Bureaucrats Driving Inequality in Access: Experimental Evidence from Colombia

Supporting Information

Tara Slough*

January 17, 2020

Contents

*Ph.D. Candidate, Columbia University

A1 Theoretical Model: Extensions and Proofs

A1.1 Proofs of Propositions in the Main Text

Proof of Proposition 1: $\mathbb{E}[e_g^*]$ is the expectation of bureaucratic effort for citizens of group g. Recall that c_C is a realization of the random variable C_g with the cdf $F_g(\cdot)$. As such, the proportion of citizens that would complain if service is not granted is given by $F_g(\frac{b}{c_P})$. Assume that effort is interior for all citizens, $e_g^* < 1$. It is straightforward to calculate $\mathbb{E}[e_g^*]$ from Equation 7:

$$\mathbb{E}[e_g^*] = \frac{\mathbb{E}[\gamma_B^g]}{c_B} + \frac{r}{c_B c_P}\left[S + \mathbb{E}[\gamma_P^g] + F_g\left(\frac{b}{c_P}\right)\right] \tag{1}$$

where $\mathbb{E}[\gamma_B^g]$ and $\mathbb{E}[\gamma_P^g]$ are the expectations of the bureaucrat's and politician's tastes for group g, respectively. The expectation of bias between groups x and y in the aggregate is defined as $\Delta = \mathbb{E}[e_x^*] - \mathbb{E}[e_y^*]$ and is calculated:

$$\Delta = \frac{\mathbb{E}[\gamma_B^x] - \mathbb{E}[\gamma_B^y]}{c_B} + \frac{r}{c_B c_P}\left[\mathbb{E}[\gamma_P^x] - \mathbb{E}[\gamma_P^y] + F_x\left(\frac{b}{c_P}\right) - F_y\left(\frac{b}{c_P}\right)\right] \qquad \blacksquare \tag{2}$$

Proof of Proposition 2: Without loss of generality, assume that $\eta_S + \eta_Q > 0$. This implies that per citizen, the politician would devote more effort to monitoring the bureaucrat's service to group x when service is not provided. Note that if $\eta_B \geq 0$, the higher probability of oversight must increase bias. I solve for the magnitude of bureaucrat's taste-driven bias at which the magnitude (absolute value) of bias with and without oversight is equivalent. Comparison of the magnitude implies:

$$\frac{\eta_B}{c_B} + \frac{r(\eta_P + \eta_Q)}{c_B c_P} = -\frac{\eta_B}{c_B} \tag{3}$$

Simplifying:

$$\frac{r(\eta_P + \eta_Q)}{2c_P} = -\eta_B \tag{4}$$

This implies that for $\frac{r(\eta_P+\eta_Q)}{2c_P} > -\eta_B$, the magnitude of bias in effort is increased by the biased probability of oversight toward group x. For $\frac{r(\eta_P+\eta_Q)}{2c_P} = -\eta_B$, the magnitude of bias is not affected by oversight, and only when $\frac{r(\eta_P+\eta_Q)}{2c_P} < -\eta_B$, does the threat of oversight decrease bias. $\blacksquare$

Proof of Proposition 3: Comparative statics.

1. Differentiating Δ_O and Δ_B with respect to c_P yields:

$$\frac{\partial \Delta_B}{\partial c_P} = 0 \tag{5}$$

$$\frac{\partial \Delta_O}{\partial c_P} = \frac{r}{c_B}\frac{b(f_y(\frac{b}{c_P}) - f_x(\frac{b}{c_P})) - c_P(\eta_P + F_x(\frac{b}{c_P}) - F_y(\frac{b}{c_P}))}{c_P^3} \tag{6}$$

The sign of $\frac{\partial \Delta_O}{\partial c_P}$ depends on the shape of the densities $f_x(\cdot)$ and $f_y(\cdot)$. However, Δ_B does not vary in c_P while Δ_O may. Thus, $\frac{\partial \Delta}{\partial c_P} \neq 0$ indicates that $\Delta_O \neq 0$.

Further, note that $F_x(\frac{b}{c_B}) - F_y(\frac{b}{c_B}) \in [0,1]$, by the assumption $F_y(\cdot)$ FOSD $F_x(\cdot)$. This implies that $\eta_Q \in [0,1]$. As such, for a sufficient increase in c_P, Δ_O attenuates toward zero.

2. Differentiating Δ_B and Δ_O with respect to r yields:

$$\frac{\partial \Delta_B}{\partial r} = 0$$
$$\frac{\partial \Delta_O}{\partial r} = \frac{\eta_Q + \eta_P}{c_B c_P} \begin{cases} > 0 & \text{if } \eta_Q + \eta_P > 0 \\ < 0 & \text{if } \eta_Q + \eta_P < 0 \end{cases}$$

This implies that if $\frac{\partial \Delta}{\partial r} > 0$ if $\Delta_O > 0$ and $\frac{\partial \Delta}{\partial r} < 0$ if $\Delta_O < 0$.

3. Differentiating Δ with respect to η_Q yields $\frac{\partial \Delta}{\partial \eta_Q} = \frac{r}{c_B c_P} > 0$.

4. Differentiating Δ with respect to η_P yields $\frac{\partial \Delta}{\partial \eta_P} = \frac{r}{c_B c_P} > 0$.■

A1.2 Bias in Effort and Inequality in Outputs

Defining Inequality in Outputs: The model implies *inequality in outputs* as a second quantity measuring bias beyond *bias in effort*. Bias in effort is given by Proposition 1.

Inequality in outputs considers differences in expectation of the ultimate levels of service provision by group. Service provision, $\mathcal{S}$ is given by:

$$\mathcal{S}(e,q) = \begin{cases} e + (1-e)\frac{S+\gamma_P+1}{c_P} & \text{if } q = 1 \\ e + (1-e)\frac{S+\gamma_P}{c_P} & \text{if } q = 0 \end{cases}$$

I measure inequality in outputs as the difference in the expectation of service provision for each group. Define the expectation of service provision for a member of group g as:

$$\begin{aligned} \mathbb{E}[\mathcal{S}_g] =& F_g\left(\frac{b}{c_P}\right)\mathcal{S}_g(e_g^*, 1) + \left(1 - F_g\left(\frac{b}{c_P}\right)\right)\mathcal{S}_g(e_g^*, 0) \\ =& \frac{S + \gamma_P^g}{c_P} + \frac{e_g^*(c_P - S - \gamma_P^g)}{c_P} + F_g\left(\frac{b}{c_P}\right)\frac{1 - e_g^*}{c_P} \end{aligned}$$

Inequality in outputs for an individual from each group x and y is therefore defined as:

$$\begin{aligned} \mathbb{E}[\mathcal{S}_x] - \mathbb{E}[\mathcal{S}_y] =& \frac{S + \mathbb{E}[\gamma_P^x]}{c_P} + \frac{e_x^*(c_P - S - \mathbb{E}[\gamma_P^x])}{c_P} + F_x\left(\frac{b}{c_P}\right)\frac{1 - e_x^*}{c_P} - \\ & \left(\frac{S + \mathbb{E}[\gamma_P^y]}{c_P} + \frac{e_y^*(c_P - S - [\gamma_P^y])}{c_P} + F_y\left(\frac{b}{c_P}\right)\frac{1 - e_y^*}{c_P}\right) \\ =& \frac{-r(n_Q + 2\mathbb{E}[\gamma_P^y])(2S + \eta_Q + 2F_y(\frac{b}{c_P}) + \eta_P + \mathbb{E}[\gamma_P^2] - c_P)}{c_P^2} + \\ & \frac{\mathbb{E}[\gamma_B^x](-F_x(\frac{b}{c_P} - \mathbb{E}[\gamma_P^x]) + \mathbb{E}[\gamma_B^y](F_y(\frac{b}{c_P}) + \mathbb{E}[\gamma_P^y]) - S\eta_B + c_P\eta_B + c_B(\eta_Q + \eta_P)}{c_P} \end{aligned}$$

A1.3 Endogenous Requests for Service

In the text, citizen requests of service are treated as exogenous. In this extension, I consider equilibrium levels of effort when citizens pay a cost to request a service. This adds a first step to the sequence presented in the main text, in which citizens request the service or not, denoted $R \in \{0, 1\}$. Preceding the bureaucrat's decision to exert effort, this extension includes:

1. Citizen chooses whether or not to request service.

The cost of requesting a service, ξ, is proportional to c_C. The cost for requesting service is thus ξc_C.

Denote the ex-ante expected utility for citizens that would complain and citizens that would not complain, as $\mathcal{C}$ and $\mathcal{N}$:

$$\mathbb{E}[U_C^{\mathcal{C}}] = eb + (1-e)\left(b\frac{S+\gamma_P^g+1}{c_P} - c_C\right) - \xi c_C$$
$$\mathbb{E}[U_C^{\mathcal{N}}] = eb + (1-e)\left(b\frac{S+\gamma_P^g}{c_P}\right) - \xi c_C$$

Given that the citizen will only complain if $c_C < \frac{b}{c_P}$, clearly $\mathbb{E}[U_C^{\mathcal{C}}] \geq \mathbb{E}[U_C^{\mathcal{N}}]$. The citizen's expected utility is decreasing linearly in c_C. For interior e_g^*:

$$\frac{\partial \mathbb{E}[U_C^{\mathcal{C}}]}{\partial c_C} = -1 + \frac{\gamma_B^g}{c_B} + \frac{r(\gamma_P^g + S + 1)}{c_B c_P} - \xi < 0, \qquad \frac{\partial^2 \mathbb{E}[U_C^{\mathcal{C}}]}{\partial c_C^2} = 0$$
$$\frac{\partial \mathbb{E}[U_C^{\mathcal{N}}]}{\partial c_C} = -1 + \frac{\gamma_B^g}{c_B} + \frac{r(\gamma_P^g + S)}{c_B c_P} - \xi < 0, \qquad \frac{\partial^2 \mathbb{E}[U_C^{\mathcal{N}}]}{\partial c_C^2} = 0$$

There are three cases, that I characterize in terms of cut points in the cost of requesting service, ξ. I denote these cut points as $\bar{\xi}$ and $\underline{\xi}$ where $0 \leq \underline{\xi} \leq \bar{\xi}$. Assume, without loss of generality, that the upper support of f_g exceeds $\frac{b}{c_P}$.[1]

Case 1: $\xi \geq \bar{\xi}$: There exists some cost $\bar{\xi} \geq 0$ which makes a citizen for whom $c = \frac{b}{c_P}$ is indifferent between requesting the service and not. Substituting e_g^* and $c_C = \frac{b}{c_P}$ into $\mathbb{E}[U]_C^{\mathcal{C}} = 0$ and solving for $\bar{\xi}$ yields:

$$\bar{\xi} = S + \gamma_P^g + \frac{(c_P - S - \gamma_P^g)(r(S + \gamma_P^g + 1) + c_P \gamma_B^g)}{c_B c_P}$$

Case 2: $\xi \in (\underline{\xi}, \bar{\xi})$: Denote the upper support of f_g as $\bar{c}_C$. There exists some threshold $\tilde{c}_C \in (\frac{b}{c_P}, \bar{c}_C)$. A citizen for whom $c_C \leq \tilde{c}_C$ will request service, while a citizen for whom $c_C > \tilde{c}_C$ will not.

Case 3: $\xi \leq \underline{\xi}$: There exists some cost $\underline{\xi} \geq 0$ that makes a citizen with cost $\bar{c}_C$ indifferent between requesting the service and not. Substituting e_g^* and $c_C = \bar{c}_C$ into $\mathbb{E}[U]_C^{\mathcal{N}}$ and solving for $\underline{\xi}$ yields:

[1] If this assumption does not hold, the solution collapses to Case 1.

$$\underline{\xi} = \frac{r(S+\gamma_P^g)+C_P\gamma_B^g}{c_P c_B} + \frac{b(c_B c_P(S+\gamma_P^g)+(c_P-S-\gamma_P^g)(r(S+\gamma_P^g)+c_P\gamma_B^g))}{c_B c_P^2 \bar{c}_C} - 1$$

Proposition A1. *Equilibrium effort with costly requests. If, within a group g, the bureaucrat's tastes, politician's tastes, and citizen costs are independent, as the cost of requesting service, ξ, increases, the proportion of citizens requesting service decreases. The expectation of equilibrium effort exerted by the bureaucrat on each request increases (piecewise) in ξ.*

Proof:
To characterize the share of citizens in a group g that request service and the expectation of equilibrium service provided, I assume that $Cov(\gamma_B^g, \gamma_P^g) = 0, Cov(\gamma_B^g, C_g) = 0$, and $Cov(\gamma_P^g, C_g) = 0$ and consider each case.

Case 1: $\xi > \bar{\xi}$: For any $\xi > \bar{\xi}$, conditional on requesting service ($R = 1$), the expectation of equilibrium effort for group g is:

$$\mathbb{E}[e_g^*|R=1] = \frac{\mathbb{E}[\gamma_B^g]}{c_B} + \frac{r(S+1+\mathbb{E}[\gamma_P^g])}{c_B c_P}$$

and the proportion of group g that requests the service is given by:

$$F_g\left(\frac{b(c_B c_P(S+\gamma_P^x+1)+(c_P-(S+\gamma_P^x+1))(r(S+\gamma_P^x+1)+c_P\gamma_B^x))}{c_P(-r(S+\gamma_P^x+1)-c_P\gamma_B^x+c_B c_P(1+\xi))}\right) \leq F_g\left(\frac{b}{c_P}\right)$$

Case 2: $\xi \in (\underline{\xi}, \bar{\xi}]$: For any $\xi \in (\underline{\xi}, \bar{\xi})$, conditional on requesting service ($R = 1$), the expectation of equilibrium effort for group g is:

$$\mathbb{E}[e_g^*|R=1] = \frac{\mathbb{E}[\gamma_B^g]}{c_B} + \frac{r\left(S+\frac{F_g(\frac{b}{c_P})}{F_g(\tilde{c}_C)}+\mathbb{E}[\gamma_P^g]\right)}{c_B c_P}$$

Because $\tilde{c}_C \in (\frac{b}{c_P}, \bar{c}_C)$, $\frac{F_g(\frac{b}{c_P})}{F_g(\tilde{c}_C)} \in (F_g(\frac{b}{c_P}), 1)$. The proportion of group g that requests the service is given by $F_g(\tilde{c}_C)$.

Case 3: $\xi < \underline{\xi}$: For any $\xi \leq \underline{\xi}$, conditional on requesting service, ($R = 1$), the expectation of equilibrium effort for group g is:

$$\mathbb{E}[e_g^*|R=1] = \frac{\mathbb{E}[\gamma_B^g]}{c_B} + \frac{r(S+F_g(\frac{b}{c_P})+\mathbb{E}[\gamma_P^g])}{c_B c_P}$$

This is equivalent to the equilibrium with exogenous (costless) requests characterized in the main text. As was assumed in the paper, all citizens request service.

Collectively this analysis implies that the proportion of citizens requesting the service decreases in ξ and that the expectation of equilibrium effort, conditional on requesting the service, increases (piecewise) in ξ. ■

A1.4 Bias in Effort and Inequality in Outputs

The mapping from bias in effort to inequality in outputs depends on the composition of the bias in effort (between tastes and complaint-driven bias). For this analysis, make two simplifying assumptions. First, within a group g, the distribution of tastes and costs is independent, i.e. $Cov(\gamma_B^g, \gamma_P^g) = 0, Cov(\gamma_B^g, C_g) = 0$, and $Cov(\gamma_P^g, C_g) = 0$. Second, normalize the tastes and likelihood of complaint for group y to 0, i.e. $\mathbb{E}[\gamma_P^y] = 0$, $\mathbb{E}[\gamma_B^y] = 0$, and $F_y(\frac{b}{c_P}) = 0$.

