcal{E}} + \delta, \sigma_1^{\mathcal{I}} + \delta, \sigma_1^{\mathcal{M}} + \delta)] \geq E[U_P(m_1^*; \pi_1, \sigma_1^{\mathcal{E}}, \sigma_1^{\mathcal{I}}, \sigma_1^{\mathcal{M}})]$, *the politician replaces the bureaucrat and monitors at rate* m_1^**. If the bureaucrat is a technocrat, she exerts effort* e_1^{T*} *and if she is a loyalist, she exerts effort* e_1^{L*}*. Else, the politician keeps the bureaucrat and monitors at rate* m_2^**. If the bureaucrat is a technocrat, she exerts effort* e_2^{T*}*, and if she is a loyalist, she exerts effort* e_2^{L*}*.*

Figure 1 presents a number of comparisons of interest that stem from this analysis. First, the bureaucrat is replaced with a new appointee in the blue region in panel (a). This depicts selection into treatment (from the perspective of the empirical analysis). The fact that retention is more valuable as learning increases (higher δ) or when turnover in the pool of bureaucrats is lower (lower ρ) should not be especially surprising, but this determination is not simply mechanical. It is influenced by the politician's monitoring determination (panel (b)) and the bureaucrat's effort (panel (c)). The latter two are measures of equilibrium actions with the equilibrium selection decision superimposed in blue.

Panels (b)-(c) show that for this set of feasible parameters, in the region in which the politician replaces the bureaucrat in equilibrium, oversight is greater for appointees and, for most of the region, equilibrium effort is higher among appointees.[5] In contrast, in the region in which the politician retains the incumbent bureaucrat in equilibrium, equilibrium monitoring could be higher or lower among appointees and equilibrium effort is greater among retained bureaucrats in much of the region. When we compare retained to reappointed bureaucrats empirically, thus, the politician's strategic retention decision thus should generally *attenuate* differences in equilibrium oversight and effort relative to what we would see if the retention decision were random.

Empirical implications: This equilibrium suggests a number of empirical questions[6] for the study of the politician's monitoring and the bureaucrat's effort allocation strategies in addition to two equilibrium outcomes: data quality and the transfer allocation. The model helps to guide interpretation of the differences in the behavior and outputs of *newly appointed* versus *retained* bureaucrats. Within the notation above, this means making comparisons between $t = 1$ and $t = 2$.

Consider first the politician's choice of oversight strategy. Is oversight more frequent/punitive for new or retained bureaucrats? Monitoring *decreases* in the share of technocrats which, in isolation, would imply less oversight of retained bureaucrats. However, there is a countervailing force. Bureaucrats work harder when they know more and when they are monitored more stringently. Politicians gain more when bureaucrats work harder, which induces a strategic complementarity between bureaucratic knowledge and the monitoring rate. This implies that for sufficient gains in knowledge, the politician is actually better off monitoring the second bureaucrat more intensively. It can be shown that when there is sufficient patronage—i.e., a large enough change in the pool of new appointees versus incumbents willing to continue their employment—monitoring should be more intensive for new bureaucrats.

[5] I have not imposed probability measure over δ or ρ, so it is not sensible to make statements about the likelihood of these occurrences.

[6] By empirical question, I refer to a prediction that is theoretically ambiguous in sign, but for which empirical findings may be informative about the equilibrium in question.

Remark 1. *For a sufficiently large change in the pool of bureaucrats from $t = 1$ to $t = 2$, i.e., for sufficiently large ρ, politicians impose stronger oversight among new bureaucrats, $m_1^* > m_2^*$. (Proofs in appendix.)*

Do new bureaucrats exert more effort than their retained counterparts? There are three forces at work. First, recall that, all else equal, retained bureaucrats are more likely to be technocrats who exert less effort than loyalists. But all else is not equal. Second period bureaucrats are more accurate, which generates greater incentives to work via their preferences over policy outcomes. These forces are therefore counterveiling. Third, Remark 1 shows that when a system is sufficiently patronage-laden, monitoring is higher among new bureaucrats than existing bureaucrats, which elicits greater effort. Which of these forces prevails is therefore an empirical question. Nevertheless, for a sufficiently patronage-laden personnel system, newly appointed bureaucrats should work harder than their reatined counterparts.

Remark 2. *For a sufficiently large change in the pool of bureaucrats from $t = 1$ to $t = 2$, i.e., for sufficient ρ, new bureaucrats exert greater effort in expectation $E[e_1^*] > E[e_2^*]$.*

There there are two further empirical questions related to an equilibrium outcome: the share of poor or qualifying households recorded in the register data. Recall that the share of households classified as eligible by the technocrat is:

$$e_t^{T*}\left[F(\widehat{a})\sigma_t^{\mathcal{E}} + (1 - F(\widehat{a}))(1 - \sigma_t^{\mathcal{I}})\right],$$

whereas the share of households classified as eligible by the loyalist is:

$$e_t^{L*}\left[F(\widehat{a})\sigma_t^{\mathcal{E}} + (1 - F(\widehat{a}))(1 - \sigma_t^{\mathcal{I}} + \sigma_t^{\mathcal{M}})\right].$$

By assumption, loyalists classify more households as poor (holding fixed effort) and are overrepresented among new appointees relative to retained bureaucrats. In a sufficiently patronage-laden

personnel system, new bureaucrats should classify more households as poor over time. This is particularly the case when new bureaucrats exert more effort than their retained counterparts.

Remark 3. *For a sufficiently large change in the pool of bureaucrats from $t = 1$ to $t = 2$, i.e., for sufficient ρ, new appointees report more households below the eligibility threshold than retained bureaucrats.*

The final empirical implication considers the effect of increased expertise within an individual bureaucrat's data collection. Suppose that we can observe more subtle increases in knowledge as a bureaucrat gains experience. Treating the politician's monitoring strategy as sticky the short term (i.e., for a given bureaucrat), monotonic gains in knowledge should drive bureaucrats to work incrementally harder (i.e., expand the register). However, conditional on a household entering the register, the eligibility classification should have diverging patterns among the two bureaucrat types. For technocrats, gains in knowlege help to correctly classify more households of each type (eligible and ineligible). However, under the assumption that less than half of households should be eligible, these accuracy gains avoid more errors among ineligible households. In contrast, for loyalists, gains in accuracy are concentrated in eligible households, yielding an increase in households classified as "qualified." Thus, if the share of loyalists is sufficiently high, we should see aggregate increases in qualification in a bureaucrat's tenure.

Remark 4. *As knowledge increases over a bureaucrat's tenure, individual bureaucrats classify more households as eligible for the social program when there is a sufficient share of loyalists in the pool of initial bureaucrats, i.e., for sufficient π_1.*

2 Contexts

This study examines agency problems in the production of social registries in Brazil and Colombia. Specifically, I examine Brazil's Cadastro Único para Programas Sociais (CadUnico) and Colombia's Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (SISBÉN).

Both registries are produced and maintained by bureaucrats working in municipal governments. Similarly, both are national in scope and govern entrance to social welfare programs funded and administered by their respective national governments. Thus, whereas Brazil is a federal state and Colombia is a unitary state, the principal intergovernmental interaction is between the national (federal) and local (municipal) governments in both contexts. In practice, municipal governments in both contexts are highly decentralized (Falleti, 2005, 2012).

2.1 Municipal Bureaucracies

Municipal governments in Brazil and Colombia are led by an elected mayor (the local principal) and staffed by local bureaucrats. Mayors are elected every four years in local elections. In Brazil, mayors can serve a maximum of two consecutive terms whereas in Colombia, incumbent mayors are barred from immediate re-election. As a consequence of these term limits, Colombian municipalities experience routine mayoral turnover every four years, whereas some Brazilian mayors serve eight years (two terms) consecutively. Given data availability (see below), I study data production under Brazilian mayors elected in late 2008 who served from 2009-2012 and Colombian mayors elected in late 2019 who served from 2020-2023.

In both countries, mayors oversee the hiring and performance of local bureaucrats. Bureaucrats can be hired as civil servants or as contractors. Civil servants are afforded some tenure protections whereas contractors are typically hired (and rehired) on short contracts. In Colombia, for example, these contracts last an average of three months, but are often renewed (Slough, 2022). Since mayors typically oversee contracting (Rueda and Ruiz, 2022), they typically have greater power to select and fire contractors than civil servants. Peterlevitz (2023) argues that contractors can be largely viewed as patronage positions in the Brazilian context as well. He further documents substantial variation across municipalities in rates of contracting in some sectors (e.g., healthcare).