Bias in effort: As in Proposition in the main text, bias in effort is given by:

$$\Delta = \frac{\eta_B}{c_B} + \frac{r(\eta_P + \eta_Q)}{c_B c_P} \tag{7}$$

Inequality in outputs: Following the definition in Section A1.2, inequality in outputs simplifies to:

$$\mathbb{E}[\mathcal{S}_x] - \mathbb{E}[\mathcal{S}_y] = \frac{(\eta_P + \eta_Q)(rc_P - r(2S + \eta_P + \eta_Q) + c_P(c_B - \eta_B)) + \eta_B c_P(c_P - S)}{c_P^2 c_B} \tag{8}$$

$$= \underbrace{\frac{\eta_B}{c_B} + \frac{r(\eta_P + \eta_Q)}{c_P c_B}}_{\Delta} + \frac{(\eta_P + \eta_Q)(-r(2S + \eta_P + \eta_Q) + c_P(c_B - \eta_B)) - c_P(\eta_B S)}{c_P^2 c_B} \tag{9}$$

While the sign of the remaining term in the sum is ambiguous, it can be shown that within the parametric assumptions of the model, bias in effort Δ and inequality in outputs $\mathbb{E}[\mathcal{S}_x] - \mathbb{E}[\mathcal{S}_y]$ must share the same sign. Rewriting:

$$\mathbb{E}[\mathcal{S}_x] - \mathbb{E}[\mathcal{S}_y] = \frac{\eta_B(c_P - 1 - S - \eta_Q - \eta_P)}{c_B c_P} + \frac{(\eta_P + \eta_Q)(c_P(c_B + r) - r(2S - \eta_P - \eta_Q))}{c_P^2 c_B}$$

Because $c_P > S + 2$, $r \in (0, 1)$, and $c_B > 1$, the final term must maintain the same sign as $\eta_Q + \eta_P$.

Remark A1: *Given a neutral bureaucrat, $\eta_B = 0$, or a bureaucrat favoring the same group as that favored by oversight, $sign(\eta_B) = sign(\eta_P + \eta_Q)$, bias in effort favoring group g is sufficient to ensure that inequality in outputs favor g.*

A1.5 Implications for Complaint Rates in the Observational Data

Here I derive the expected rate of complaint per the baseline model in which the request for service is exogenous. Assume that $Cov(\gamma_B^g, C_g) = 0$, and $Cov(\gamma_P^g, C_g) = 0$. The rate of complaint by group is thus given by:

$$\underbrace{F_g\left(\frac{b}{c_P}\right)}_{\text{Share ``complainers''}} \underbrace{\left[1 - \left(\frac{\mathbb{E}[\gamma_B^g]}{c_B} + \frac{r}{c_B c_P}(\mathbb{E}[\gamma_P^g] + S + 1)\right)\right]}_{\text{Share of ``complainers'' receiving service}}$$

The difference in rate of complaint, between groups x and y is thus:

$$\underbrace{\left[F_x\left(\frac{b}{c_P}\right) - F_y\left(\frac{b}{c_P}\right)\right]\left(\frac{c_B c_P - r(S+1)}{c_b c_P}\right)}_{\geq 0} - F_x\left(\frac{b}{c_P}\right)\left(\frac{\mathbb{E}[\gamma_B^x]}{c_B} + \frac{rE[\gamma_B^x]}{c_B c_P}\right) + F_y\left(\frac{b}{c_P}\right)\left(\frac{\mathbb{E}[\gamma_B^y]}{c_B} + \frac{rE[\gamma_B^y]}{c_B c_P}\right)$$

In equilibrium, we should see more complaints from the group with a higher likelihood of complaint if:

$$\left[1 - \frac{F_y(\frac{b}{c_P})}{F_x(\frac{b}{c_P})}\right]\left(\frac{c_B c_P - r(S+1)}{c_b c_P}\right) + \frac{F_y(\frac{b}{c_P})}{F_x(\frac{b}{c_P})}\left(\frac{\mathbb{E}[\gamma_B^y]}{c_B} + \frac{rE[\gamma_B^y]}{c_B c_P}\right) > \frac{\mathbb{E}[\gamma_B^x]}{c_B} + \frac{rE[\gamma_B^x]}{c_B c_P}$$

The implication of this expression is that if the rate of observed complaint is higher for group x (with lower costs of complaint), it must be the case that the tastes of the bureaucrat and politician must not favor group x by too large of a magnitude. If this were the case, the difference in the rate at which citizens of group x receive service over those from group y would yield a higher rate of complaint from group y given the lower rate of service provision.

A2 Cross-National Data

A2.1 Survey Data

I refer to rates of contact with street-level bureaucrats using cross-national survey data from LAPOP AmericasBarometer (2014) and AfroBarometer (2016). The sample of countries includes all surveyed countries in Latin America, the Caribbean, and Africa with a Polity IV score above 0 in 2015. This includes both hybrid and democratic regimes.

Figure A1 depicts the proportion of citizens reporting that they sought service from a street-level bureaucrat in the last twelve months.

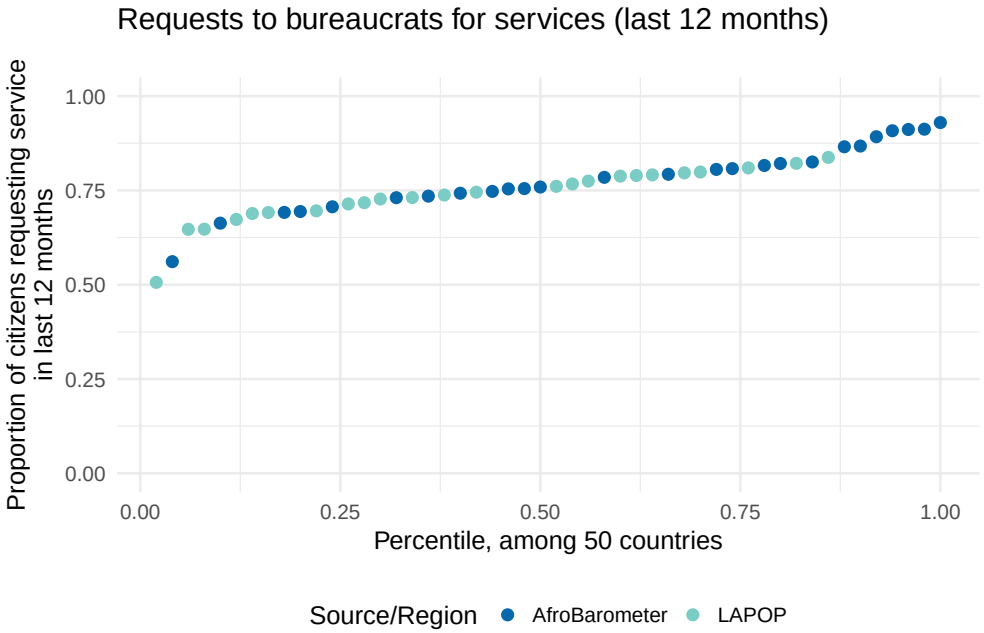


Figure A1: The proportion of survey respondents that sought government services within the last year. Data is from AmericasBarometer (LAPOP) (2014) and AfroBarometer (2016).

A2.1.1 AmericasBarometer (2014)

The measure of contact with street-level bureaucrats is constructed from four questions, as follows. Note that these questions are two parts; the second part of the question asks about petty corruption (bribes). If the answer indicates that the individual "had dealings" with the relevant entity, I code this as having sought a service.

1. **EXC11** "In the last twelve months, did you have any official dealings in the municipality/local government?"
2. **EXC14** "In the last twelve months, did you have any dealings with the courts?"
3. **EXC15** "Have you used any public health services in the last twelve months?"

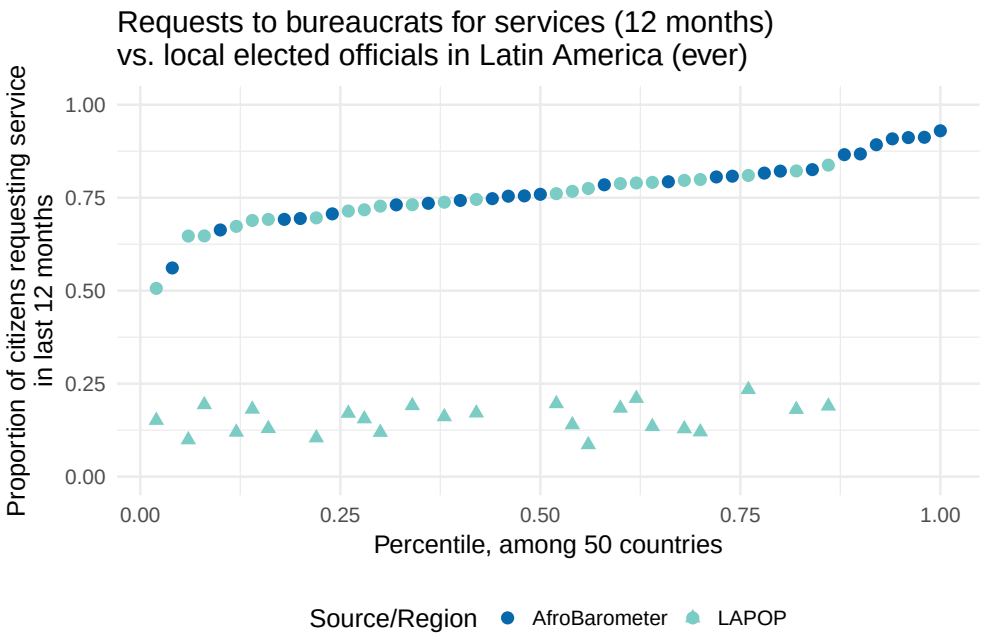


Figure A2: The proportion of survey respondents that sought government services within the last year (circles) compared to the proportion of survey respondents from Latin America that have ever sought the services of a municipal politician.

4. **EXC16** "Have you had a child in school in the last twelve months?"

The proportion of individuals "seeking service" is the proportion of individuals that answered affirmatively to any of these questions.

The countries included in the graph are: Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Trinidad and Tobago, Uruguay, and Venezuela.

To justify the claim that interactions with street-level bureaucrats are the "most common form" of citizen-government interaction, I compare the rate of individuals seeking a service in the last twelve months (circles) to those citizens who have ever sought assistance from an elected official (△'s) in Figure A2. The latter are calculated from the question:

CP4a: "In order to solve your problems have you ever requested help or cooperation from a local public official or local government: for example, a mayor, municipal council, councilman, provincial official, civil governor or governor?"

Figure A2 suggests that the proportions engaging bureaucrats in the last 12 months far exceeds the proportion that have *ever* engaged a local elected official (triangles).

A2.1.2 AfroBarometer (2016)

The measure of contact with street-level bureaucrats is constructed from five (largely parallel) questions, as follows:

1. **Q55A** " In the past 12 months have you had contact with a public school?"
2. **Q55C** "In the past 12 months have you had contact with a public clinic or hospital?"
3. **Q55E** "In the past 12 months have you tried to get an identity document like a birth certificate, driver's license, passport or voter's card, or a permit, from government?"
4. **Q55G** "In the past 12 months have you tried to get water, sanitation or electric services from government?"
5. **Q55K** "In the past 12 months have you had contact with the courts?"

The proportion of individuals "seeking service" is the proportion of individuals that answered affirmatively to any of these questions.

The countries included in the graph are: Algeria, Benin, Botswana, Burkina Faso, Cabo Verde, Gabon, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritius, Mozambique, Namibia, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Tanzania, Tunisia, Zambia, and Zimbabwe.

AfroBarometer includes an assessment of the difficulty of obtaining service, conditional on having sought it. For these measures, I consider assessments of public administrators – those disbursing identity documents or working at the tax authority (questions Q55E, Q70A, and Q70B).[2] I average perceptions across the three questions. Where answers are missing, I average the assessments that are reported. I examine the bivariate relationship between this assessment and education graphically in Figure A3. Citizens perceive services to be relatively difficult to access. This difficulty declines somewhat in educational attainment.

A2.2 World Bank Worldwide Governance Indicators

The World Bank Worldwide Governance Indicators (WGI) allows for cross-national characterizations of governance, broadly conceived. Some of these measures measure bureaucratic outputs. The intent in providing this information is simply to demonstrate where Colombia ranks globally in standard measures of governance. Using data from 2016 (the most recent WGI data) evaluate Colombia's rank, as a percentile, among:

- All countries in the World Bank WGI data ($n = 214$).
- All Spanish, French, and Portuguese-speaking countries in Latin America ($n = 19$).
- All OECD countries ($n = 37$). Note that Colombia joined the OECD in July 2018.

[2]The tax questions are as follows. **Q70A**:"Based on your experience, how easy or difficult is it to obtain the following services from government? Or do you never try and get these services from government: To find out what taxes and fees you are supposed to pay to the government?" **Q70B**:"Based on your experience, how easy or difficult is it to obtain the following services from government? Or do you never try and get these services from government: To avoid paying the income or property taxes that you owe to government?"

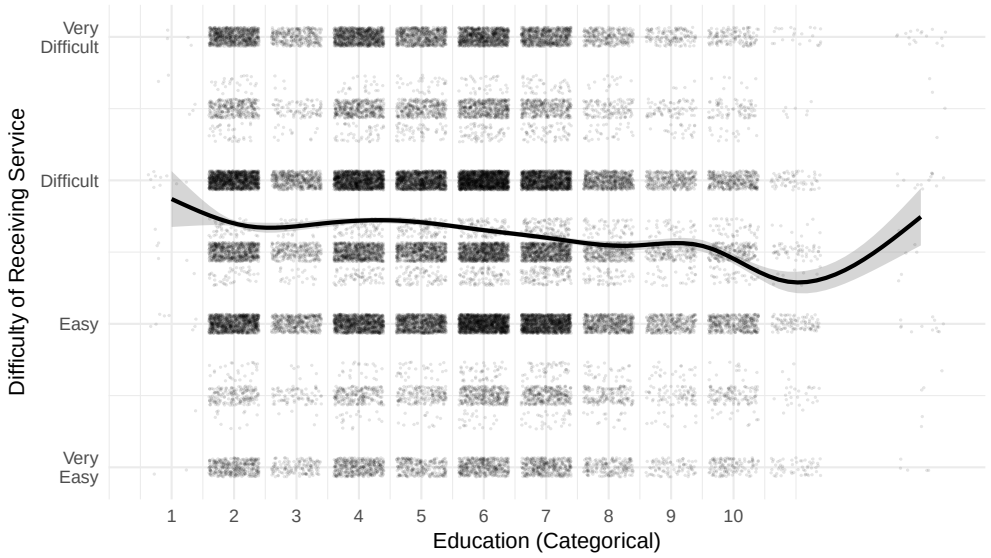


Figure A3: Perceptions of the difficulty of obtaining services from public administrators by level of educational attainment. $n = 50,758$.

Table A1 shows Colombia's rank among the three comparison groups for each of three indicators. While Colombia unsurprisingly performs quite poorly relative to all reference groups on the "Political Stability and Absence of Violence/Terrorism" measure, the other indicators which are plausibly more relevant measures of bureaucratic outputs. In general, Colombia generally performs somewhere around the median of all countries, in the top tercile of Latin American countries, and the bottom decile of OECD countries.

	Colombia's percentile among ...		
	World	Latin America	OECD
Indicator	($n = 214$)	($n = 19$)	($n = 37$)
Political Stability and Absence of Violence/Terrorism	13.81	5.56	2.78
Rule of Law	41.35	72.22	5.56
Control of Corruption	44.23	72.22	5.56
Voice and Accountability	49.75	55.56	5.56
Government Effectiveness	54.33	66.67	2.78
Regulatory Quality	67.31	77.78	11.11

Table A1: Colombia's rankings on each World Bank World Governance Indicator (2016) as a percentile within the relevant comparison group.

A3 Original Survey of Bureaucrats

The manuscript cites one descriptive finding from an original survey of street-level bureaucrats in *alcaldiías* in Bogotá and Cundinamarca. The details of the survey are as reported here. The survey was conducted in October and November, 2016. Two parallel surveys were conducted: one of citizens awaiting service and one of bureaucrats providing service in select entities. The surveys were conducted in:

- *Alcaldías*: local *alcaldías* in Bogotá and municipal *alcaldías* in Cundinamarca
- CADES/SUPERCADES: These are District (Bogotá) Centers for Public Service where citizens can seek many public services.
- Local offices of the Registraduría Nacional del Estado Civil
- Local notaries (public/private).

The relevant sample cited in the paper includes 73 surveys of bureaucrats from 14 *alcaldías*. The *alcaldías* were purposefully sampled but the timing of the visit was randomly assigned. The sampling included the first 5-8 bureaucrats that we encountered starting at reception (e.g. street level bureaucrats) that were willing to take the survey. In this sense, the sample is not random, but encompasses street-level bureaucrats in these entities present at the time of the survey.

The relevant question cited in the descriptive statistic in the paper was an open response question, enumerated and translated as follows:

- "¿Si usted tomara una decisión que su supervisor no apoyara, cuál sería la consecuencia?"
- Translation: "If you made a decision that your supervisor did not support, what would be the consequence?"

The responses ranged from verbal admonishment to more formal admonishments (in the form of a memorandum) to unwillingness to renew a contract (contractors only).

A4 Bogotá Complaint Data

Data on formal complaints from Bogotá is collected by the Veeduría Distrital, an oversight organ of the city government. Data is available at `tablerocontrolciudadano.veeduriadistrital.gov.co:3838/BogotaDashboard/`. The data consist of 464,387 PQRS petitions submitted to city entities in Bogotá between January 1, 2017 and June 30, 2018. PQRS stands for "*peticiones, quejas, reclamos, y sugerencias*," translated "petitions, complaints, [another word for] complaints, and suggestions)." These comprise represent formal written requests, not verbal or informal complaints. Note that the per capita rate of PQRS submission during this period is 5.68%, or one submission per ≈17.5 people.[3]

The PQRS are characterized by type, as in Table A2. Note that there are more words for complaints in Spanish than in English. I focus on the first three categories (the complaints) in the subsequent analysis $n = 440,803$.

PQRS Type	Translation	n	Proportion
Denuncia	Report (of complaint)	2,501	0.005
Queja	Complaint	99,302	0.214
Reclamo	Complaint	339,000	0.730
Sugerencia o Felictación	Suggestion or congratulation	23,584	0.051

Table A2: PQRS submitted in Bogotá, January 2017-June 2018. The type designation is made by the Veeduria (or receiving entities). Translations by author.

Of the complaints, 63,330 were registered by *alcaldías locales*, the entities audited in the experiment. Other complaints were directed to district-wide entities. To assess the correlation between class and propensity to complain, I examine the relationship between the relative wealth of a locality and the per-capita rate of complaint submission. To measure the wealth of a locality, I examine the average *estrato* (strata) of all residential properties. Strata range from 1 (very low/*bajo bajo*) to 6 (very upper/*alta alta*). While these zoning designations are technically made to properties, citizens identify *estrato* with class. Equating the two implies an assumption that lower-class Bogotanos are priced out of rich neighborhoods/dwellings and few middle- and upper-class Bogotanos choose to live substantially below their means. The marked degree of differentiation of localities in average *estrato* implies high levels of residential segregation.

Figure A4 plots the rate of complaint submission by wealth of localities. There is a clear positive relationship between the wealth of the locality and the rate of complaint submission. This occurs despite the fact that service is believed to be better in wealthy localities, suggesting that this analysis *understates* the relationship between class and propensity to complain. The outlier, La Candelaria, is a very small locality in the center of Bogotá with a vastly disproportionate tourist/foreigner ("ex-pat") population. To the extent that "ex-pats" (foreigners from rich countries) choose to live in a locality with a relatively low *estrato*, resident wealth is understated by the *estrato* designation in La Candelaria.

[3]Calculated based on an estimated population of 8,181,047 residents.

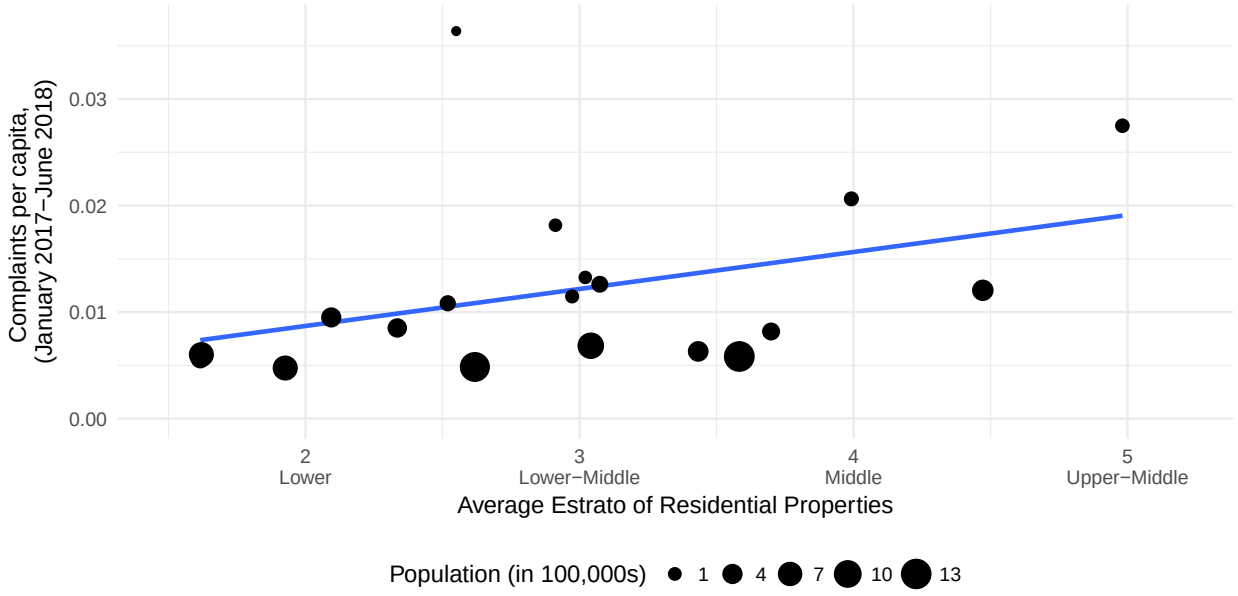


Figure A4: Rate of complaints filed by locality in Bogotá by average wealth of the locality.

A5 Demographic Data Related to Bias Treatments

Colombia's last national census was conducted in 2005. A new census will be conducted this year (2018). Given the vintage of the data, I use population projections where relevant and available. Otherwise, this section describes historical trends through 2005. All aggregate data in this section comes from the Departamento Administrativo Nacional de Estadística and all microdata comes from IPUMs.

A5.1 Regional Accents

In this section, I describe the geographic coverage and prevalence of the three accents utilized in the audit experiment. Figure A5 shows the departments to which these (generalized) accents are native. These are among the most densely populated regions of Colombia.

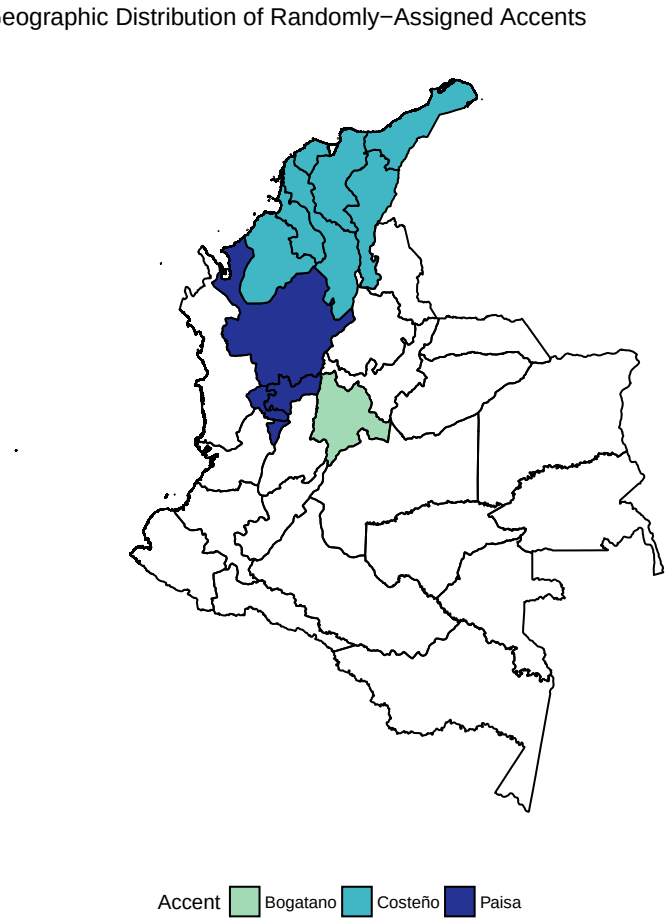


Figure A5: Geographical coverage of the Bogotano, Costeño, and Paisa accents at the departmental level. Note that the map of the Bogotá accent includes the department of Cundinamarca. Some portion of Cundinamarca's population speaks with a different accent (Cundinaboyacense).

I approximate the number of speakers of each accent in Colombia in Figure A6.[4] To approximate these

[4]Note that several of the accents are also spoken in neighboring countries, e.g. the Llanero accent in Venezuela. These counts

quantities, I consider the most widely-spoken accent in each department. I aggregate the projected population (2017) by DANE for each department and sum across the departments associated with each accent. This exercise indicates that nearly 60% of the Colombian population speaks one of the three regional accents utilized in the audit experiment.

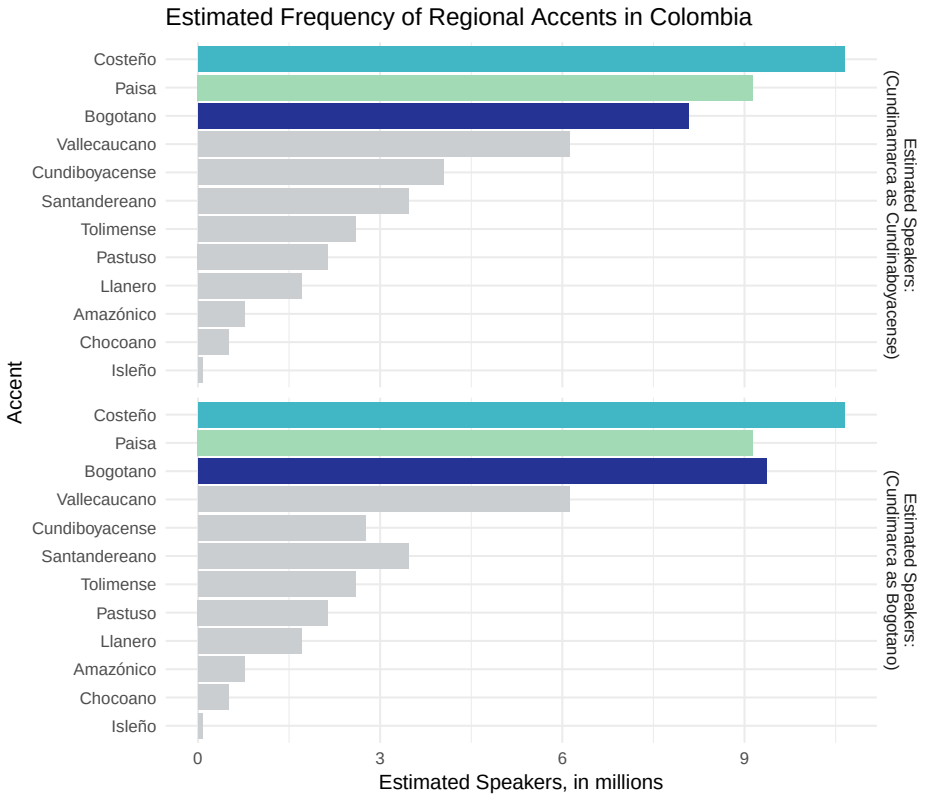


Figure A6: Approximate number of speakers of Colombia's twelve major accents. The graph shows that the most widely-spoken accents are the Bogotano, Costeño, and Paisa accents used in the experiment. Approximately 60% of the Colombian population speaks one of these three accents. The panels differ in the classification of the accent in the department of Cundinamarca (Bogotano or Cundinaboyacense).

A5.2 Socioeconomic Class

Detailed data on socioeconomic class is not available in Colombia. As such, I present the distribution of Colombians by class as per the 2005 census in Figure A7.

only include speakers of the accent in Colombia.

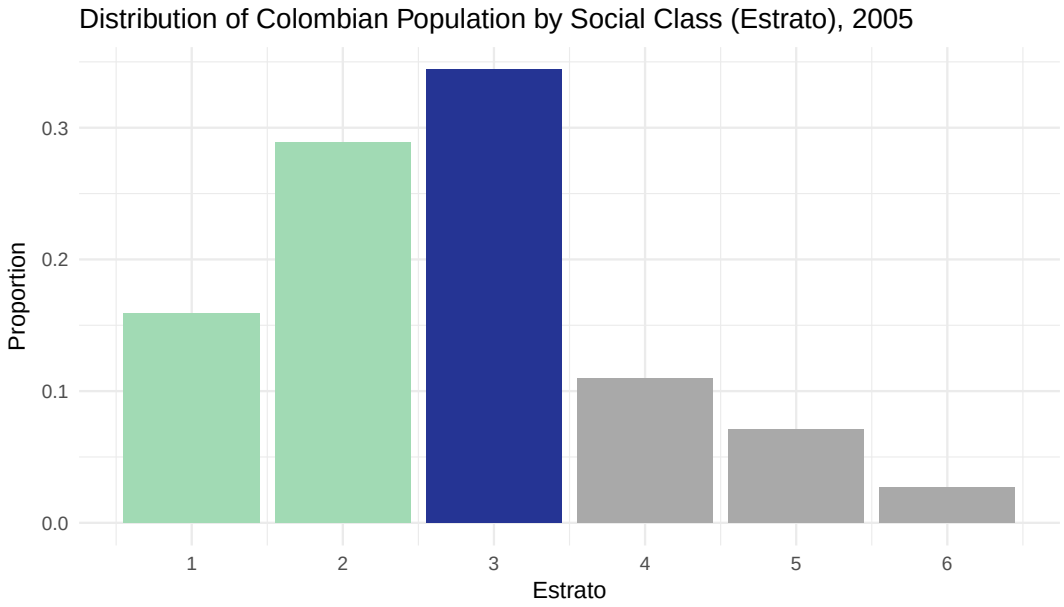


Figure A7: Distribution of classes in the 2005 Colombian census. The social class of the treatments are denoted by the colored bars.

A5.3 Migration

In Table A3, I place lower bounds on the proportion of Colombians that have (a) ever migrated and (b) migrated within the last five years using census microdata subsamples available from IPUMs. I bound the share of Colombians that have ever migrated by examining the share of individuals residing in their municipality of birth. Note that this is a lower bound as reverse-migrants will not be counted as migrants. I bound the share of individuals who report migrating to a different geographical unit in the past five years. This similarly does not count reverse-migrants and should be considered a lower bound.