Figure 2 plots the tenures of bureaucrats who run or contribute to the production of social register data in Brazil and Colombia. As detailed below, the Brazilian data comes from records from

the Annual Social Information Survey (RAIS) employer-employee linked dataset. The Colombian data was reported in an original survey of these bureaucrats. The graphs have two common features. First, tenures are relatively short. The bureaucrat with the median tenure was hired into local government at the time of (potential) mayoral turnover (early 2009 in Brazil and early 2020 in Colombia). The median tenure as a CadUnico interviewer/SISBÉN administrator is about two years. Second, we observe spikes in hiring and, in Colombia, reassignment to the SISBÉN administrator position at this point of political turnover.[7] This increase is more dramatic in Colombia than Brazil. However, this appears to be at least partially driven by the prohibition on consecutive re-election in Colombia. To this end, Figure A1 supports this hypothesis by showing that in Brazil, that mayoral turnover (as opposed to re-election) induces short-term increases in hiring of CadUnico interviewers on both the intensive and extensive margins. In general, this variation in bureaucratic tenure points to the importance of the selection, retention, and assignment of bureaucrats to their posts as a decision of local mayors.

2.2 Social Registries

The process of producing and maintaining social registries is distinct from other forms of state data production because it involves both service provision to citizens (interviews) and the reporting of the resultant survey data to the national government. When we examine the resultant administrative data, the two actions are not fully distinguishable. In other words, variation in data quality captures both variation in effort to interview citizens and in accurately reporting observations to the national government. To the extent that these processes have been studied in earlier iterations of the Colombian means-testing program, Slough (2022) documents uneven service provision that may limit access to the survey and Camacho and Conover (2011) shows variation in reporting. In Brazil, the citizen-facing service provision by CadUnico interviewers makes them attractive targets

[7]Note that both ECDFs cross among recently-hired employees, meaning that some agents were contributing to social registries before they were hired by local government. This is likely due to the use of interns/contractors for registry interviews prior to formal contracting, as documented in Brazil by Frey and Santarrosa (2024).

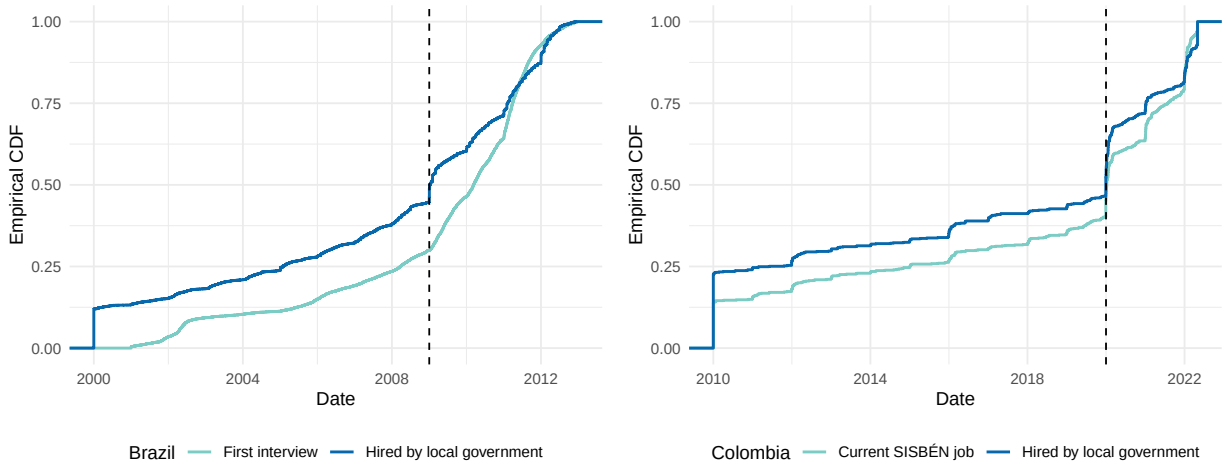


Figure 2: Tenure of bureaucrats (a) in local governments and (b) as agents running/administering social registries in Brazil (left) and Colombia (right). Tenures are censored to 12 years. In Brazil, 12.5% of bureaucrats were working for local governments before 2000; in Colombia, 22.5% of bureaucrats were working for local governments before 2010. The Brazilian data is weighted such that each municipality is weighted equivalently for comparability with the Colombia data.

for politicization by incumbent parties (Frey and Santarrosa, 2024).

The principal differences between these registers are the measures used by the Brazilian and Colombian national governments to determine eligibility for social programs are different.[8] In Brazil, eligibility for transfers or programs is a direct function of declared household income. For example, to qualify for the Bolsa Familia conditional cash transfer, two distinct, public thresholds determine access to unconditional and conditional (family-size dependent) benefits. While these thresholds are adjusted over time to account for inflation and programmatic goals, the eligibility determination on the basis of CadUnico data is a widely-known function of declared income.[9]

In contrast, in Colombia, local register administrators conduct a survey of household assets and conditions. These data are then sent to the National Department of Planning (DNP), which uses a private algorithm to generate an index score. Social programs—including but not limited to the CCT Familias en Acción—use different cut points in this index to determine eligibility. The

[8]These distinct measures correspond to different choices of a_i and $\widehat{a}$ in the model.

[9]Figure A13 shows how reporting of income cutoffs changes in response to changes in these thresholds.

privacy of the algorithm is seen as key to protecting the means-testing policy (and the rolls of social program beneficiaries) from distortion by local politicians. Past iterations of the index have been undermined by manipulation of scores by local politicians and bureaucrats. In their study of SISBÉN-I—the first iteration of the means-testing index—Camacho and Conover (2011) show bunching under threshold for social program eligibility (to increase the number of households that qualify) after the formula was released to local bureaucrats. Subsequent iterations of SISBÉN were designed in response to manipulation or perceived limits to the performance of the indices in the targeting of social programs. The current iteration, SISBÉN-IV, was introduced in March 2021. To prepare for the introduction of SISBÉN-IV, national contractors aided in resurveying all households. The number of enrolled households consequently decreased by an average of 33% in each municipality (see Figures A2-A4). Some of this decline came from de-duplication[10] or actualization of records; some other omissions resulted in a loss of benefits.

In sum, CadUnico and SISBÉN are produced by the same actors and for the same purposes in Brazil and Colombia, respectively. The principal differences relate to the Brazilian and Colombian national governments' choice of how to construct the means-testing measures. These differences include: (1) the nature of the measure (income vs. an index of assets and household conditions); (2) whether the measure is directly constructed/reported by bureacrats or generated by a private formula; and (3) the nature of changes in the means-testing measure over time.

3 Research Design

3.1 Data

The analysis employs a combination of administrative data from Brazil as well as an original survey and administrative data from Colombia. Given differences in timing, data sources, data structure, and the function of CadUnico and SISBÉN registries, I examine the empirical implications using

[10] Appendix A3.1 provides descriptive evidence that some of the duplication came was produced by households re-entered in SISBÉN-III when they were surveyed for SISBÉN-IV in 2018 or 2019.

the available data. Given the differing data structures, there are no directly comparable quantities of interest. Nevertheless, both contexts offer testing of different implications of the theory.

3.1.1 Social Register Data

The primary outcomes of interest come from the social register data from CadUnico and SISBÉN. The reported data allow for an examination of the behavior of data-producing bureaucrats within Brazilian and Colombian local governments. However, there are substantial constraints on access to these registries.

In Brazil, following Frey and Santarrosa (2024), I rely on a single cross section of the CadUnico registry from December 2012. This data reports the content of the most recent survey from each of the 30,171,166 households ever registered CadUnico as of 2012.[11] Most observations (60%) record the identification number of the CadUnico interviewer who conducted the most recent survey of household income and assets. This rate increases to 75% among interviews conducted from 2009-2012. The remainder of analysis focuses on the 18,150,550 surveyed households that can be linked to an interviewer in the most recent survey.

In Colombia, I use public, anonymized microdata from repeated cross sections of SISBÉN-IV from 2021 and 2022. The public microdata includes a large stratified random sample of households from both the central (*cabecera*) and rural (*rural disperso*) zones of each municipality. In urban zones, this sample ranges from 0.05% of the SISBÉN register in Bogotá to 97% of the SISBÉN register of small municipalities; in rural zones, this sample ranges from 4% (in Pasto, Nariño) to 97% in small municipalities. The microdata samples are large: the 2021 data include 1,319,585 households sampled from the 8,496,846 registered households. Enrollment increased by nearly 23% to 10,437,045 households by December 2022.[12]

I aggregate data to the level of the bureaucrat (interviewer) in Brazil and the municipality

[11] In total, 79.4% of households were resurveyed at least once after initial enrollment.

[12] This growth should be interpreted in light of the 33% reduction in household enrollment in the transition from SISBÉN-III to SISBÉN-IV in March 2021.

Attribute	CadUnico (Brazil)	SISBÉN-IV (Colombia)
	PANEL A: RAW DATA	
Cross-section	Household	Household or individual
Time	December 2012	Annual from 2021-2022
Sample	Universe of households	Random sample by municipality-zone
	PANEL B: PROCESSED DATA	
Unit of analysis	Bureaucrat	Municipality-year
Years analyzed	2009-2012	2021, 2022
Total municipalities	5,568	1,102
Municipalities in sample	5,385	752
Sample restrictions	Municipality present in CadUnico ($n = 5554$), municipal government present in RAIS ($n = 5,554$), in which any interviewer from 2009-2012 term is located in RAIS data ($n = 5,385$).	SISBÉN administrator answered survey ($n = 752$).