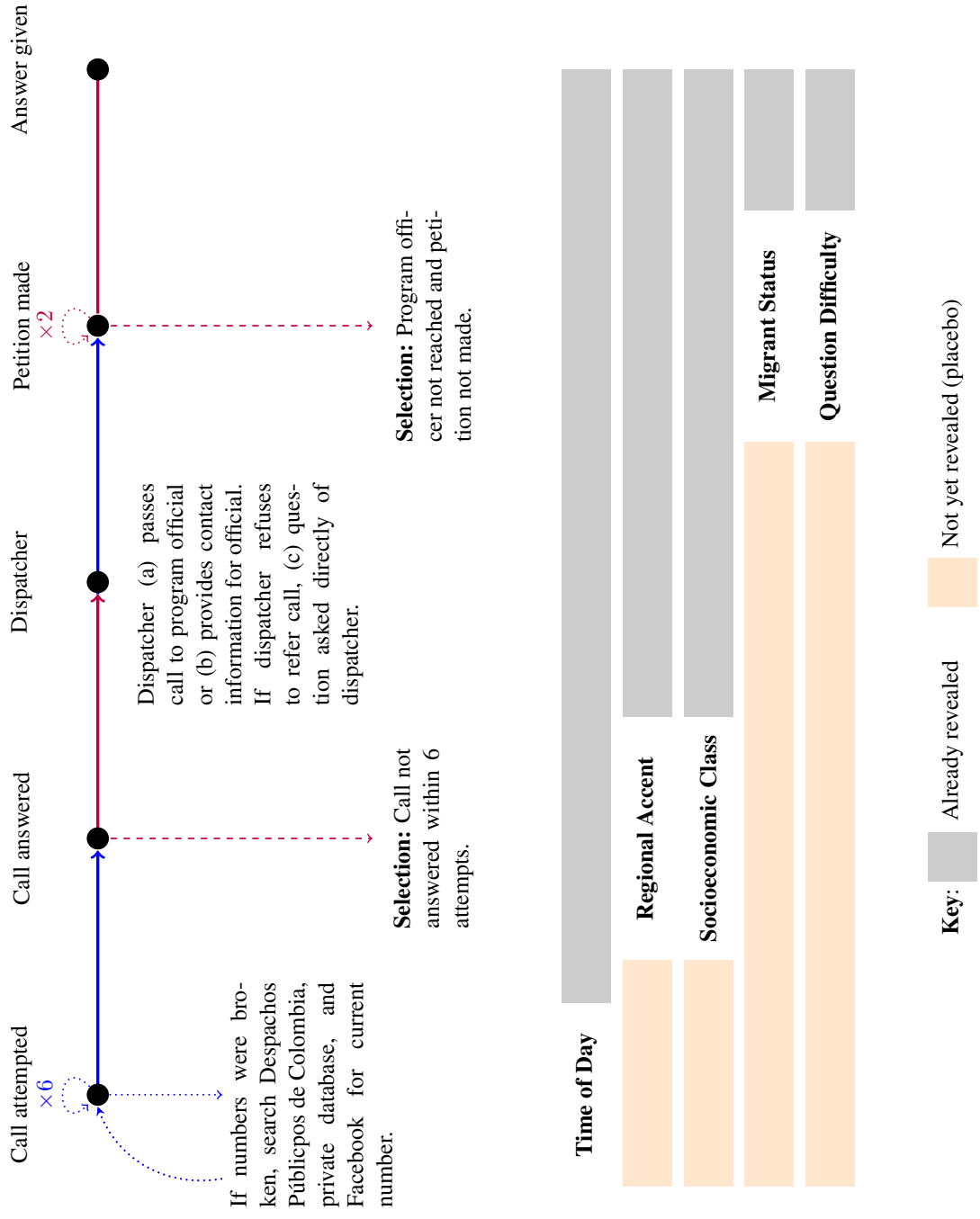
In general, the bounds suggest that at least 37% of the Colombian population (2005) has migrated at some point in their lives. This remains relatively stable across time. Furthermore, in past censuses, 10-17% of Colombians report migrating within the past five years. Note that migration to other departments is slightly more common than migration within the same department, though both forms of migration are common. These totals include both ordinary and conflict-induced migration.

	Census of				
	1964	1973	1985	1993	2005
A: MICRODATA SAMPLE SIZE (FROM IPUMS)					
Sample Size	349,652	1,988,831	2,643,125	3,213,657	4,006,168
B: LIFETIME MIGRATION (LOWER BOUND), RESIDES IN:					
Municipality of Birth	0.635	0.626	0.642	0.594	0.623
Same Department, Different Municipality	0.180	0.160	0.145	0.170	0.152
Different Department	0.181	0.209	0.210	0.225	0.204
(Born Abroad)	0.004	0.004	0.003	0.003	0.003
No Response	0.000	0.000	0.000	0.007	0.019
C: MIGRATION IN LAST FIVE YEARS (LOWER BOUND):					
Has not Migrated			0.869	0.834	0.905
Migrated within Department			0.056	0.062	0.031
Migrated from a Different Department			0.069	0.084	0.041
Immigrated			0.006	0.004	0.002
No response			0.000	0.016	0.021

Table A3: Lower bound on lifetime migration and migration within the last five years in Colombian censuses since 1964. Migration within the last five years was not included in the 1964 or 1973 censuses. Population-weighted means estimated from census microdata.

A6 Call Sequencing Flow Chart

Confederates were trained, in part, by memorizing a basic flowchart for each call which mirrored the instrument that they filled in. This graphic provides a translated and vastly simplified version of the sequence of calls, denoting the point in the call at which each factor was revealed and delineating the subsamples. As in the main manuscript three samples are defined and temporally delineated, as depicted below the flowchart.



A7 Petitions, Correct Answers

The following table lists the questions and correct answers used in the audit experiment translated to English. There are eight total questions. Because the difficulty of the question and migrant status are conveyed in the petition, there is effectively a 2×2 for each of the two programs, SISBÉN and MFA. The petitions appear in Tables A4 and A5, respectively.

All questions were presented in the third person to minimize detection (e.g. a request for the petitioner's identification number). Empirically, a substantial proportion of observed calls in government call centers were made on behalf of someone else. Further, in piloting, the responses to first person versus third person requests were qualitatively similar with the exception of petitions for an identification number.

	Program	Difficulty	Migrant	Question	Correct Answer
1	SISBÉN	Easy	Migrant	I am doing my neighbor a favor and asking about SISBÉN. She just moved to this municipality and wants to register for SISBÉN. What does she need to do to register?	1. Go to the SISBÉN office. 2. Ask for an application for the survey for the first time. 3. She must be a resident of the home, older than 18 years, and present her identification card. 4. (At this time, we are not doing new registrations for SISBÉN.)
2	SISBÉN	Easy	–	I am doing my neighbor a favor and asking about SISBÉN. She wants to register for SISBÉN. What does she need to do to register?	1. Go to the SISBÉN office. 2. Ask for an application for the survey for the first time. 3. She must be a resident of the home, older than 18 years, and present her identification card. 4. (At this time, we are not doing new registrations for SISBÉN.)
3	SISBÉN	Difficult	Migrant	I am doing my neighbor a favor and asking about SISBÉN. She just moved to the municpality and tried to enter *Colombia Mayor* and they did not let her. She has a score of 45. What can she do to lower her score? She is 65 years old.	1. Go to the SISBÉN office. 2. Check the form of the person to verify that the data in the system are correct. 3. If there are differences on the form, complete a petition of disagreement. 4. For the request to be approved, she will need to ask for a new survey. 5. She must be registered for SISBÉN in this municipality and her score must correspond to the guidelines in this [category of] municipality.
4	SISBÉN	Difficult	–	I am doing my neighbor a favor and asking about SISBÉN. She tried to enter *Colombia Mayor* and they did not let her. She has a score of 45. What can she do to lower her score? She is 65 years old.	1. Go to the SISBÉN office. 2. Check the form of the person to verify that the data in the system are correct. 3. If there are differences on the form, complete a petition of disagreement. 4. For the request to be approved, she will need to ask for a new survey.

Table A4: List of SISBÉN petitions, translated to English.

Program	Difficulty	Migrant	Question	Correct Answer
MFA	Easy	Migrant	I am doing my sister a favor and asking about MFA. She just moved to this municipality and wants to register for MFA. She has a 6 year old son and a 10 year old daughter. What does she need to do to register?	1. She must come to the office of the municipal *enlace* or point of service. 2. She must bring her document of identification. 3. She must bring the civil registration of all children under 7 and the identification card for all children between ages 7 and 18. 4. For children under 6, she should bring a certificate of children's attendence at (medical) exams of growth and development, certified with the name of the attending official. 5. For children in school, the mother should bring proof of enrollment in school. 6. (At this time, we are not doing new registrations for MFA). 7. The mother must register for SISBÉN in this municipality and have a qualifying score for MFA.
MFA	Easy	–	I am doing my sister a favor and asking about MFA. She wants to register for MFA. She has a 6 year old son and a 10 year old daughter. What does she need to do to register?	1. She must come to the office of the municipal *enlace* or point of service. 2. She must bring her document of identification. 3. She must bring the civil registration of all children under 7 and the identification card for all children between ages 7 and 18. 4. For children under 6, she should bring a certificate of children's attendence at (medical) exams of growth and development, certified with the name of the attending official. 5. For children in school, the mother should bring proof of enrollment in school. 6. (At this time, we are not doing new registrations for MFA). 7. The mother must register for SISBÉN in this municipality and have a qualifying score for MFA.

MFA	Difficult	Migrant	I am doing my neighbor a favor and asking about MFA. She just moved to this municipality. How does she change her registration with the program?	1. Go to the MFA office. 2. She must present a signed written request documenting that she has moved to the municipality. 3. The mother must register for SISBÉN in this municipality and have a qualifying score for MFA.
MFA	Difficult	–	I am doing my neighbor a favor and asking about MFA. Her sister, who was a MFA recipient died and left her in charge of her nephew. How does she become the MFA recipient for her nephew?	1. Go to the MFA office. 2. She must present her identification document. 3. She must turn in the document of custody and personal care of the child, issued by the competent authority: the defender or commisary of the family. 4. She must present the document from the civil registry documenting her sister's death. 5. The mother must be registered for SISBÉN in this municipality and have a qualifying score for MFA.

Table A5: List of MFA petitions, translated to English.

A8 Sampling, Construction of Weights

A8.1 Sampling

I use stratified random sampling in an effort to maximize within variation while limiting the probability of detection. Table 2 defines the characteristics of the strata. Here, I provide additional details on sampling of municipalities. The sampling probability for medium and large municipalities is clearly 1. However, the sampling probability for small municipalities is heterogeneous since sampling was proportional to estimated population (2018). While Eframidis and Spirakis (2006) provide a numerical solution for sampling probabilities for weighted sampling without replacement, the closed form probabilities require construction of all combinations, this solution is computationally infeasible in the present application with $\binom{898}{400}$ possible samples.

I approximate the sampling probabilities for the small stratum as follows:

1. Sample 400 municipalities with weights proportional to estimated population 100,000 times.
2. Calculate the probability that each municipality was sampled across the 100,000 draws.
3. Fit the a function by regressing these probabilities on 2018 population and polynomial terms (up to a tertic term) in OLS.
4. Predict the sampling probabilities from the fitted model.

Figure A8 depicts the sampling probabilities for municipalities as a function of population size. It demarcates the three strata by color. Note that Bogotá, the largest municipality, is divided into municipalities for the purposes of the experiment.

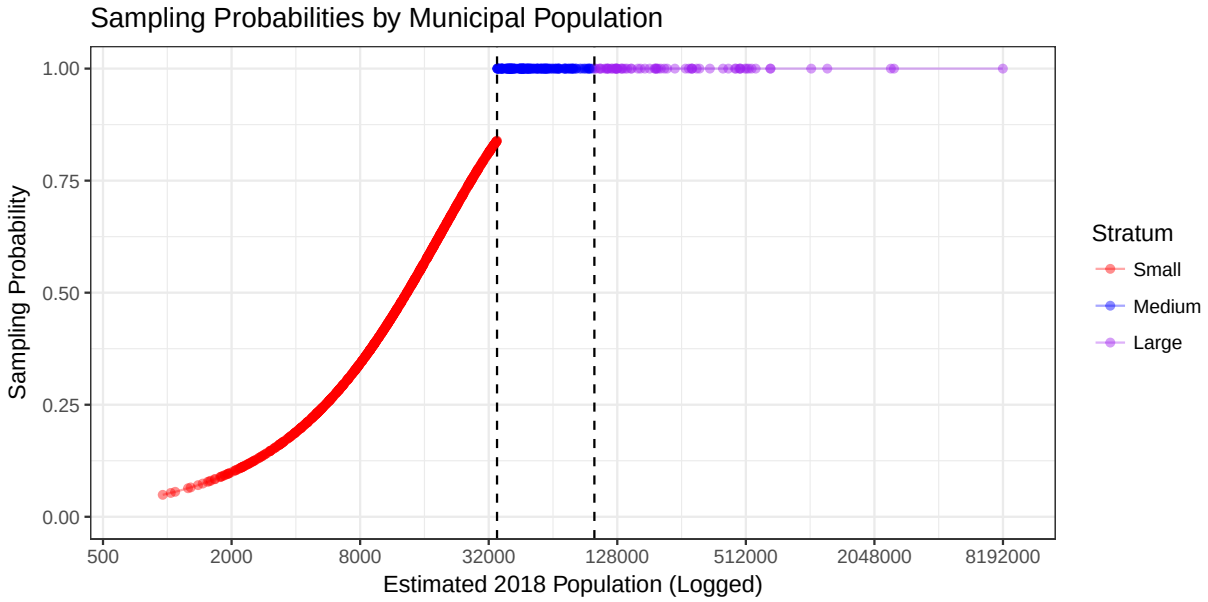


Figure A8: The sampling probabilities for municipalities as a function of population. The points represent individual municipalities in the universe. Note that these points are municipalities; localities in Bogotá are not represented in the present graph.

Predictably, this sampling procedure gives rise to a sample that is larger, on average, than the pool of municipalities as a whole, but one that provides support across the distribution of municipal populations, as depicted in the density plots in Figure A9. Note that in the experiment, 16 of Bogotá's localities are sampled according to the same rules for a total of 618 entities. Bogotá is represented as a whole in Figure A9.