Table 2: Raw administrative social register data and aggregate units of analysis.

zone-year in Colombia. Note that because the survey includes a single SISBÉN administrator per municipality in Colombia, the municipality level is observationally equivalent to the "bureaucrat" level in this context. Table 2 summarizes the properties of the raw and processed social register data that serve to measure bureaucratic reporting behavior.

3.1.2 Survey of Colombian SISBÉN Administrators

I conducted an original survey of Colombian SISBÉN administrators in municipal governments. The goal of the survey was to measure characteristics of these administrators, their terms of employment, incentives, and management practices. To identify the relevant administrator in each municipality, the research team submitted a freedom of information request to the DNP, the national agency that administers SISBÉN. The resultant data included the complete contact information (both email addresses and phone numbers) for 1,074 SISBÉN administrators. This covers 97.5% of Colombia's 1,102 municipalities.[13]

In order to maximize the response rate, we administered the survey by web and phone. Importantly, this mirrors the two principal means through which the national government solicits data

[13]In missing municipalities, there was a vacancy at the time of the FOI request or the national government did not have a current contact.

from other government entities, as documented in Garbiras-Díaz and Slough (2025). The survey was initially distributed to all subjects by email. Two weeks after the distribution of the survey, enumerators began to conduct phone surveys with subjects who had not already completed the online version of the survey. After the phone surveys, we distributed a second invitation to participate in the web survey. In sum, 74% of the surveys were completed online and 26% of the surveys were completed by phone. All surveys were completed between May 1 and August 10, 2022.

In sum, the survey data contain responses from 752/1,074 municipalities in the sampling frame, a 70.2% response rate. This response rate compares favorably to many existing elite surveys. The overall response rate falls at the the 85^{th} percentile of response rates among the 68 elite surveys published in three leading political science journals for which response rates are reported (Kertzer and Renshon, 2022).

Municipal characteristics are generally poor predictors of selection into survey response. Figure A6 and Table A3 show widespread geographic representation of every region of the country within each survey. Table A3 further shows that the characteristics of responding municipalities (e.g., population and municipal category) and metrics of municipal governance compiled by the DNP are very similar to those of all municipalities in Colombia. In particular, the survey sample closely resembles all Colombian municipalities on these characteristics. Collectively, these figures suggest that the combination of phone and web surveys undertaken was successful in cultivating a high response rate that is representatitive of the distribution of Colombian municipalities. See Table A3 reports summary statistics on individual respondents.

The survey consisted of several modules intended to measure the characteristics and incentives of data-producing bureaucrats. First, the surveys sought to understand employees' incentives within the *alcaldía* in which they are employed. It measures whether the employee is a contractor or civil servant, as well as who the employee reports to, and their observation of penalties used against other employees. A second module, distinct to each program, measures the objectives that have been specified for the collection and reporting of data. This module echoes recent surveys

of politician management practices used in Italy by Carreri (2021) and in the US by Carreri and Payson (2023). A third module measures time use following best practices established by Kalaj, Rogger, and Somani (2020). A fourth module measures contact between data producers and the respective national governments that they report data to. A final module collects data on demographic characteristics and the officials' career tenures within the *alcaldía* in which they work.

3.1.3 Employer-Employee Data in Brazil

In Brazil, I rely on the RAIS employer-employee dataset to measure the characteristics of bureaucrats and the employment of interviewers that appear in CadUnico. To do so, I begin the universe of employees of local governments between 2008 and 2012, at the bureaucrat-locality level ($n = 8,181,729$). Of these observations, the vast majority of individual bureaucrats (89%) worked for a single municipality and are therefore unique in the dataset. Individual ID numbers provide a link between the CadUnico records and the employee who submitted the data. This merge yields a universe of $n = 33,669$ unique bureaucrats who conducted at least one CadUnico interview between 2009 and 2012 for whom CadUnico and RAIS data can be linked. The merged sample covers $5,385$ municipalities. These bureaucrats range substantially in their tenures and productivity with a median of 122 (IQR: [22, 407]) household entries per bureaucrat.[14]

The RAIS data contains information on (1) duration of employment; (2) contract type; and (3) monthly salary in addition to individual-level covariates. Thus, these data provide alternative measures of bureaucratic incentives to those in the survey. The RAIS data also include covariates that measure bureaucrats' demographic characteristics and individual qualificaions. These data facilitate descriptive comparision between survey-measured attributes of SISBÉN administrators in Colombia and RAIS-measured attributes of CadUnico interviewers in Brazil.

Table 3 compares features of the bureaucrats in each context. In general, the samples are broadly similar: the median bureaucrat is mid-career, female, and a high school graduate. The ad-

[14]When a household is reinterviewed, only the most recent interviewer is recorded, so this number is a lower bound on the number of interviews completed.

	Brazil (2009-2012) CadUnico Interviewers				Colombia (2022) SISBÉN Administrators			
Variable	Min.	Max.	Mean	St. Dev.	Min.	Max.	Mean	St. Dev.
Age	15	79	34.98	10.04	19	65	39.35	9.98
Female	0	1	0.75	0.43	0	1	0.63	0.48
High school complete	0	1	0.89	0.31	0	1	0.99	0.08
Undergraduate degree complete	0	1	0.32	0.47	0	1	0.40	0.49
Postgraduate degree complete	0	1	0.004	0.068	0	1	0.08	0.28
Resident of municipality	–	–	–	–	0	1	0.94	0.24
Civil servant (indicator)	0	1	0.93	0.25	0	1	0.66	0.47
Tenure in local government (years)	0	42.84	5.56	6.33	0	39.49	7.36	8.89
Tenure in job (years)	0	12	3.26	2.94	0	30.32	5.19	6.57

Table 3: Characteristics of CadUnico interviewers and SISBÉN enumerators from the RAIS data (Brazil) and survey data (Colombia).

ministrators are slightly better educated in Colombia than in Brazil, though the educational qualifications of Brazilian municipal bureaucrats increased during the intervening decade (2012 to 2022) (Slough, 2024), which may account for the modest observed difference. On average, bureaucrats had served in local governments longer than in their current post, meaning that some bureaucrats were transferred into their appointments working on social registries.

3.2 Mapping of Empirical Questions and Implications

With both the Colombian and Brazilian data, the core contrast of interest is whether a register administrator was appointed by the mayor vs. retained by the mayor from a past administration. These measures are coded from the survey data in Colombia and the RAIS employer-employee data in Brazil. Table 4 summarizes how the data sources discussed above are used to generate outcome measures to evaluate the empirical questions in Remarks 1-4.

3.3 Estimation and Inference

Politician oversight: To examine the prediction that mayors oversee their appointees more stringently than holdovers from previous mayors, I use the Colombian survey data to estimate the

	Implication	Brazil	Colombia
Q1	For sufficiently large change in the pool of bureaucrats, politicians oversee appointed bureaucrats more stringently than retained bureaucrats.	–	Survey of SISBÉN administrators measuring oversight experience.
Q2	For sufficiently large change in the pool of bureaucrats, newly appointed bureaucrats exert more effort than retained bureaucrats.	Bureaucrat-level counts of households surveyed from CadUnico.	Yearly change in municipal zone-level SISBÉN enrollment data.
Q3	For sufficiently large change in the pool of bureaucrats, newly appointed bureaucrats score more households as eligible for the cash transfer.	Bureaucrat-level household income classifications.	Yearly changes in the distribution of scored households in municipal zone- level SISBÉN scores.
Q4	As bureaucratic tenure increases, they score more households as eligible for the transfer.†	Household-level income classifications.	

Table 4: Data sources and outcome measurement related to empirical implications. †Note that the contrast of interest for implication 4 is a bureaucrat's tenure or experience, not whether they are appointed or retained.

following OLS specifications:

$$\text{Oversight}_m = \beta_1 \text{Appointed}_m + \beta_2 \text{Contractor}_m + \boldsymbol{\gamma}\mathbf{X}_m + \varepsilon_i \tag{4}$$

I use survey-based measures of oversight: a measure of the frequency of oversight, a measure of whether oversight is conducted directly by the mayor, a measure of whether goals are set by supervisors, and a reverse-coded measure of freedom in their job.[15] The coefficient β_1 aims to measure the difference in reported oversight between appointed and retained bureaucrats, conditional on the controls. Because appointees are more likely to be contractors, a second panel also includes a contractor indicator to ensure that observed differences are not solely attributable the differences in the terms of employment alone. The remaining covariates in $\mathbf{X}_i$ are bureaucrat characteristics including age bin indicators, (five) education category indicators, and gender, which aim to purge systematic differences in the pools of bureaucrats beyond their (1) preferences and (2) experience/knowledge. Since the measures of oversight are are oriented such that higher values indicate more stringent oversight, implication #1 holds that $\beta_1 > 0$.