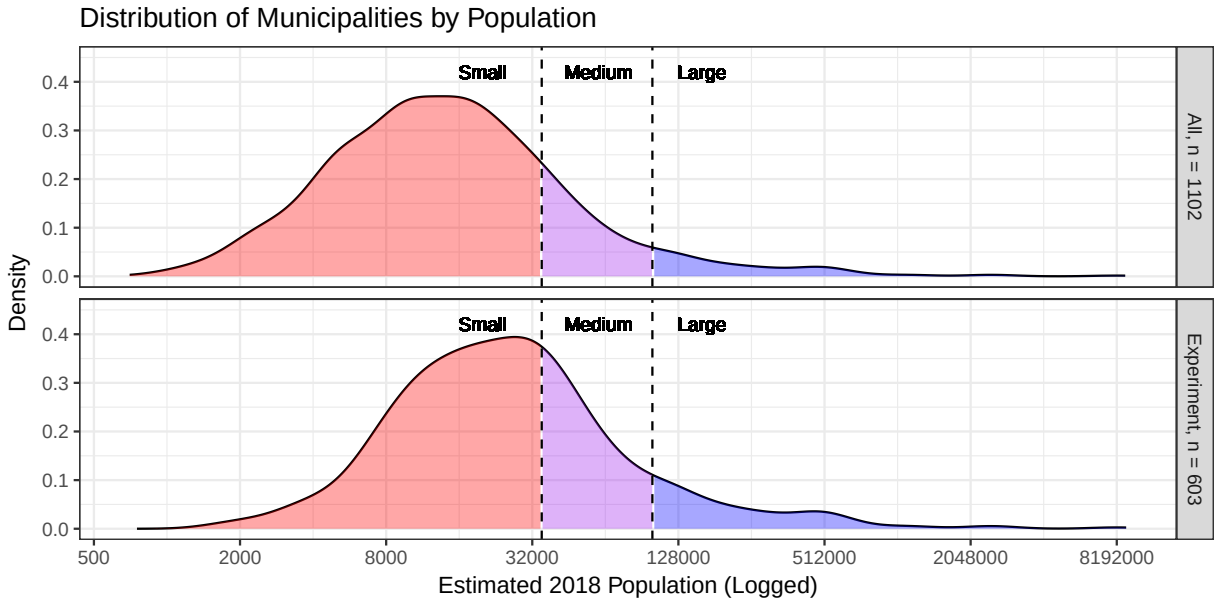


Figure A9: The municipalities in the sample compared by population compared to the full distribution of Colombian municipalities.

A8.2 Construction of Weights

Sample estimates can easily be estimated. Additionally, because the probabilities of selection into the sample are known and the (approximate) population of the municipalities is known, They can also be reweighted as follows:

- Municipal-level measures: To estimate municipal-level quantities, calculate the probability, p_j, that a municipality enters the sample. This probability is 1 for medium and large municipalities. The probability can be estimated numerically for the small municipalities using the method outlined above.This is then multiplied by $1/n_j$ where n_j is the number of petitions directed to the municipality. The weight for each municipality is thus $\frac{1}{n_j p_j}$. This weights small municipalities much more heavily than large municipalities.
- Population-weighted measures: To estimate population-weighted estimates, the number of calls per population is $n_j/\text{Population}_j$. The inverse probability weight is thus $\frac{Population_j}{n_j}$. This will weight large cities much more heavily than small municipalities.

A9 Random Assignment

A9.1 Factorial Design

The random assignment is generally blocked at the entity level (including both petitions) in order to maximize the within variation within each entity. The specifications by municipality size stratum are indicated in Table A6.

A9.2 Random Assignment of Enumerators

Enumerators come from each of the three regions corresponding to the regional accents. All enumerators are actresses and voice both the low- and neutral-class scripts. Within an accent, calls were block random-assigned to enumerators. As such, all enumerators cover all of the other factors of randomization.

A9.3 Random Assignment of Order of Calls

Calls were randomly assigned to an order for the original call. The assignment of the order proceeds as follows:

- Assignment to portions of the order distribution:
  - Within the large stratum (denoted $\mathcal{L}$), block randomly assign calls within each entity to six blocks. These correspond to sextiles of the rollout. Denote these blocks as $b_i^L \in \{1, 2, 3, 4, 5, 6\}$
  - Within the medium stratum (denoted $\mathcal{M}$), block randomly assign calls within each entity to four blocks. These correspond to quartiles of the rollout. Denote these blocks as $b_i^M \in \{1, 2, 3, 4\}$.
  - Within the small stratum (denoted $\mathcal{S}$), block randomly assign calls within entity to two blocks. These correspond to halves of the rollout. Denote these blocks as $b_i^S \in \{1, 2\}$
- Within each block in the rollout (defined above), randomly assign an integer ordering to the calls, denoted $O_i \in \{1, ..., \frac{|\mathcal{L}|}{6}\}$ for the large stratum, where $|\mathcal{L}|$ is the cardinality of the set of calls in the large stratum.
- Use the following formula to convert the rollout to a continuous measure between 0 and 1, shuffling the calls from the strata:

$$R_i = \frac{\mathbb{I}_{i\in\mathcal{L}}[(\frac{|\mathcal{L}|}{6}(b_i^L - 1) + O_i]}{|\mathcal{L}|} + \frac{\mathbb{I}_{i\in\mathcal{M}}[(\frac{|\mathcal{M}|}{4})(b_i^M - 1) + O_i]}{|\mathcal{M}|} + \frac{\mathbb{I}_{i\in\mathcal{S}}[\frac{|\mathcal{S}|}{2}(b_i^S - 1) + O_i]}{|\mathcal{S}|} \tag{10}$$

- Assign calls to enumerators (assigned as above) based on their order in the distribution.

Stratum	Attribute	Details	Restrictions
Large	Class	• 3 neutral • 3 low	Blocked by municpality/locality with each of 6 accent+ class combinations represented
	Accent	• 2 Bogotá • 2 Paisa • 2 Costeño	Blocked by municpality/locality with each of 6 accent+ class combinations represented
	Difficulty	• 4 easy • 2 hard	Blocked by program (SISBEN and MFA): easy questions (all) with one SISBEN and one MFA hard question.
	Migrant Status	• 3 migrant • 3 non-migrant	Both easy migrant questions; one hard question includes migrant line.
	Time of day	• 2-4 morning • 2-4 afternoon	Blocked by program (SISBEN, MFA) and municipality/locality
Medium	Class	• 2 neutral • 2 low	Blocked by program (SISBEN, MFA) and municipality/locality
	Accent	• 1 Bogotá • 1 Paisa • 1 Costeño • (1 repeated)	Blocked by municipality/locality. Repeated accent appears in both programs
	Difficulty	• 2 easy • 2 hard	Blocked by program (SISBEN/MFA) and municipality/locality
	Migrant Status	• 2 migrant • 2 non-migrant	Blocked by program (SISBEN/MFA) and municpality/locality
	Time	• 2 morning • 2 afternoon	Blocked by program (SISBEN/MFA) and municipality/locality
Small	Class	• 1 neutral • 1 low	Blocked by municpality/locality.
	Accent	• 2 distinct accents	Two distinct accents assigned to each municipality/locality
	Difficulty	• 1 easy • 1 hard	Blocked municipality/locality.
	Migrant Status	• 1 migrant • 1 non-migrant	Blocked by municipality/locality
	Time of day	• 1 morning • 1 afternoon	Blocked by municipality/locality

Table A6: Factorial design by stratum with restrictions on the randomization intended to maximize within-municipality variation. Note that the total number of petitions reflects those sent to both SISBÉN and MFA (combined).

A10 Distribution of Treatments

The frequency with which treatment cells were utilized in the experiment is reported in Table A7. Note that There was a higher probability of assignment to easy than hard questions in the large stratum, and thus in the experiment as a whole. Otherwise, all cells (within easy and hard) had equal probabilities of assignment.

Easy Petition								
			Migrant Accent			Non-Migrant Accent		
			Bogotano	Costeño	Paisa	Bogotano	Costeño	Paisa
Morning	Class	Lower	44	37	47	46	44	33
		Middle	57	37	50	45	48	47
Afternoon	Class	Lower	28	39	40	42	26	37
		Middle	50	52	49	32	36	32
Difficult Petition								
			Migrant Accent			Non-Migrant Accent		
			Bogotano	Costeño	Paisa	Bogotano	Costeño	Paisa
Morning	Class	Lower	25	34	28	44	32	43
		Middle	36	36	26	27	37	24
Afternoon	Class	Lower	43	38	35	41	42	40
		Middle	29	33	35	34	39	37

Table A7: Distribution of the frequency of treatment cells in the $2 \times 3 \times 2 \times 2 \times 2$ factorial design.

A11 Experimental Design Validation

A11.1 Compliance

With a relatively complex audit experiment and a large team of confederates, compliance with treatment assignment is a concern for the analysis and interpretation of findings. To address such concerns, all calls were recorded. Subsequent to the experiment, two trained research assistants listened to all the calls (a full time job for over a month) and marked what they heard in the calls. The research assistants were not apprised of the schedule of treatment assignment.[5] I examine compliance factor by factor in Table A8.

Factor	Validation Data	Assignment	Coding in Validation data				% Compliers
Time	Phone log times		Morning	Afternoon	Mix with $\geq$ 1 intent		
		Morning	98.7%	0	1.3%		98.7%
		Afternoon	0.3%	98.0%	1.65%		98.0%
Accent	Double entry		Bogotá	Costeño	Paisa	Indeterminate	
		Bogotá	99.7%	0%	0%	0.3%	99.7%
		Costeño	0.3%	98.7%	0%	1.0%	98.7%
		Paisa	0.8%	0.5%	98.4%	0.3%	98.4%
Class	Double entry		Lower	Neutral	Indeterminate		
		Lower	76.7%	11.2%	13.1%		76.7%
		Neutral	6.7%	79.3%	14.0%		79.3%
Difficulty	Double entry		Technical	Easy			
		Technical	99.3%	0.7%			99.3%
		Easy	0.8%	99.2%			99.2%
Migrant	Double entry		Migrant	Resident			
		Migrant	97.3%	2.7%			97.3%
		Resident	5.0%	95.0%			95%

Table A8: Rates of compliance by treatment. Double entry refers to the hand coded data by outsiders listening to recordings after the fact. The phone log times were automatically recorded and outside the purview of confederates.

[5] Calls for 10 petitions were lost by the software doing the recordings. These calls represent less than 1% of the total calls and I have no reason to believe that the missingness is systematic.

A11.2 Detection

Of the 1194 answered calls, bureaucrats appeared to detect six of the calls as audits, per the classification of double coders. These calls are detailed below. Note that all calls that were detected were detected prior to the statement of the petition. There are no systematic differences in detection by municipal population stratum (as defined in the sampling of municipalities) or destination of the calls (department). Further, there are no systematic differences in the characteristics of the petition or petitioner except that these calls occurred later in the sequence within a given *alcaldía*.

	Stratum	Department	Call Order	Program	Time	Accent	Class	Difficulty	Migrant
1	Large	Bogotá	4/6	MFA	PM	Costeño	Lower	Easy	Resident
2	Large	Bogotá	5/6	MFA	AM	Paisa	Lower	Easy	Migrant
3	Large	Cundinamarca	5/6	MFA	PM	Paisa	Lower-Middle	Technical	Migrant
4	Medium	Bolívar	4/4	SISBÉN	PM	Paisa	Lower-Middle	Technical	Resident
5	Small	Bolívar	2/2	MFA	PM	Costeño	Lower	Technical	Resident
6	Small	Cauca	2/2	SISBÉN	PM	Paisa	Lower-Middle	Technical	Resident

Table A9: This table documents the calls that were detected, as denoted by the double coders. Both calls detected in Bogotá were detected in the same locality (*alcaldía local*).

A11.3 Joint Test of Interactions between Identity Treatments

The empirical strategy employed in this paper analyzes "along the margins" of the factorial experimental design. In this section, I allay concerns of substitutability or complementarities between the identity-based attributes used to measure bias. In this analysis, I utilize F-tests of the joint significance of the relevant interactions. To conduct this analysis, I use the IPW model from the main analysis as the restricted model:

$$Y_{ipm} = \sum_{j \in \mathbf{Z}} \beta_j Z_i^j + \kappa_p + \epsilon_{ipm} \tag{11}$$

I then specify an unrestricted model, as in Equation. Note that M_i indicates lower-middle class, R_i indicates a resident, B_i indicates a Bogotá accent, and C_i indicates a Costeño accent.

$$\begin{aligned} Y_{ipm} = \sum_{j \in \mathbf{Z}} \beta_j Z_i^j + \kappa_p + \chi_1 M_i R_i + \chi_2 M_i B_i + \chi_3 M_i C_i + \chi_4 R_i B_i + \chi_5 R_i C_i + \\ \chi_6 M_i R_i B_i + \chi_7 M_i R_i C_i + \epsilon_{ipm} \end{aligned} \tag{12}$$

I test the null hypothesis that $\chi_1 = \chi_2 = \chi_3 = \chi_4 = \chi_5 = \chi_6 = \chi_7 = 0$ using an F-test. The results of this test are reported in Table A10. I fail to reject the null hypothesis for 9/10 outcomes at the $\alpha = 0.1$ level, providing no evidence that of interactions between the identity-based treatments.

	N			
Outcome	Unrestricted	Restricted	F-statistic	p-value
Dispatcher Gave Name	1186	1179	0.17	0.99
Petition Made	1186	1179	0.43	0.88
Second Official	1186	1179	1.24	0.28
Program Officer	1186	1179	0.23	0.98
Complete	1186	1179	1.44	0.19
Incomplete	1186	1179	0.90	0.50
Any Info.	1186	1179	1.04	0.40
No Info.	1186	1179	1.24	0.28
Alcaldía Only	1186	1179	2.71	0.01
Red Tape	1186	1179	0.80	0.58

Table A10: Results of F-tests of the significance of interactions between identity-based characteristics for each of the main outcomes reported in Table 4.

A12 Personnel Data

A12.1 Source

The administrative data comes from the compilation of five administrative datasets, obtained by various means. Table A11 documents the datasets and how they were obtained.

Dataset	Population	Coverage (Experimental)	Method Obtained
SIGEP	Civil Servants	>800 municipalities	Obtained from Departamento Administrativo de la Función Pública by *derecho de petición* (≈ FOIA request). Request granted in April 2018.
SIGEP	Civil Servants	798 employees in experimental *alcaldías*	Identified by hand search and scraped from online database, May 2018.
SIDEAP	Civil Servants	20 Bogotá *alcaldías locales*	Downloaded in April 2018.
SECOP-I	Contractors	1100 *alcaldías*, Metro SALUD Medellín	Downloaded in April 2018.
SECOP-II	Contractors	20 Bogotá *alcaldías*	Downloaded in April 2018.
SIDEAP	Contractors	Bogotá *alcaldías* (for cross-referencing)	Downloaded in April 2018.

Table A11: Administrative datasets, coverage, and method used to obtain the data.

At the aggregate level, this data overlaps with the experimental sample of 600 *alcaldías* and 18 *alcaldías locales* as follows. "No contractor data coverage" indicates that there are no current contracts (as of the experiment) in the system.

	Contractor Data Coverage (SECOP-I, SECOP-II)	
	Coverage	No Coverage
Civil Servant Coverage	550	22
No Civil Servant Coverage	40	6

A12.2 Identification of Relevant Personnel

The individual data comes from two sources. First, I have a list of program officers furnished by PS (MFA) and downloaded from DNP (SISBÉN). Second, the list of names comes from the double entered calls. Enumerators wrote the names shared by officials. Most names in the administrative data contain four names, two first names and two last names, as is standard in administrative documents.

The names from the calls are often much shorter (often one first name and one last name) and spelling is approximated by the research assistants. While most Spanish names are quite straightforward to spell, some

Colombians have adopted English names. Spelling of these names varies substantially. For example, in my data the common name pronounced "Jason" (to an English speaker) is spelled "Jeison," "Jeisson," "Yeison," and "Yeisson." This poses a challenge for string matching. To maximize information before matching to the administrative datasets, I matched the calls to the administrative data, which would match the hypothetical the name "Yohana Díaz" to "Johanna Luz Díaz Guerrero."

I then took the list of unique names (some full and some short) to the dataset of civil servants and contractors. To identify relevant personnel, I implemented a q-gram matching algorithm, selecting matches that maximize the number of common 3-grams within entity. I then revised the matches with the most q-grams in common by hand, applying a 90% (of the shorter name) threshold or a phonetic match (i.e. the multiple spellings of "Jason") to identify employees.

Where no match was located, I searched the online dataset of SIGEP by hand, since the coverage of the original civil servant dataset was much more limited. If the employee was located on SIGEP, I wrote a Python program to scrape their information. I applied the same rule for identification to the SIGEP searches as to the q-gram matches.