[15] See Table A4 for English translations of these survey questions.

Bureaucratic effort: Turning to measures of bureaucratic effort, consider first the data from Colombia. Using municipal enrollment statistics inferred from the random sample of households, I examine the change in the number of households in a given classification from year t to year $t+1$. I estimate specifications of the form:

$$\text{Share enrolled}_{m(z),t} - \text{Share enrolled}_{m(z),t-1} = \beta_1 \text{Appointed}_m + \beta_2 \text{Contractor}_m + \boldsymbol{\kappa}\mathbf{X}_m + \epsilon_{mz} \tag{5}$$

These specifications examines changes in reporting from year t to $t+1$, directly measuring actions taken by the SISBÉN administrator and, where relevant, their office. To measure effort, I focus on the change in households entered in SISBÉN as a share of population from the 2018 Census. This provides a direct measure of effort with interpretation that is aligned with the interpretation of the e choice variable in the model. I report estimates by zone (population center versus rural area), z, and aggregated at the municipal level. In this specification, β_1 corresponds to the difference in effort exerted by appointed versus retained SISBÉN administrators. As in (4), I include a contractor indicator in some specifications to ensure that this difference is not solely being driven by the difference in contracts within the two pools. I also include individual bureaucrat fixed effects—age, education, and gender—and municipal covariates—municipal category fixed effects and population quintiles.

In Brazil, I measure effort at the level of the bureaucrat through the the number of interviews completed. This data allows for comparisons between bureaucrats within the same municipality under the same mayor. The specification in (6) mirrors those used on the cross-sectional Colombian data. The employment data and covariates—again, age bins, educational attainment, and gender—are measured at the level of the bureaucrat (i). However, since multiple bureaucrats can enroll households in a given municipality-month, we can compare individual bureaucrats' behavior within the same municipality and time through the inclusion of municipal ($\boldsymbol{\psi}_m$) and bureaucrat experience ($\boldsymbol{\kappa}_t$) fixed effects. The experience fixed effects ($\boldsymbol{\kappa}_t$) are indicators for the total num-

ber of interviews conducted, which are generated by binning the distribution of interviews into 5 percentile bins.[16]

$$\ln(\text{Interviews}_{imt}) = \beta_1 \text{Appointee}_i + \beta_2 \text{Contractor}_i + \boldsymbol{\psi}_m + \kappa_t + \boldsymbol{\gamma}\mathbf{X}_i + \varepsilon_{imt} \tag{6}$$

Scoring Households: To understand whether bureaucratic appointment characteristics influence the distribution of social transfers—the ultimate outcome of interest—I examine the distribution of SISBÉN scores in Colombia. Specifically, there are thirty ranked categories, broken into four super categories: extreme poverty, moderate poverty, vulnerable, and neither poor nor vulnerable. Increases in the share of residents in lower categories by rank increase the likelihood of qualification for specific programs.[17] To estimate changes in scores, I estimate a specification analogous to (5). However, instead of using the total number of new households measured as an outcome, I examine the change in the cumulative share of households poorer than a given threshold. Formally, for category $c \in \{1, ...30\}$, the outcome is the difference:

$$\sum_{1}^{c} \text{Share of enrolled households}_{m(z),t} - \sum_{1}^{c} \text{Share of enrolled households}_{m(z),t-1}$$

I then estimate these specifications for each category, c. The results provides a measure of the difference in the distribution of SISBÉN scores as a function of the employment of the SISBÉN coordinator that is analogous to each of the black vertical lines in Figure 3.

In Brazil, recall that qualification for Bolsa Familia depends on two predetermined income thresholds, one for unconditional and one for conditional transfers.[18] I therefore consider the share

[16]These bins will naturally capture some of the on-the-job learning. But by partialling out one determinant of knowledge, they offer a more stringent test for comparing the behavior of mayor-appointed versus retained bureaucrats.

[17]SISBÉN scores are one of multiple qualification types, so sufficiently low SISBÉN scores should be viewed as a necessary but not sufficient condition for enrollment.

[18]The cutoffs increased in August 2009, so I consider the threshold on the date of the interview to construct these outcomes.

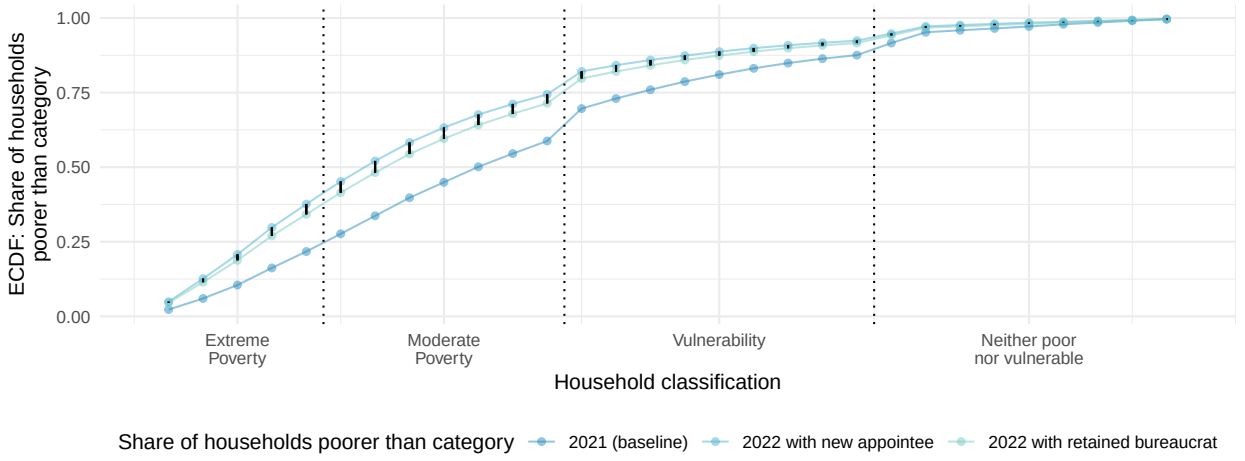


Figure 3: The black vertical segments illustrate the outcome measure of interest from the specifications that measure differences in the change in scores. The data plots the ECDF of raw scores/changes in scores from rural zones from 2021-2022.

of registered households scored below each threshold as well as the share of scores exactly at the threshold (bunching). While one could imagine that there are a non-trivial number of workers at these thresholds because they are multiples of ten, bunching need not be indicative of deceptive reporting. However, it is not clear how households would be allocated across bureaucrats within a municipality such that appointed versus retained bureaucrats would see systematically different numbers of such households. I also examine a measure of Bolsa Familia enrollment—which is largely based on the register scores. I use these five outcomes as dependent variables in the specification in (6) to examine how mayor appointments change the distribution of scores.

Learning: Because the Brazilian data attributes household scores to individual bureaucrats, I examine changes in a bureaucrat's behavior over their tenure to assess assumptions/implications about bureaucrat learning. I first examine the data descriptively by examining how five scoring outcomes change with respect to the order of the interview in a bureaucrat's history. Specifically, I examine means of these outcomes over bins of five interviews. Thus, if a bureaucrat has conducted 200 household interviews, I look at the mean of these outcomes in interviews 1-5, 6-10, etc. Comparison of these means show how scores evolve through a bureaucrats' tenures. However, they do

not account for the selection problem of interest: a politician can retain or remove a bureaucrat from conducting interviews at will. To this end, I leverage within-bureaucrat variation in scoring by estimating a regression of the form:

$$Y_{ibmt} = \beta_1 \text{Order}_i + \boldsymbol{\phi}_b + \boldsymbol{\tau}_m + \boldsymbol{\eta}_t + \varepsilon_{ibmt}, \tag{7}$$

where i indexes the household interviewed, b indexes the bureaucrat that conducted the interview, m indexes the municipality, and t indexes the month of the interview. The inclusion of municipal fixed effects beyond bureaucrat fixed effects accounts for bureaucrats who interviewed in multiple municipalities during the 2009-2012 term. The estimator β_1 measures how outcome Y_{ibmt} changes as a function of a bureaucrat's experience. To lessen the influence of outliers (highly productive interviewers), I evaluate this specification over a bureaucrats' first K interviews, where $K \in \{275, 1275\}$, the 75^{th} and 95^{th} percentiles of interviews conducted, respectively.