In total, I was able to match the following:

Classification	Observations	Unique Employees
Contractor	1328	1028
Civil Servant	1658	1218
Individual not found	463	375
Neither data source	10	9
No contractor data	40	34
No civil servant data	121	97
Proportion Identified	0.825	0.813
Proportion Identified | $\geq$ 1 Call answered in municipality	0.831	0.821
Proportion Identified | Administrative Data	0.843	0.843
Proportion Identified | Call data	0.808	0.789

Table A12: Matched sample of bureaucrats from the calls and program officer data.

Government interviews suggested that the data quality is worse in small and poor municipalities. Based on their comments, I analyze the correlates of missingness/unidentifiability in the data briefly showing that the highest rates of unidentifiability occur in very poor municipalities. In Figure A11 I examine an index of municipal poverty (2005 census) and estimated 2018 population as predictors of missingness. Unidentifiability is increasing (monotonically) in the poverty rate in each municipality.

A12.3 Individual Measures

The main individual measure is an indicator of civil service versus patronage. Within each subset, I have some covariates (with varying degrees of missingness). For the purposes of contracts, I measure the tenure of the individual's employment as the date since the first contract was signed. I also note whether this occurred before or after the last transition of municipal government (January 1, 2016).

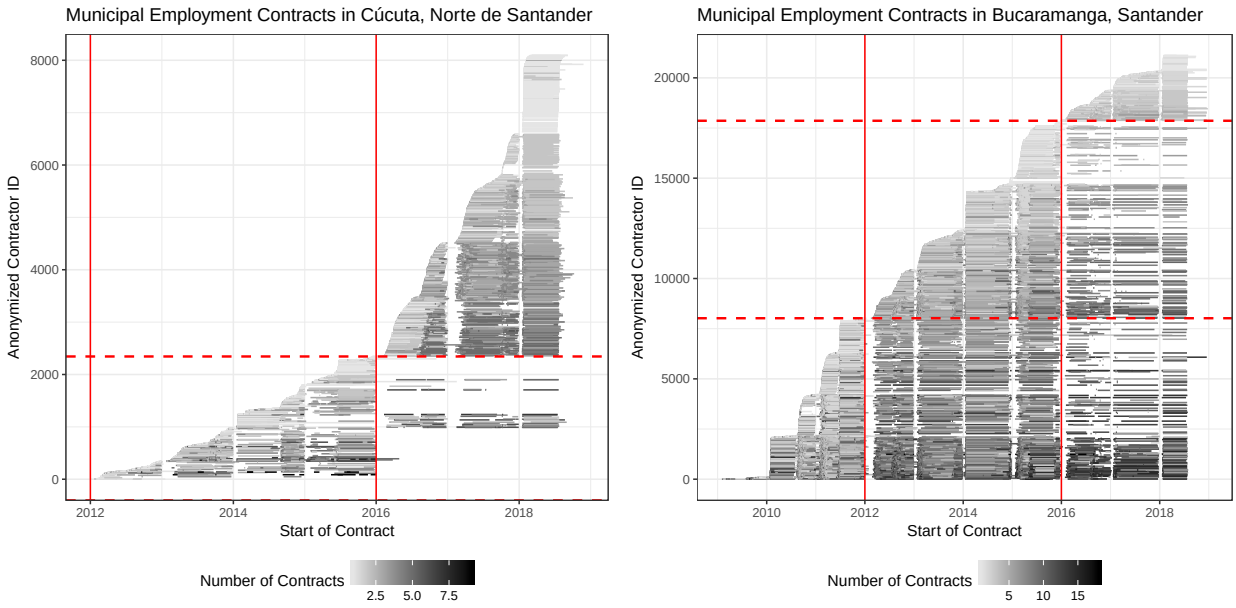


Figure A10: All contracts in SECOP-I for Cúcuta and Bucaramanga, Colombia.

A12.4 Municipal-Level Measures

At the municipal level, I measure two variables from the contract data:

- Median length of contracts, in days. Principals can theoretically exercise more control by holding bureaucrats to shorter contracts.
- Degree of turnover following the last municipal transition of power. I examine the degree of turnover after the last change of municipal government in 2016. This quantity is the share of contractors hired during 2016 and 2017 that had previously served as a contractor in public administration in the *alcaldía*. Stylized examples from Bucaramanga and Cúcuta two large cities in the same region illustrate the difference in proportion graphically in Figure A10. A much larger share of Bucaramanga's employees carried on after the election.

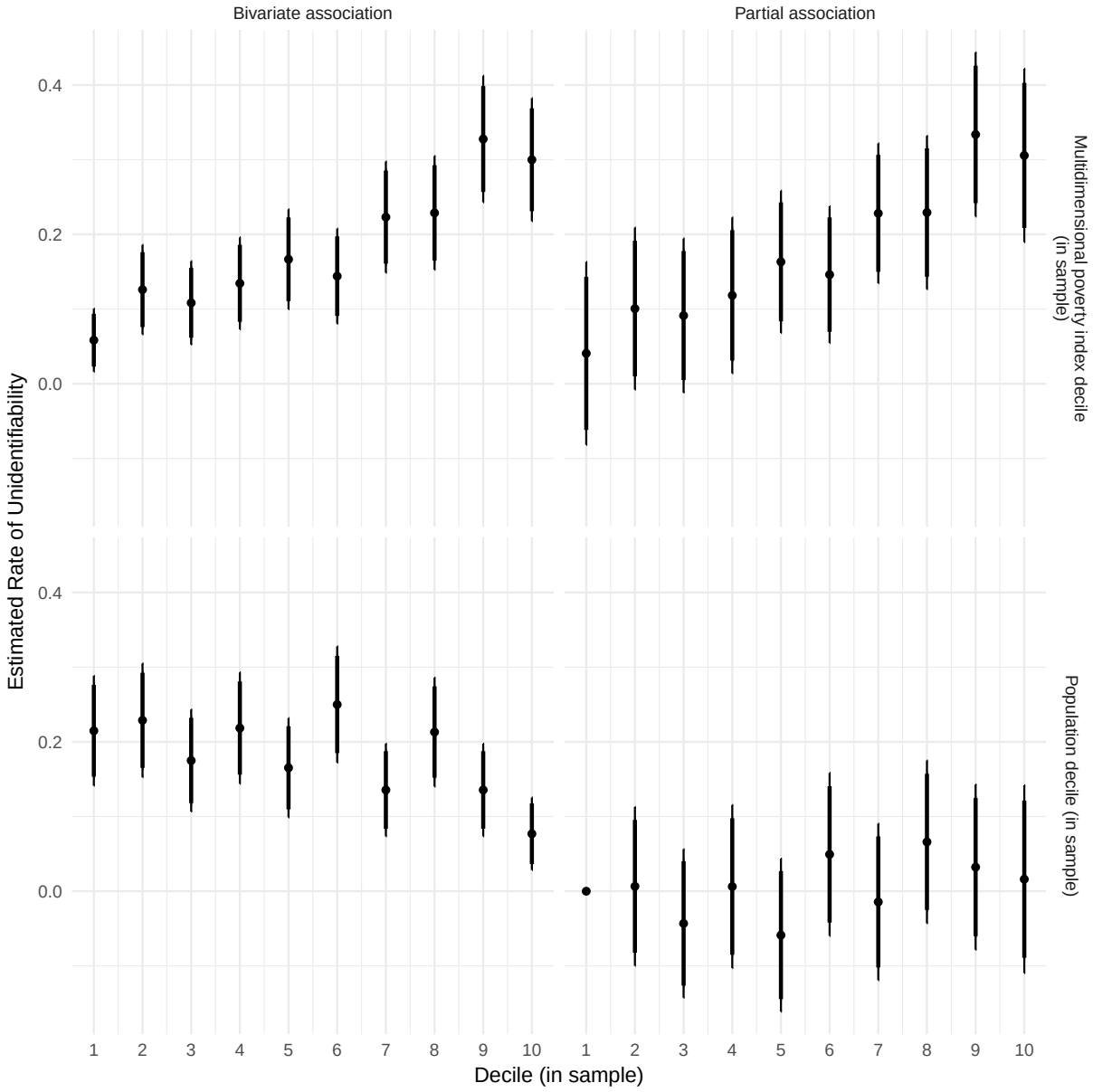


Figure A11: The association between the DANE multidimensional municipal poverty index (2005), estimated population (2018), and failure to identify employees in the database. The left column is estimated from two separate bivariate regressions. The right column depicts partial associations from a model with both sets of indicators. Confidence intervals are estimated on the basis of heteroskedasticity-robust standard errors.

A13 Demographic Data

The demographic data, measured at the municipal level, that is used in analyses come from:

- Population projections as of 2018. I use DANE's 2018 population projections for municipalities and the Alcaldía Mayor's projections of Bogotá's population, by locality, in 2018.
- Census of 2005: Municipal education levels and municipal multidimensional poverty index.
- Census microdata: Rates of migration are obtained from all census microdata available through IPUMS.
- SISBÉN registrations: Data downloaded from SISBÉN's open data as of November, 2017. Available at `https://www.sisben.gov.co/Territorios/Paginas/Reportes%20Base%20Certificada/ano2018.aspx`.
- Population of internally displaced persons (IDPs), by municipality. I obtain data on the number of IDPs per municipality from Unidad de Victímas as of July 1, 2018. Available at `http://cifras.unidadvictimas.gov.co/Home/Desplazamiento`.

A14 Election Data

Election data came from the dataset at Universidad de los Andes. This data is compiled from the Registraduría Nacional del Estado Civil by the Centro de Estudios sobre Desarollo Económico. All candidates in municipal elections are available since 1997. I examine on the municipal elections of 1997, 2000, 2003, 2007, 2011, and 2015.

Extreme weakness of parties limited serial correlation render standard measures of competitiveness poor measures. The correlation between mayoral margins of victory at time $t-1$ and t are depicted in Figure A12.

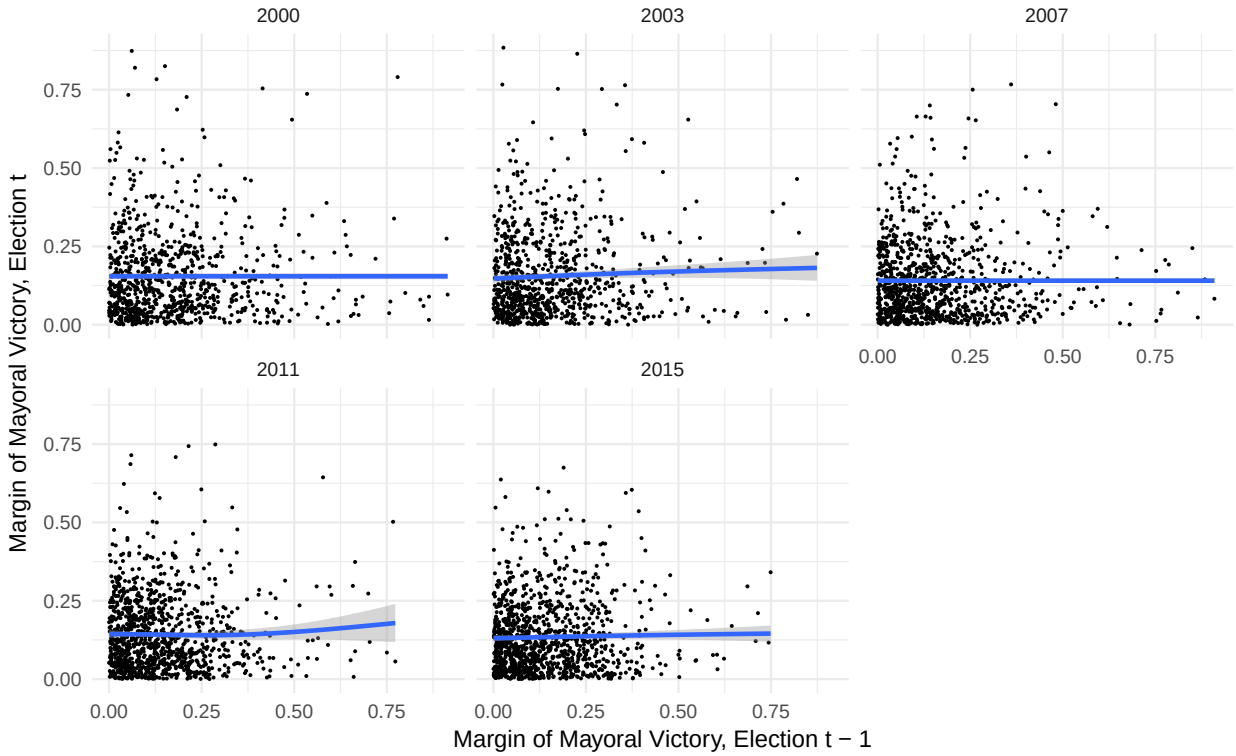


Figure A12: There is effectively no correlation between mayoral margin of victory at time $t-1$ and time t. Each point represents a municipality. The lines are fit with GAM.

Instead I use three measures of municipal political competition from both local mayoral and council elections.

- Mayoral election effective number of candidates (ENC). This measure is increasing in competitiveness. It is calculated via the following formula, where i index candidates and p_i is each candidate's vote share.

$$ENC = \frac{1}{\sum_{i=1}^{n} p_i^2}$$

- Number of unique council members since 1997. Council members are not term limited. This is calculated by the following formula.

$$\text{Ratio of Unique Concejales} = \frac{\#\text{ Unique winners, 1997-2015}}{\#\text{ Council seats, 1997-2015}}$$

- The Colombian naming tradition is to give a child two last names. The first is the father's first last name; the second is the mother's first last name. Most candidates use both last names. To measure concentration of power by family, I examine the ratio of unique last names to total last names ($\approx 2\times$ number of candidates).

$$\text{Ratio of Unique Last Names} = \frac{\#\text{ Unique last names of concejales, 1997-2015}}{\#\text{ Total last names of consejales, 1997-2015}}$$

A15 Sample Selection

As indicated in Table 3, I estimate the bias outcomes on the sample of answered calls. While the blocking scheme ensures balance across municipalities and within the sample of calls, it is worth assessing whether or not the resultant sample is imbalanced across the experimental factors. This is assessed through an F-test of the joint significance of the four unrevealed factors: class, accent, migrant status, and question difficulty.

I show in Table A13 that there is no evidence that entering this sample (Columns 3-4) is endogenous to the unrevealed factors. The test of this claim is the F-test on the placebos, for each estimator. Further, Columns 1-2 and 5 show that the unrevealed factors do not predict how soon or late a call was answered. In sum, there is no evidence that answering is endogenous to any of the analyzed treatments.

	Did not Answer First Call		Did not Answer Any Calls		Number of Calls
	(1)	(2)	(3)	(4)	(5)
Afternoon Call	0.031*	0.022	0.039**	0.030**	0.291*
	(0.023)	(0.018)	(0.023)	(0.015)	(0.188)
DV mean, Morning	0.639	0.639	0.330	0.330	4.48
Estimator	OLS	OLS	OLS	OLS	Tobit
IPW	✓		✓		✓
Entity, Enumerator FE		✓		✓	
Program Indicator	✓	✓	✓	✓	✓
Placebos	✓	✓	✓	✓	✓
Placebo F-test*, p-value	0.898	0.381	0.705	0.163	0.619
Hypothesis test	Upper	Upper	Upper	Upper	Upper
DV range	{0, 1}	{0, 1}	{0, 1}	{0, 1}	[1,7]
Censoring					Right
Observations	1,836	1,836	1,836	1,836	1,836

Note: *p<0.1; **p<0.05; ***p<0.01

Table A13: The AMCEs of an afternoon call on measures of absenteeism of the dispatcher. Six calls were attempted over the course of at least three days; Columns 3-4 indicate that none of the six calls were answered. Heteroskedasity-robust standard errors for the OLS models in parentheses. *For the tobit model, the p-value on the test of joint significance of the placebos is calculated from a Pearson χ^2 likelihood ratio test. The mean outcome in the morning in the fifth column refers to the latent outcome and is estimated by tobit.