4 Results

4.1 Mayors monitor new appointees more stringently

Table 5 reports the association between bureaucratic appointment and multiple measures of oversight. The primary measure of oversight corresponds to the frequency of oversight. Columns 1-2 show that SISBÉN administrators who were initially appointed by the current mayor report higher levels of oversight. The magnitude of these differences corresponds to 0.25 to 0.35 standard deviations of the outcome measure. This increase in oversight is not simply a change in who oversees the bureaucrat (columns 3-4): appointees are not more likely to be overseen by the mayor (rather than another civil servant) relative to retained bureaucrats. However, oversight does correspond to a higher share that experience goal setting by their superiors (columns 5-6). The estimated coefficient represents an increase of 0.15 to 0.26 standard deviations of the goal setting measure. Respondents reported that goals are usually expressed in terms of new households reg-

	Oversight frequency		Oversight by mayor		Goals set		(Lack of) freedom	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PANEL A: BUREAUCRATIC TENURE								
Appointed by mayor	0.441***	0.418***	0.052	-0.011	0.085*	0.117**	0.071	-0.021
	(0.096)	(0.109)	(0.036)	(0.040)	(0.035)	(0.039)	(0.083)	(0.095)
Observations	738	734	750	745	741	737	739	735
PANEL B: BUREAUCRATIC TENURE AND CONTRACTS								
Appointed by mayor	0.314**	0.323**	0.050	-0.003	0.070+	0.095*	0.017	-0.037
	(0.108)	(0.114)	(0.040)	(0.042)	(0.038)	(0.042)	(0.090)	(0.100)
Contractor	0.351***	0.372***	0.005	-0.034	0.037	0.072+	0.108	0.029
	(0.095)	(0.102)	(0.041)	(0.043)	(0.038)	(0.041)	(0.094)	(0.101)
Observation	715	711	727	722	718	714	716	712
Bureaucrat characteristics		✓		✓		✓		✓
DV scale	$\{0, ..., 5\}$		$\{0, 1\}$		$\{0, 1\}$		$\{1, ..., 5\}$	
DV mean, std. dev.	3.133 (1.249)		0.367 (0.482)		0.703 (0.457)		2.333 (1.101)	

$^{+}p < 0.1,^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$

Table 5: The relationship between bureaucratic tenure and contract type and perceptions of oversight in Colombia. Covariates included in specifications bureaucrat characteristics include age bins, five education category bins, and gender of the bureaucrat. Heteroskedasticity-robust standard errors in parentheses.

istered/actualized (68%), the length of the municipal register (65%), and/or resolution of citizen complaints/concerns (66%). The composition of these specific goals does not vary detectably by type of bureaucratic appointment. This finding suggests that mayors (principals) set higher expectations for effort of appointees relative to retained bureaucrats. Finally, I do not observe clear evidence of changes in perceived freedom in the workplace.

One potential concern with this interpretation of the results in Table 5 is that these results could simply reflect differences in oversight behavior in municipalities where bureaucrats are replaced/transferred at higher rates instead of differences in oversight targeted to new versus retained administrators. Two additional findings provide evidence against this alternative interpretation. First, eliciting beliefs about punishments applied to contractors (more likely to be appointees) versus civil servants (comparatively less likely to be appointees) reveals substantial variation in the application of penalties (Figure A7) that do not vary in the employment status of the respondent.

Second, the survey was fielded in parallel for (up to) three administrators of different data collection processes per municipality. This allows for estimation in the within-municipality correlation of responses to the outcomes in Table 5. Table A5 that the ICC is very low (-0.04 to 0.07 across outcomes), suggesting that targeting of oversight to individual bureaucrats or classes thereof, e.g., new appointees, is a reasonable representation of politician strategies.

An second alternative interpretation of this finding holds that bureaucrats appointed by the mayor are simply monitored more because they are new to the job. If this were the case, we might expect that appointees with longer tenures in the job ($\approx 2 - 2.5$ years) would be overseen with less intensity than very new appointees. Figure A8 reveals no systematic variation in levels of oversight or goal setting as a function of tenure among mayoral appointees.

4.2 Effort

I evaluate bureaucratic effort through measures of output: the interviews conducted. In Colombia, I measure the growth in share of households in SISBÉN between 2021 (the first use of SISBÉN-IV) and 2022. Table 6 shows that, relative to reappointed bureaucrats, mayoral appointees increase the rolls by an average of 2 percent of households (panel C, column 5). This represents an increase of 12.8% relative to the average change in the share of households registered (17.9 percent). Thus, while the rolls grew in 98.1% of municipalities between 2021—the introduction of SISBÉN-IV—and 2022, they grew by more in municipalities with a SISBÉN administrator appointed by the current mayor. Furthermore, these changes were not driven by differences in baseline registration. Table A6 reports analogous specifications that use baseline registration as the outcome of interest and shows no systematic differences between municipalities with an mayoral appointee versus a retained bureaucrat.

Panels A and B break down this increase into the two zones, reflecting the most granular analysis permitted by the data. All estimates are positive, though only those in rural areas are statistically significant at current thresholds. However, since growth in SISBÉN was disproportionately con-

	Δ Share of households enrolled, 2021-2022				
	(1)	(2)	(3)	(4)	(5)
PANEL A: HOUSEHOLDS IN POPULATION CENTER					
Appointed by mayor	0.016	0.019^{+}	0.019	0.018	0.017
	(0.010)	(0.011)	(0.012)	(0.012)	(0.012)
DV mean (Std. dev.)	0.141 (0.135)	0.141 (0.135)	0.141 (0.135)	0.141 (0.135)	0.141 (0.135)
Observations	748	748	748	748	748
PANEL B: HOUSEHOLDS IN RURAL AREA					
Appointed by mayor	0.031**	0.036***	0.032**	0.029*	0.029*
	(0.010)	(0.011)	(0.012)	(0.012)	(0.012)
DV mean (Std. dev.)	0.179 (0.155)	0.179 (0.155)	0.179 (0.155)	0.179 (0.155)	0.179 (0.155)
Observations	747	747	747	747	747
PANEL C: ALL HOUSEHOLDS					
Appointed by mayor	0.023**	0.036***	0.032**	0.021**	0.020*
	(0.008)	(0.011)	(0.012)	(0.008)	(0.008)
DV mean (Std. dev.)	0.156 (0.113)	0.156 (0.113)	0.156 (0.113)	0.156 (0.113)	0.156 (0.113)
Observations	747	747	747	747	747
Contractor		✓	✓	✓	✓
Bureaucrat covariates			✓	✓	✓
Municipal covariates				✓	✓
Share of households enrolled in 2021					✓
Missingness indicators		✓	✓	✓	✓

$^{+}p < 0.1,^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$

Table 6: Changes in enrollment in SISBÉN, 2021-2022 as a function of bureaucratic appointments. The denominator for calculating the share of households comes from the 2018 census. Robust standard errors in parentheses.

centrated in rural areas, the point estimates suggest similar effects relative to the overall rate of change (12.0 in population centers vs. 16.8% in rural areas). In sum, this analysis suggests that appointed SISBÉN administrators exerted more effort from 2021-2022, thereby registering more households, than did their reappointed colleagues.

Turning to Brazil, Table 6 reports that bureaucrats similarly exert greater effort. In the top panel, I focus on entries in CadUnico between 2009 and 2012, e.g., during the mayor's term. The most demanding specification (column 4) suggests that relative to retained CadUnico bureaucrats, mayoral appointees in the same municipality conduct 46% more interviews than retained bureaucrats. Relative to the median of 65 interviews, thus, this corresponds to an additional 30 interviews. This provides suggestive evidence that, as in Colombia, appointed bureaucrats in Brazil exert greater effort than their retained counterparts.

However, interpretation of Panel A is complicated by the fact that households should update their registrations at least every two years. If, for example, new appointees were tasked with outreach while existing bureaucrats were tasked with updating existing records that they had previously entered, this measure would overstate the productivity of appointees. Thus, comparison of records entered in the last two years of the mayor's term, 2011 and 2012, offer cleaner comparisons since updating and entering new households are counted in a more similar fashion. In Panels B-C again see that appointed bureaucrats complete more interviews. In these assessments, the most conservative estimates (column 4) suggest that appointees entered an additional 5% (in 2011) and 17% (in 2012) of households. This narrows the sample to approximately 60% of the bureaucrats examined in Panel A. These bureaucrats also appear (mechanically) more productive since substantially fewer records have been updated. The 5% in 2011 corresponds to an increase of 2.5 interviews (relative to a median of 50) whereas the 17% in 2012 corresponds to an additional 16 interviews (relative to a median of 98).