A16 Robustness of Bias Estimates

A16.1 Complier AMCE Estimates for the Class Treatment

The results on the class treatment are central to findings of bias. There are two main concerns about this treatment, with responses below.

1. Can people judge a caller's class by voice alone?

    - The results from the double coding of compliance in Table A8 suggest as much: 77.5% of calls were correctly identified; 13.5% were indeterminate; and only 9% were opposite of the assignment indicator.

2. Class is a compound treatment in any case in Colombia. The scripts that distinguish the classes include different salutations, different vocabularies, and different presentation of the question.

    - The pervasiveness of social class in the Colombian context is important; isolating class from its correlates or constituent parts is not particularly feasible, nor is it particularly useful in this context.
    - The results show that class-based bias happens in certain processes and not others. There is no evidence that bureaucrats did not understand questions from poor petitioners, as rates of correct responses are not substantially different between lower- and lower-middle class Colombians. These facts suggest that there was no systematic response to the characteristics used to connote class.

To test further concerns about the excludability of treatment assignment in Table A14, I estimate complier AMCEs on the class treatment, instrumenting an observed lower-middle class exchange with assignment to the lower-middle class treatment condition. If results are driven by perceptions of class and coder ratings of class are correlated with how bureaucrats perceived class, then (non-zero) point estimates of complier AMCEs should be higher than intent-to-treat AMCEs. I report the class estimates from Table 4 (Panels A and B) in the main text along their complier analogues.

Note that this test serves as an informal test of excludability. It is also possible that if some characteristic driving the results that is highly positively correlated with observer judgments of class is driving bias in behavior, the complier estimates would be higher than the intent-to-treat estimates. All of the point estimates on the outcomes where there is evidence for bias in the ITT AMCE estimates in Table 4 in the main text increase in magnitude.

A16.2 Regional Accent and Red Tape Robustness Test

Table 4 documents the disproportionate rate at which red tape was demanded from petitioners with a Paisa regional accent. Because the regional accent did not vary within an enumerators' calls, I conduct tests that drop enumerators one by one as well as in trios (one enumerator per accent) to examine the robustness of this finding. Figures A13 and A14 show the point estimates when dropping one enumerator and one trio of enumerators, respectively. They suggest that the inference is generally robust and that the effect does not seem to be driven by any single enumerator or pair/trio of enumerators.

	Access/Process					Response to Petition				
	Dispatcher Gave Name	Petition Made	Second Official	Program Officer	Complete	Incomplete	Any Info.	Red Tape	*Alcaldía* Only	extra
PANEL A: INTENT-TO-TREAT AMCE ESTIMATES, IPW ADJUSTMENT										
Lower-Middle Class	0.011	0.014	0.017	0.044	0.021	0.048*	0.069**	−0.019	−0.036**	0.003
	(0.020)	(0.025)	(0.029)	(0.028)	(0.018)	(0.029)	(0.029)	(0.021)	(0.017)	(0.024)
PANEL B: COMPLIER AMCE ESTIMATES, IPW ADJUSTMENT										
Lower-Middle Class	0.028	0.019	0.033	0.069	0.035	0.065	0.100**	−0.028	−0.053**	0.003
	(0.028)	(0.037)	(0.044)	(0.042)	(0.027)	(0.043)	(0.043)	(0.031)	(0.025)	(0.036)
PANEL C: INTENT-TO-TREAT AMCE ESTIMATES, ENTITY FIXED EFFECTS										
Lower-Middle Class	−0.015	0.016	−0.004	0.049*	0.024	0.057**	0.081***	−0.027	−0.039**	0.002
	(0.019)	(0.025)	(0.028)	(0.025)	(0.019)	(0.028)	(0.028)	(0.021)	(0.017)	(0.026)
PANEL D: COMPLIER AMCE ESTIMATES, ENTITY FIXED EFFECTS										
Lower-Middle Class	0.002	0.028	0.009	0.083**	0.038	0.087**	0.125***	0.006	−0.055**	0.006
	(0.025)	(0.036)	(0.041)	(0.037)	(0.028)	(0.040)	(0.040)	(0.037)	(0.025)	(0.037)
Mean DV, Paisa Accent	0.842	0.723	0.53	0.296	0.119	0.414	0.533	0.124	0.066	0.266
All Factors	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DV range	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
Observations	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194

Note: *p<0.1; **p<0.05; ***p<0.01

Table A14: Estimates of the intent-to-treat and complier AMCEs of the class treatment. Heteroskedasticity robust standard errors in parentheses. "Observed" indicates that the accent was observed by double coders that listened to all call recordings.

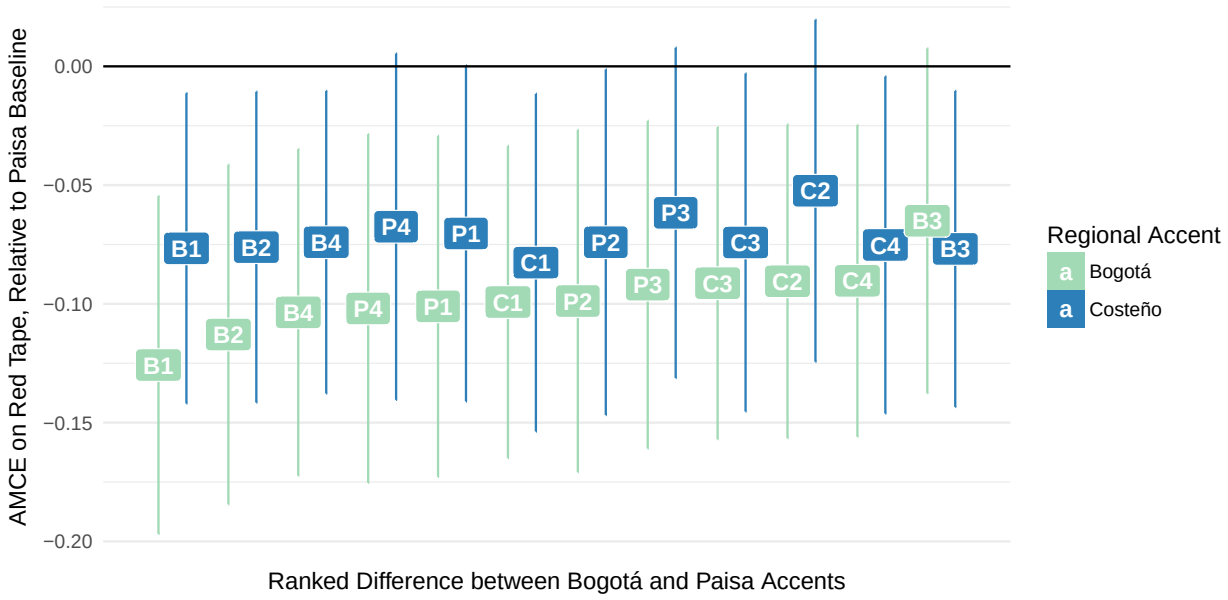


Figure A13: AMCE of Bogotá and Costeño accents relative to a Paisa accent when dropping one enumerator. The enumerator codes are labeled on the points and indexed by accent ("B", "C", or "P") and number. The bars indicate 95% confidence intervals calculated on heteroskedasticity-robust standard errors.

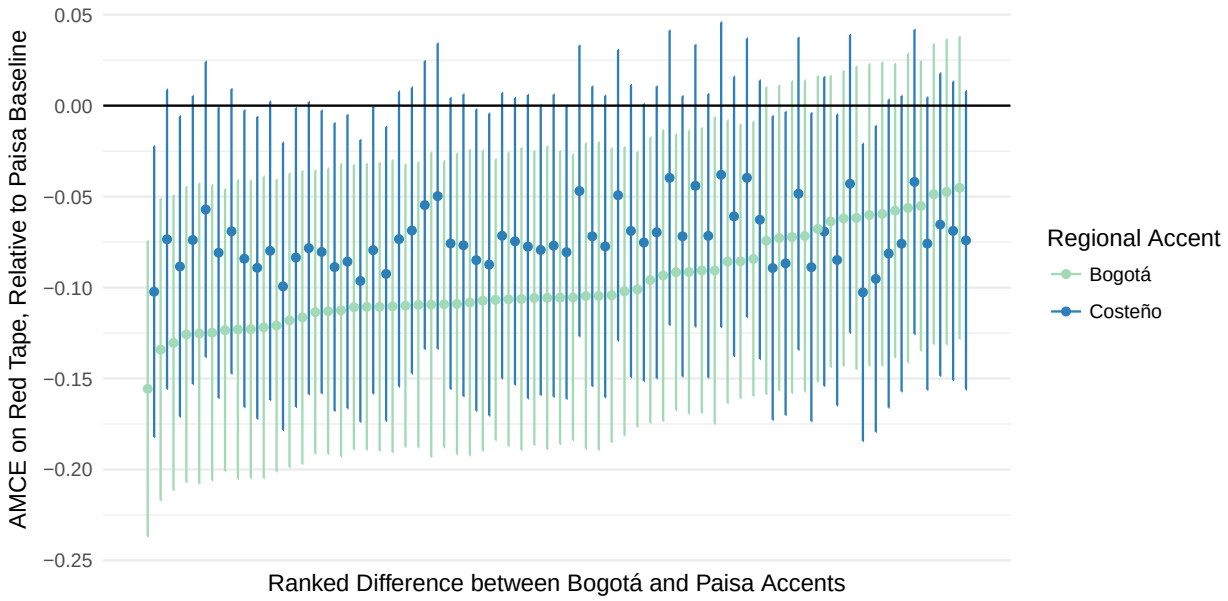


Figure A14: AMCE of Bogotá and Costeño accents relative to a Paisa accent when dropping each permutation of three enumerators (one per accent). Note that by dropping three enumerators simultaneously, the effective sample is approximately 75% of the main sample reported in Table 4, inflating standard error estimates accordingly. The bars indicate 95% confidence intervals calculated on heteroskedasticity-robust standard errors.

A16.3 Regional Accents in Home Region

In the main results, I analyze regional accents without regard to the match between the accent of the petitioner and the accent native to the municipality where the audit was conducted. This analysis redefines the treatment indicator for accent as being an "in-region" (home) accent. Because the experimental accent treatment only includes 3 of $\approx$ 12 regional accents, I subset to the regions of Colombia in which there is common support for the treatment. Table A15 clarifies the definition of the region for purposes of analysis as well as the total number of petitions in the subsample. In sum, this sample represents about 55% of all answered calls.

				n of Petitions	
Accent	Region	Departments	Total	"In-region"	"Out-region"
Bogotá	Centro Oriente (subset)	(Bogotá) Cundinamarca	198	65	133
Costeño	Caribe	Atlántico Bolívar Cesar Córdoba La Guajira Magdalena San Andrés Sucre	189	58	131
Paisa	Eje Cafetero - Antioquia	Antioquia Caldas Quindío Risaralda	270	85	185

Table A15: Definition of regions for the analysis of "in-region" accents. The sample from which these municipalities are drawn is the 1194 answered calls. Note that by construction, $\frac{1}{3}$ of calls should be "in-region"; this proportion is maintained in this subsample.

With both adjustment strategies (IPW and entity fixed effects), the region indicator is interacted with all factors and an indicator for the program. Results are reported in Table A16. The main takeaways are as follows:

- We cannot reject the null hypothesis that the conditional AMCE = 0 for the pooled subsample or any subgroup therein.
- For some outcomes, there are statistically significant differences between regions. If anything, these differences seem to be driven by the fact that, Costeño confederates seem to receive slightly "worse" service "in region" in the Atlantic Coast (Caribe). This was not anticipated.

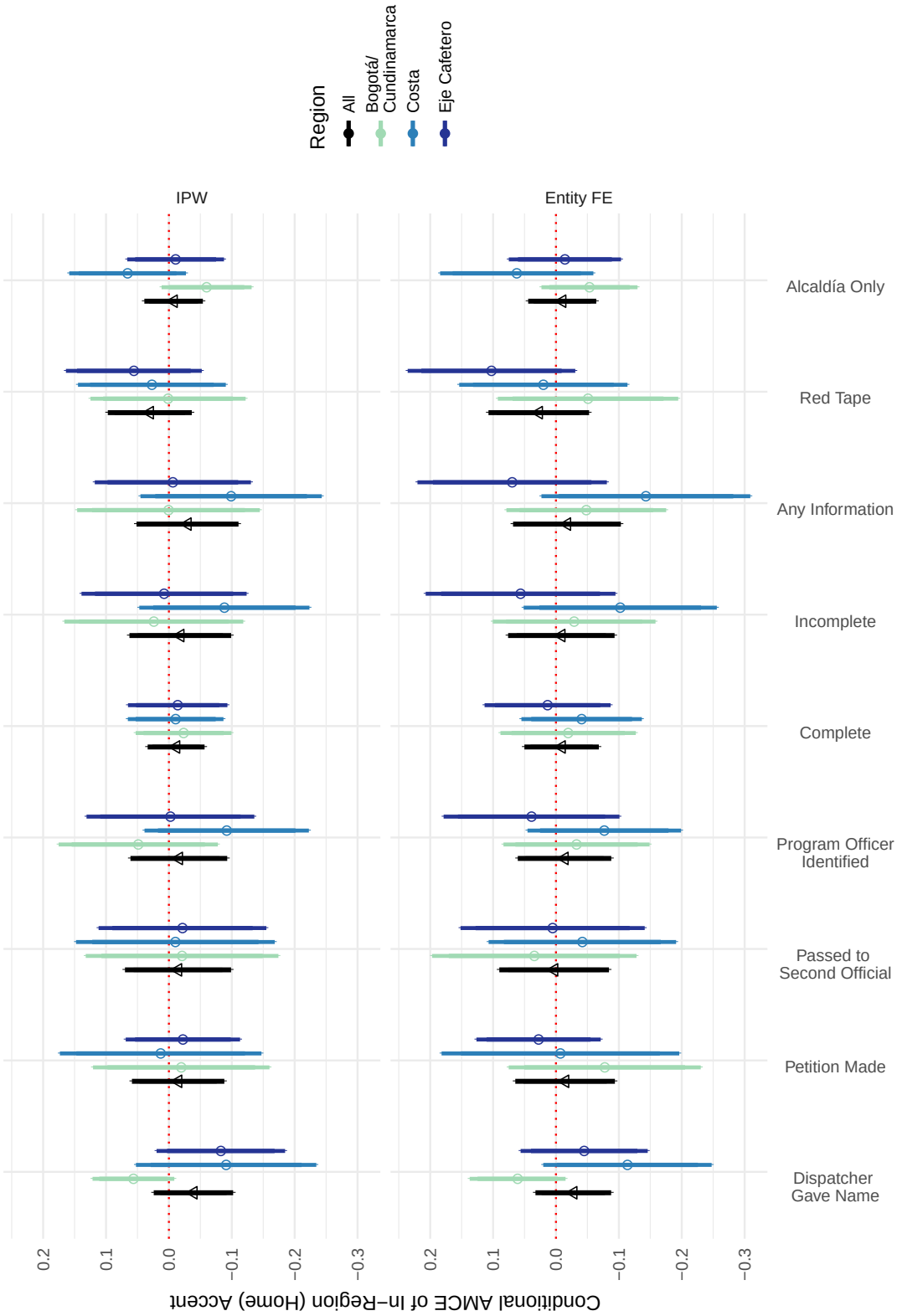


Table A16: Estimates of the conditional AMCE of an "in region" accent petition. Bars represent 95% confidence intervals constructed on heteroskedasticity robust standard errors. The panels correspond to the adjustment strategy used in estimation.