Interestingly, the *relative* effect sizes for appointed relative to retained SISBÉN and CadÚnico administrators is of roughly similar magnitude when focusing on a single year period. In Colombia,

	Log Households Interviewed			
	(1)	(2)	(3)	(4)
PANEL A: TOTAL INTERVIEWS IN REGISTER, DECEMBER 2012				
Appointed by mayor	0.781***	0.776***	0.549***	0.461***
	(0.043)	(0.044)	(0.040)	(0.029)
Observations	41253	41253	41253	41253
PANEL B: HOUSEHOLDS INTERVIEWED IN 2011				
Appointed by mayor	0.120**	0.122**	0.097*	0.053*
	(0.041)	(0.042)	(0.042)	(0.022)
Observations	25316	25316	25316	25316
PANEL C: HOUSEHOLDS INTERVIEWED IN 2012				
Appointed by mayor	0.203***	0.205***	0.189***	0.168***
	(0.041)	(0.042)	(0.041)	(0.023)
Observations	25680	25680	25680	25680
Municipal FE	✓	✓	✓	✓
Contractor Indicator		✓	✓	✓
Bureaucrat Characteristics			✓	✓
Interview experience FE				✓

$^{+}p < 0.1,^{*}\ p < 0.05,^{**}\ p < 0.01,^{***}\ p < 0.001$

Table 7: Bureaucratic appointment and output in Brazil. Standard errors are clustered at the municipality level.

recall that the marginal effect of a newly-appointed administrator in charge of SISBÉN was 12.8%. In Brazil, the (roughly) analogous marginal effects average 11% over the two years for which the CadÚnico data permit credible comparisons.

4.3 Qualification for Social Programs

A bureaucrat can influence social register data both by (1) reaching more households and (2) scoring households differently. Focusing on the set of households scored by bureaucrats, how do bureaucratic influence scores? In Colombia, appointed bureaucrats report that more households are poor or vulnerable. Figure 4 show that the share of households classified as poor expands under appointed bureaucrats since positive estimates suggest more of the population is classified as poorer than category c. This is not driven by baseline differences across municipalities. Recall that baseline 2021 SISBÉN scores were largely collected by national government contractors associated with the rollout of SISBÉN-IV. The share of individuals

Given the increases in the effort of mayoral appointees documented in Table 6, one may be concerned that the results in Figure 4 reflect only increases in effort directed at poorer populations or areas in a municipality. However, an examination of changes in the *number* of households classified in each category from 2021 to 2022 reveals a reduction in the number of households classified as "vulnerable" or "neither vulnerable nor poor" (Table A7). These reductions are larger in municipalities with a mayoral appointee (Figure A11). These reductions show that differential rates of reclassification toward poorer categories—in addition to differential growth in enrollment—must also be driving the results in Figure 4.

Turning to the Brazilian data, mayoral appointees similarly record a greater share of household incomes below income cutoffs. Figure 5 shows that appointees report more households below the conditional threshold. The increase of 2.3 percentage points in the most conservative specification represents a 3.7% increase on the mean of 0.64 among all bureaucrats. There is suggestive evidence ($p < 0.08$) that mayoral appointees similarly report a higher share of households below

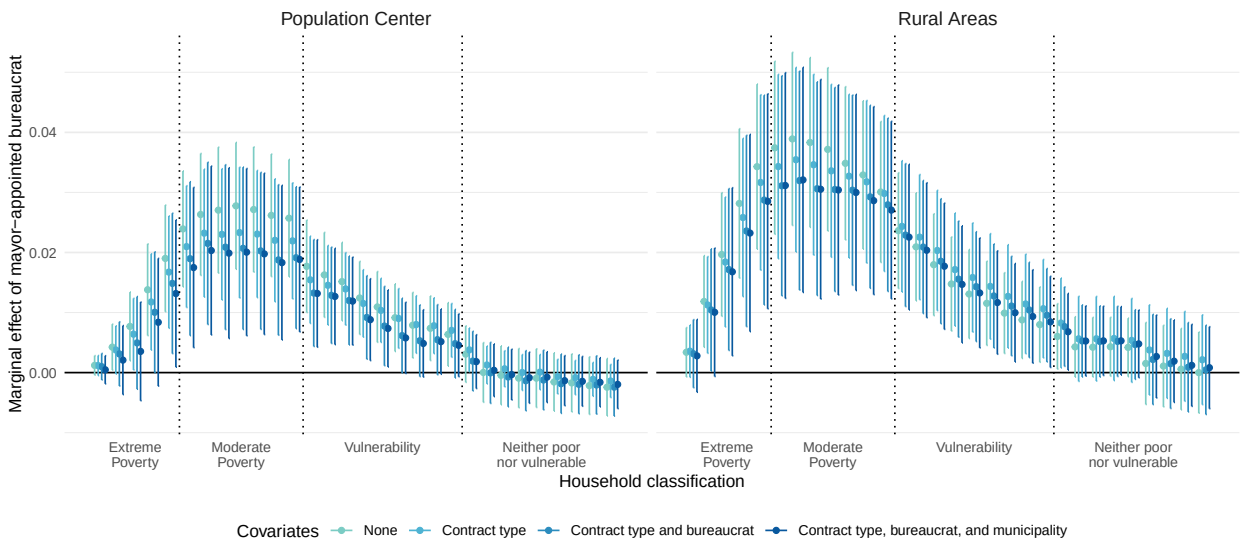


Figure 4: Marginal effects of a mayor-appointed bureaucrat on differences in the 2022 and 2021 cumulative mass functions of household classifications. Positive marginal effects indicate a shift toward a higher share of households being classified as poorer than a given category. 95% confidence intervals are calculated from heteroskedasticity-robust standard errors.

the conditional threshold—a 0.6 percentage point increase on a mean of 0.49. However, mayoral appointees are *less* likely to bunch at income thresholds than retained bureaucrats. These marginal effects are larger relative to (lower) baseline levels. Both estimates of bunching represent an 18-19% decrease on mean levels of reporting threshold values. One possible interpretation of these differences is that appointees' higher level of effort also yields more careful scoring—and/or distortion of—household incomes.

Finally, the Brazilian data allows me to link bureaucratic appointment to the outcomes experienced by households. Table A8 shows that within municipality, interview experience, and contract type, and controlling for bureaucrat attributes, households scored by appointees are 4.8 percentage points more likely to become Bolsa Familia beneficiaries. This is a sizeable effect on a baseline rate of 53% of households, and is equivalent to 22% of the within-municipality variation in the rate of enrollment. Moreover, Table A8 also provides evidence that this is not simply a consequence of finding poorer households: there is a smaller gap between the households that are reported as

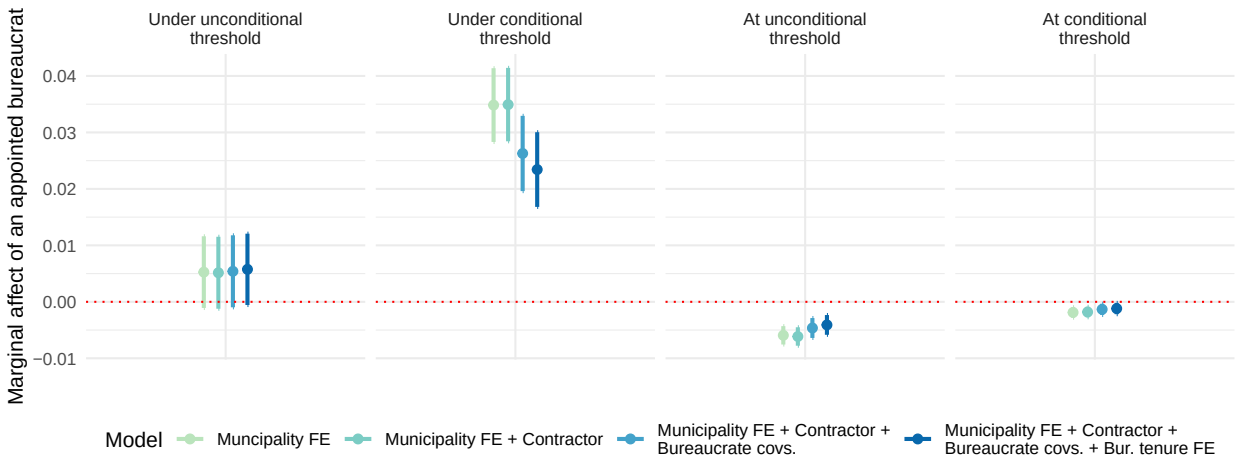


Figure 5: Estimates of the marginal effect of being appointed by the current mayor on income classification measures.

eligible (based on the conditional income qualification) and those enrolled in Bolsa Familia for mayoral appointees. The process of mapping bureaucrat-entered data into registration for the cash transfer occurs at the federal level and is beyond the scope of this paper, but these results suggests that a mayor's choice of bureaucrat and choice of oversight strategy and the bureaucrat's actions has distributive consequence in one of world's largest CCTs.

4.4 Learning

I argue that frequent replacement of bureaucrats may beget loyalists with greater alignment with the politician's policy preferences. However, this comes at a cost of limiting bureaucratic expertise. Here, I show that, consistent with model predictions for an environment with sufficient loyalists, as bureaucrats learn, they score more households as eligible for the transfer. This increase comes from the gains in the correct classification of eligible households by loyalists.[19] Figure 6 provides strongly suggestive evidence to this end. As bureaucrats become more experienced, the households they score are more likely to fall below relevant income thresholds and enter Bolsa Familia.

One limitation of Figure 6 is that it conflates learning with selective retention. If bureau-

[19]Technically, if the rate of eligiblity in a *municipality*, i.e. $F(\widehat{a}) > 0.5$, the technocrat should also increase the rate of eligible households as they learn.

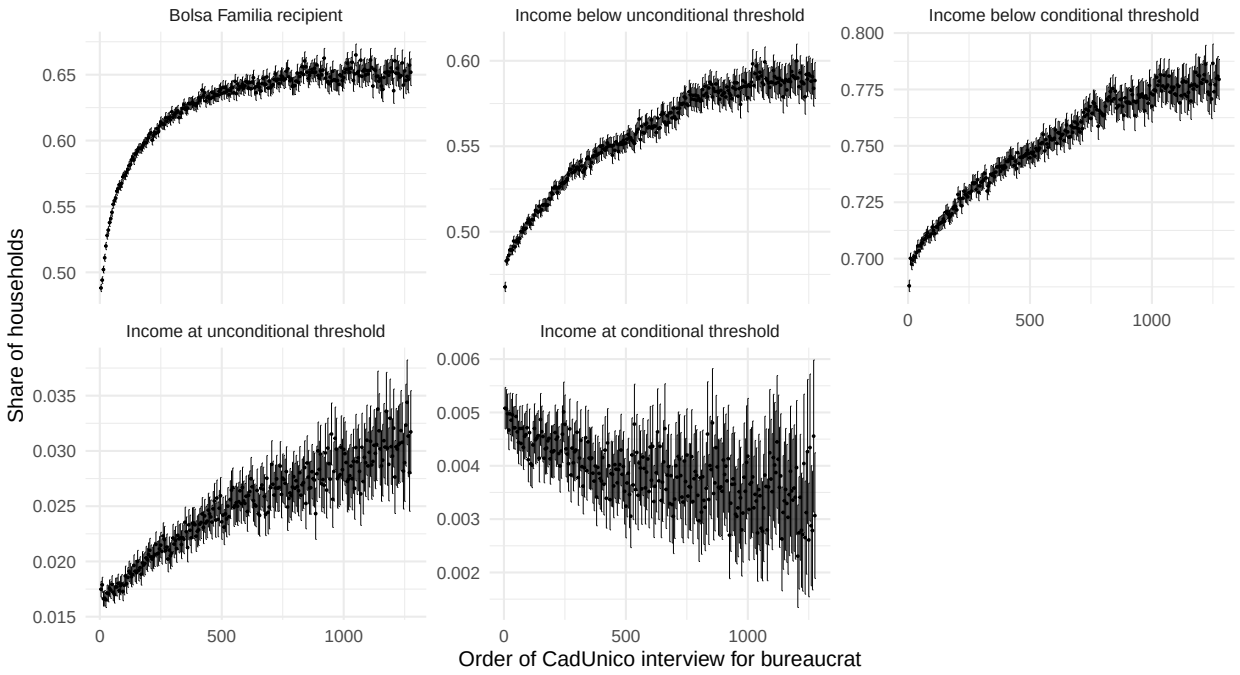


Figure 6: Changes in Bolsa Familia eligibility, qualifying incomes, and bunching as a function of the number of interviews conducted. Each estimate evaluates the mean over five-interview bins.

crats who report few qualifying households are fired or quickly transferred out of their position as CadUnico interviewers, this pattern could obtain even in the absence of learning. Figure suggests the patterns in Figure 6 are driven by both selection and learning, by comparing these trajectories as a function of how many interviews a bureaucrat conducted in total. To this end, I examine reported scores *within* bureaucrat, as a function of the order in Figure 7. Like Figure 6, the plotted estimates suggest that experience conducting interviews yields reports of lower incomes. However, the inclusion of individual bureaucrat fixed effects generally attenuates these estimates. Because the fixed effects help to parse out selection, thereby isolating learning, it suggests that both forces are at play to produce the patterns documented in Figure 6. This speaks to the importance of understanding how politicians appoint bureaucrats to administer these registers in the first place.

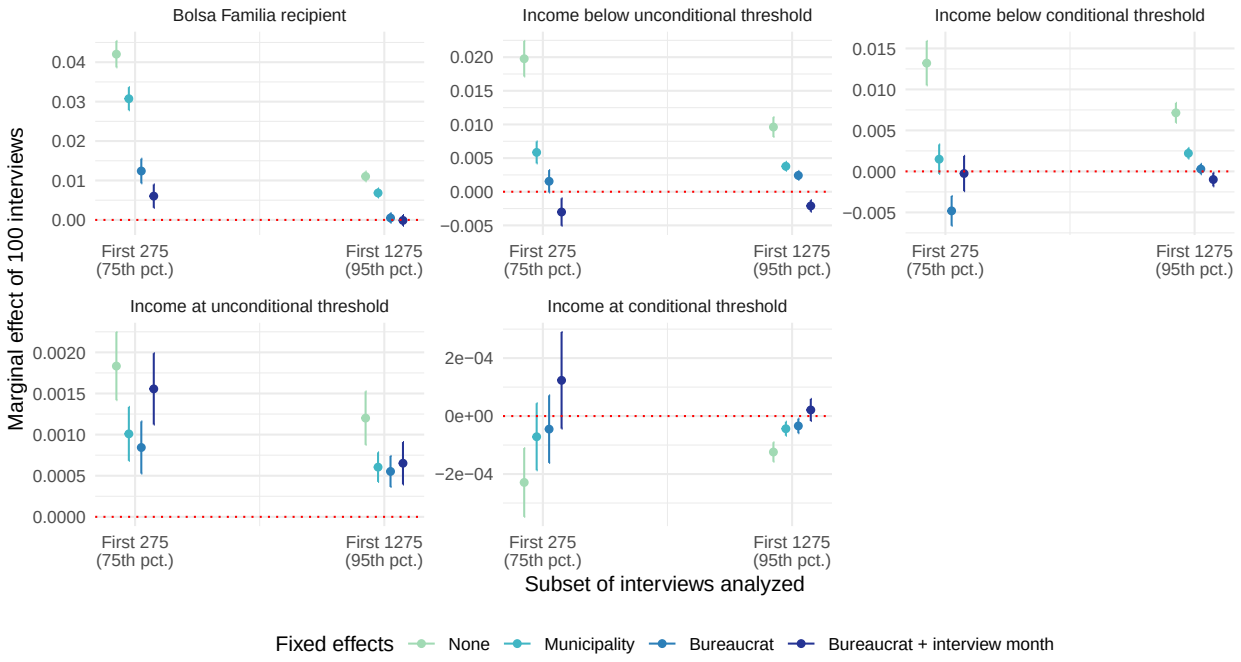


Figure 7: Estimates of the marginal effect of 100 interviews on individual bureaucrat's classification of households.

5 Implications for Social Register Design and Production

The empirical analysis takes as given the design of social policies by the central government. Within the model, these policies are given by the choice of how to quantify need/vulnerability, a_i, and the threshold for eligibility, $\widehat{a}$. However, the findings have implications for national government social policy.

Changing the classification scheme: In Colombia, the government has periodically changed the classification system (SISBÉN variants I-IV), in principle to respond to bloated rolls and/or changes in the mapping from measured assets to socioeconomic need. The standard interpretation of this strategy is that it: (1) prunes bloated rolls, to reduce the cost of these transfers; and (2) places a constraint on $\sigma_t^{\mathcal{M}}$, the bureaucrat's ability to classify ineligible households as eligible (Camacho and Conover, 2011). However, creating a new system constrain both types of bureaucrats' ability to correctly classify both types of households. If politicians are trading off aligned preferences

for expertise when selecting bureaucrats, as in the model, redesigning the classification system reduces a politician's incentive to retain an experienced bureaucrat. Thus, there is an implicit tradeoff between constraining knowledge of how to cheat the system for a personnel policy that selects on desire to cheat the classification system. In the current data, it is hard to isolate this effect. SISBÉN-IV was implemented in 2021, which coincides with the increased rate of hiring SISBÉN administrators documented in Figure 2. However, other events (i.e., the COVID-19 pandemic) and the use of personnel data a cross-section of administrators may also contribute to the observation of this increased slope.

Means testing: Social registers provide the data necessary for means testing social programs. Given the limits to the accuracy of these registers that I document, would the government be better off making the underlying programs universal? Means testing has two costs to the national government: direct costs for register maintenance and indirect costs related to the quality of the data. The direct costs consist of intergovernmental transfers to municipalities for the bureaucratic time to maintain these data and any time training these officials from afar. The indirect costs stem from the quality of the data produced. Errors of inclusion have clear costs to the national government (the cost of social benefits); errors of exclusion depend on the national government's internalization of the welfare of excluded households. A universal program costs less to administer (i.e., one need not construct and maintain a social register) and reduces drastically errors of exclusion. This becomes increasingly attractive as programs are targeted to larger shares of the population (i.e., $F(\widehat{a})$ is large) or Type-I errors of inclusion are more frequent. In other words, all else equal, universal programs may be relatively more attractive than means-testing in patronage-heavy personnel systems (high π_t) among register administrators.

6 Conclusion

Local governments and national governments often have distinct preferences over eligibility for means-tested social programs. Whereas the national government seeks to adhere to the means-

testing scheme it sets, local politicians would like to expand their rolls. Their efforts to pad the rolls are hampered to some extent by their reliance on bureaucrats to produce and report these data. I study how politicians select and monitor bureaucrats to produce these data. I show that in Brazil and Colombia, mayor-appointed register administrators exert greater effort—expanding the rolls—and report more poor/eligible households. Colombian survey data suggests that tighter oversight of appointees by political principals may play a role in inducing greater effort. Analysis of reporting trajectories in Brazil confirm that frequent turnover of appointees comes at a cost of expertise. These findings suggest that agency problems in local governments leave footprints in social registers, with implications for citizen access to programmatic social programs.

The general dynamics I describe are not unique to social registers. Agency problems within local governments are apt to affect a host of data collection processes. However, the implications of agency problems for data quality and policy outputs may be different outside the context of social registers. Specifically, in the case of social registers, agency problems insulate the national program (to varying degrees) from a local politician's desired level of political interference. Understanding the alignment of policy preferences between national policymakers, local principals, and local bureaucrats is important to understanding how related dynamics might influence other data collection processes and their policy consequences.

References

Akhtari, Mitra, Diana Moreira, and Laura Trucco. 2022. "Political Turnover, Bureaucratic Turnover, and the Quality of Public Services." *American Economic Review* 112 (2): 442–493.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102 (4): 1206–1240.

Angrist, Noam, Pinelopi Kougianou Goldberg, and Dean Jolliffe. 2021. "Why is Growth in Developing Countries So Hard to Measure?" *Journal of Economic Perspectives* 35 (3): 215–242.

Bowles, Jeremy. 2020. "The Limits of Legibility: How Distributive Conflicts Constrain State-Building." Working paper, Harvard University.

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.

Brierly, Sarah. 2020. "Unprincipled Principals: Co-opted Bureaucrats and Corruption in Ghana." *American Journal of Political Science* 64 (2): 209–222.

Brollo, Fernanda, Katja Kaufmann, and Eliana La Ferrara. 2020. "The Political Economy of Program Enforcement: Evidence from Brazil." *Journal of the European Economic Association* 18 (2): 750–791.

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.

Camacho, Adriana, Emily Conover, and Pablo Querubín. forthcoming. *Handbook on Social Protection*. chapter The Interaction of Social Protection Programs with Elections, as well as other Governance Challenges.

Carreri, Maria. 2021. "Can Good Politicians Compensate for Bad Institutions? Evidence from an Original Survey of Italian Mayors." *Journal of Politics* 83 (4): 1229–1245.

Carreri, Maria, and Julia Payson. 2023. "Local leaders and the pursuit of growth in US cities: the role of magerial skill." *Political Science Research and Methods* Forthcoming: 1–17.

Coady, David, Margaret Grosh, and John Hoddinnott. 204. *Targeting of Transfers in Developing Countries: Review of Lesons and Experiences*. Washington D.C.: World Bank.

Colonnelli, Emanuele, Mounu Prem, and Edoardo TEso. 2020. "Patronage and Selection in Public Sector Organizations." *American Economic Review* 110 (10): 3071–3099.

de la O, Ana L. 2013. "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico." *American Journal of Political Science* 57 (1): 1–14.

de la O, Ana L. 2015. *Crafting Policies to End Poverty in Latin America: The Quiet Transformation*. New York: Cambridge University Press.

Edmond, Chris. 2013. "Information manipulation, coordination, and regime change." *Review of Economic studies* 80 (4): 1422–1458.

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

Falleti, Tulia G. 2012. *Decentralization and Subnational Politics in Latin America*. New York: Cambridge University Press.

Frey, Anderson. 2012. "Do Reelection Incentives Improve Policy Implementation? Accountability versus Political Targeting." *Quarterly Journal of Political Science* 16: 35–69.

Frey, Anderson, and Rogerio Santarrosa. 2024. "The Politicization of Bureaucrats: Evidence from Brazil." *Journal of Politics* Forthcoming.

Garay, Candelaria. 2017. *Social Policy Expansion in Latin America*. New York: Cambridge University Press.

Garbiras-Díaz, Natalia, and Tara Slough. 2023. "The Limits of Decentralized Administrative Data Collection: Evidence from Colombia." Working paper, available at `http://www.taraslough.com/assets/pdf/decentralized_data.pdf`.

Garbiras-Díaz, Natalia, and Tara Slough. 2025. "Government Oversight and Inter-Institutional Legibility: Evidence from Colombia." Working paper.

Gulzar, Saad, and Benjamin Pasquale. 2017. "Politicians, Bureaucrats, and Development: Evidence from India." *American Political Science Review* 111 (1): 162–183.

Guriev, Sergei, and Daniel Treisman. 2019. "Informational autocrats." *Journal of Economic Perspectives* 33 (4): 100–127.

Hanna, Rema, and Benjamin O. Olken. 2018. "Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries." *Journal of Economic Perspectives* 32 (4): 201–226.

Hetzel, Alice M. 1997. U.S. Vital Statistics Systems: Major Activities and Developments, 1950-95. Technical report US Department of Health and Human Services Hyattsville, MD: .

Imai, Kosuke, Gary King, and Carlos Velasco-Rivera. 2020. "Do Nonpartisan Programmatic Policies Have Partisan Electoral Effects? Evidence from Two Large-Scale Experiments." *Journal of Politics* 82 (2): 714–730.

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

Joffe-Block, Jude, and Stephen Fowler. 2025. "USDA, DOGE demand states hand over personal data about food stamp recipients.".

Kain, Roger J. P., and Elizabeth Baigent. 1992. *The Cadastral Map in the Service of the State*. Chicago, IL: University of Chicago Press.

Kalaj, Jozefina, Daniel Rogger, and Ravi Somani. 2020. "Bureaucrat Time-Use and Productivity: Evidence from a Survey Experiment." *Working paper* .

Kertzer, Joshua D., and Jonathan Renshon. 2022. "Experiments and Surveys on Political Elites." *Annual Review of Political Science* 25: 529–550.

Labonne, Julien. 2013. "The Local Electoral Impacts of Conditional Cash Transfers: Evidence from a Field Experiment." *Journal of Development Economics* 104: 73–88.

Lee, Melissa M., and Nan Zhang. 2016. "Legibility and the Informational Foundations of State Capacity." *Journal of Politics* Journal of Politics (79): 1.

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

Lorentzen, Peter. 2014. "China's strategic censorship." *American Journal of political science* 58 (2): 402–414.

Manacorda, Marco, Edward Miguel, and Andrea Vigorito. 2011. "Government Transfers and Political Support." *American Economic Journal: Applied Economics* 3 (3): 1–28.

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).

Niedzwiecki, Sara. 2018. *Uneven Social Policies: The Politics of Subnational Variation in Latin America*. New York: Cambridge University Press.

Peterlevitz, Tiago. 2023. "Patronage Contracting." Working paper, available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3929169`.

Raffler, Pia. 2022. "Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from a Dominant Party Regime." *American Political Science Review* 116 (4): 1443–1459.

Rueda, Miguel B., and Nelson A. Ruiz. 2022. "How Do Electoral Outcomes Affect Campaign Contributions? The Role of Personal Loyalty and Investment Motives." Working paper available at `https://www.dropbox.com/s/is00murw518rls5/Rueda_Ruiz_Electoral_Outcomes.pdf?dl=0`.

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

Slough, Tara. 2022. "Squeaky Wheels and Inequality in Bureaucaratic Service Provision." Working paper available at `http://taraslough.com/assets/pdf/colombia_audit.pdf`.

Slough, Tara. 2024. "Bureaucratic Quality and Electoral Accountability." *American Political Science Review* 118 (4): 1931–1950.

Toral, Guillermo. 2024*a*. "How Patronage Delivers: Political Appointments, Bureaucratic Accountability, and Service Delivery in Brazil." *American Journal of Political Science* 68 (2): 797–815.

Toral, Guillermo. 2024*b*. "Turnover: How Lame-Duck Governments Disrupt the Bureaucracy and Service Delivery before Leaving Office." *Journal of Politics* 86 (4): 1348 – 1367.

Trinh, Minh. 2021. "Statistical Misreporting Debilitates Authoritarian Governance." *Working paper* .

Wallace, Jeremy L. 2016. "Juking the stats? Authoritarian information problems in China." *British Journal of Political Science* 46 (1): 11–29.

Zucco, Cesar. 2013. "When payouts pay off: Conditional cash transfers and voting behavior in Brazil 2002–10." *American Journal of Political Science* 57 (4): 810–822.