A16.4 Migrant Status in the "Call Answered" and "Petitioned" Samples

Migrant status was not revealed until petitions were made (see Table 3), however the main analysis analyzes outcomes based on migrant status within the sample of all answered calls (Table 6). This serves to inflate the effective rate of noncompliance for the migrant factor. This attenuates the resultant intent-to-treat estimates. The F-tests in Tables 4 and 6 (columns 1-4) suggest that there is no reason to believe that migrant status was revealed prior to the petition. As such, the estimates in the smaller sample should be larger in magnitude than those in the main text. Table A17 supports this interpretation.

	Response to Petition									
	Complete		Incomplete		Any Info.		Red Tape		*Alcaldía* Only	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: IPW ESTIMATES										
Migrant	−0.002	−0.003	−0.027	−0.054	−0.029	−0.057*	−0.036	−0.049	0.054***	0.070***
	(0.017)	(0.021)	(0.029)	(0.033)	(0.029)	(0.030)	(0.024)	(0.030)	(0.017)	(0.021)
PANEL B: ESTIMATES WITH ENTITY FIXED EFFECTS										
Migrant	0.003	−0.003	−0.044	−0.065*	−0.041	−0.067**	−0.041	−0.033	0.061***	0.081***
	(0.018)	(0.025)	(0.028)	(0.035)	(0.028)	(0.032)	(0.026)	(0.034)	(0.017)	(0.023)
PANEL C: ESTIMATES WITH ENTITY + ENUMERATOR FIXED EFFECTS										
Migrant	0.002	0.0003	−0.043	−0.063*	−0.041	−0.063**	−0.040	−0.029	0.061***	0.077***
	(0.018)	(0.026)	(0.028)	(0.035)	(0.028)	(0.032)	(0.026)	(0.035)	(0.017)	(0.023)
Mean DV, Resident	0.102	0.135	0.447	0.592	0.731	0.713	0.252	0.334	0.061	0.081
Program	✓	✓	✓	✓	✓	✓	✓	✓	✓	
All Factors	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DV range	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}	{0, 1}
Sample	Answered	Petitioned	Answered	Petitioned	Answered	Petitioned	Answered	Petitioned	Answered	Petitioned
Observations	1,194	911	1,194	911	1,194	911	1,194	911	1,194	911

Note: *p<0.1; **p<0.05; ***p<0.01

Table A17: Estimates of the AMCEs of the migrant treatment on both the call answered subsample from the main text as well as the petitioned sample. Heteroskedasticity robust standard errors in parentheses.

A16.5 Survey Outcomes

In this section, I examine the responses of enumerators to survey questions about their experience of each call. The survey consisted of five questions, translated to English as follows:

1. *Satisfaction:* On a scale from 1 to 5, where 1 is "not at all satisfied" and 5 is "very satisfied," how satisfied are you with the service given by the public servant during the call?

2. *Confidence (in answers):* On a scale from 1-4, where 1 is "not confident" and 4 is "very confident," how much confidence did you have in the responses given by the public servant?

3. *Actionable information:* On a scale from 1 to 5, where 1 is "very easy" and 5 is "very difficult," how hard would it be to carry out the process (service) you asked about on the basis of the information you received?

4. *Knowledge:* On a scale from 1 to 5, where 1 is "very low" and 5 is "very high," what level of knowledge did the public servant have when responding to the request?

5. *Respect:* On a scale from 1 to 5, where 1 is "very little respect" and 5 is "lots of respect," how respectful was the public servant while responding to the request?

Due to an issue in the programming of the survey, responses are missing for 59/1194 answered calls and 30/911 of the calls in which a petition was made. In these cases, the survey did not appear at the end of data input. The differential proportions of missing survey responses indicate that this error occurred as part of the sequence in the survey. This may be endogenous to some of the experimental manipulations, if through the trajectory of the call. However, missingness is balanced across the factors.

My main measure of service provision is a z-score index comprising the five measures enumerated above. Component # 3 is reversed such that higher scores on the scale map onto higher values of the index (better service). I also report the standardized measure of respect given arguments made in the paper. Estimates of AMCES of the experimental manipulations on these outcomes are reported in Table A18.

Note that the enumerator effects reported seem to correspond to idiosyncracies in how individual enumerators assess service. The estimates are *not* robust to dropping one enumerator at a time. I focus on the within-enumerator estimates in columns (3) and (6) of both panels as the main measure of the relationship between petition and petitioner attributes and service provision.

A more direct way to observe the underlying components of survey evaluations is examination of the ECDFs of the class and migrant conditions on individual outcomes. This analysis is reported in Figure A15.

	Dependent variable:					
	Service Index (Standardized)			Respect (Standardized)		
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: CLASS AND ACCENT ON ANSWERED CALLS SUBSAMPLE						
Lower Middle Class	0.109**	0.110**	0.096**	0.020	0.010	0.002
	(0.047)	(0.047)	(0.045)	(0.059)	(0.061)	(0.058)
Bogotá Accent	0.287***	0.226***		0.238***	0.191**	
	(0.058)	(0.063)		(0.075)	(0.081)	
Costeño Accent	0.062	0.064		−0.297***	−0.305***	
	(0.059)	(0.063)		(0.079)	(0.084)	
DV mean, Lower Class	-0.052	-0.052	-0.052	-0.006	-0.006	-0.006
DV mean, Paisa	-0.106	-0.106	-0.106	0.033	0.033	0.033
DV Range	[-2.957, 1.904]	[-2.957, 1.904]	[-2.957, 1.904]	[-3.767, 1.263]	[-3.767, 1.263]	[-3.767, 1.263]
Observations	1,135	1,135	1,135	1,135	1,135	1,135
PANEL B: MIGRANT STATUS AND QUESTION DIFFICULTY ON PETITION SUBSAMPLE						
Migrant	0.060	0.085	0.085	0.109*	0.143**	0.125*
	(0.049)	(0.057)	(0.055)	(0.061)	(0.072)	(0.071)
Technical Petition	−0.128**	−0.110*	−0.117**	−0.124**	−0.107	−0.114
	(0.050)	(0.058)	(0.056)	(0.063)	(0.075)	(0.072)
DV mean, Easy	-0.052	-0.052	-0.052	-0.006	-0.006	-0.006
DV mean, Resident	0.062	0.062	0.062	-0.003	-0.003	-0.003
DV Range	[-2.957, 1.904]	[-2.957, 1.904]	[-2.957, 1.904]	[-3.767, 1.263]	[-3.767, 1.263]	[-3.767, 1.263]
Observations	881	881	881	881	881	881
Estimator	IPW	Entity FE	Entity, Enum. FE	IPW	Entity FE	Entity, Enum. FE
Program	✓	✓	✓	✓	✓	✓
All Factors	✓	✓	✓	✓	✓	✓
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A18: The AMCEs of treatment conditions on enumerators' perceived treatment and service by bureaucrats. The outcome for columns 1-3 (both panels) is a Z-score index of five attributes of service. The outcome for columns 4-6 is the standardized "respect" component of that index. Heteroskedasticity-robust standard errors in parentheses.

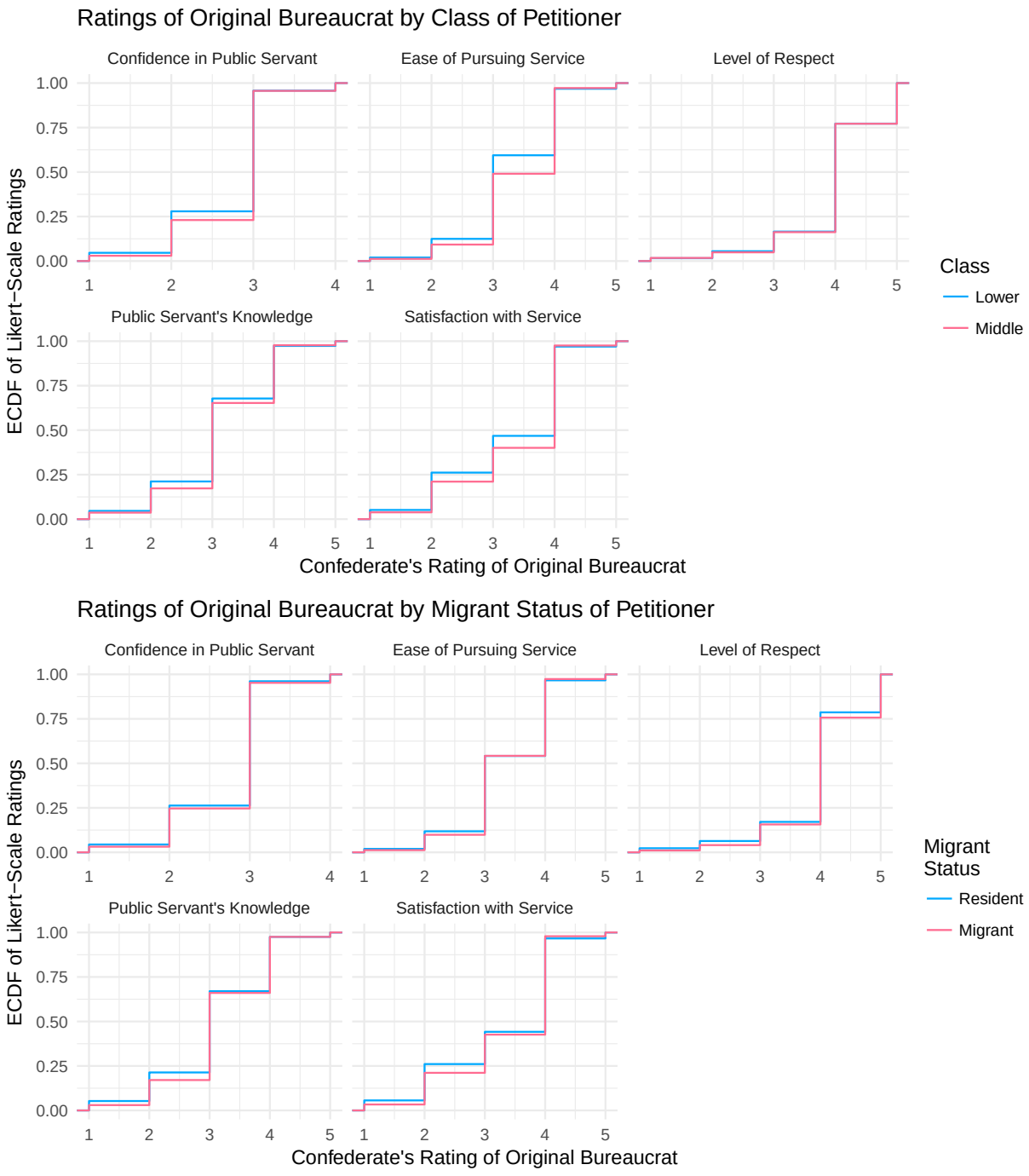


Figure A15: The ECDFs of ratings of the service provided, disaggregated by class (top panel) and migrant status (bottom panel). Importantly, the ECDFs do not cross in either panel.

A17 Supporting Tables for Tests of the Mechanism

This section provides tables to support the graphical analysis in Section 6 of the main text. The first tests examine the cost of effort and employee type (contractor/civil servant). For covariates with within-*alcaldía* variation, I do not run the specification with *alcaldía* fixed effects, since this reduces the effective sample dramatically. Instead, I run a second specification with indicators for stratum (to account for differential probabilities of assignment) and enumerator.

A17.1 Distinguishing Bureaucrat Taste-Driven from Oversight-Driven Bias

Table A19 shows that bias in information provision by class was attenuated to zero on technical questions. While the interaction is not generally significant, differences are quite stark. Note, however, that the theory implies a one-tailed test while the table reflects (conservative) two-tailed p-values.

	Complete	Incomplete	Any Info.	Alcaldía Only	Red Tape
	(1)	(2)	(3)	(4)	(5)
PANEL A: CONDITIONAL AMCE BY DIFFICULTY OF PETITION; IPW ESTIMATES					
Lower-Middle Class	0.039	0.080**	0.118***	−0.040***	0.036
	(0.026)	(0.038)	(0.038)	(0.015)	(0.034)
Hard: Lower-Middle Class	−0.034	−0.067	−0.101*	0.009	−0.064
	(0.035)	(0.057)	(0.058)	(0.034)	(0.048)
Conditional Effect, Technical Petition	0.004	0.013	0.017	-0.031	-0.029
	(0.023)	(0.042)	(0.044)	(0.030)	(0.034)
PANEL B: CONDITIONAL AMCE BY DIFFICULTY OF PETITION; STRATUM + ENUMERATOR FE					
Lower-Middle Class	0.030	0.068*	0.098***	−0.037**	0.030
	(0.026)	(0.037)	(0.038)	(0.015)	(0.034)
Hard: Lower-Middle Class	−0.021	−0.049	−0.070	0.004	−0.054
	(0.034)	(0.056)	(0.057)	(0.034)	(0.048)
Conditional Effect, Technical Petition	0.010	0.019	0.029	-0.033	-0.024
	(0.023)	(0.042)	(0.043)	(0.030)	(0.033)
Observations	1,194	1,194	1,194	1,194	1,194

Note: *p<0.1; **p<0.05; ***p<0.01

Table A19: Estimates of the conditional AMCE of a lower-middle class petitioner by the difficulty of the question. The base category is the "easy"/inscriptions questions. The experimental treatment technical petition is interacted with all factors and a program indicator. "Conditional effect" refers to the conditional effect of lower-middle class. Heteroskedasticity- robust standard errors in parentheses.

Table A20 shows that bias in information provision by class was attenuated to zero for MFA petitions. While the interaction is not generally significant, differences are large and robust to different estimators. Note that the theory implies a one-tailed test while the table reflects (conservative) two-tailed p-values.

	Complete	Incomplete	Any Info.	Alcaldía Only	Red Tape
	(1)	(2)	(3)	(4)	(5)
PANEL A: CONDITIONAL AMCE BY DIFFICULTY OF PETITION; IPW ESTIMATES					
Lower-Middle Class	0.003	0.030	0.033	−0.021	0.019
	(0.021)	(0.040)	(0.041)	(0.020)	(0.031)
SISBÉN: Lower-Middle Class	0.036	0.031	0.067	−0.029	−0.034
	(0.035)	(0.057)	(0.058)	(0.033)	(0.049)
Conditional Effect, SISBÉN	0.039	0.061	0.100**	-0.050*	-0.015
	(0.028)	(0.041)	(0.041)	(0.026)	(0.037)
PANEL B: CONDITIONAL AMCE BY DIFFICULTY OF PETITION; STRATUM + ENUMERATOR FE					
Lower-Middle Class	0.007	0.022	0.029	−0.019	0.019
	(0.020)	(0.039)	(0.040)	(0.019)	(0.031)
SISBÉN: Lower-Middle Class	0.028	0.047	0.075	−0.032	−0.029
	(0.035)	(0.056)	(0.057)	(0.031)	(0.048)
Conditional Effect, SISBÉN	0.035	0.069*	0.104***	-0.051**	-0.010
	(0.029)	(0.040)	(0.040)	(0.025)	(0.037)
Observations	1,194	1,194	1,194	1,194	1,194
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A20: Estimates of the conditional AMCE of a lower-middle class petitioner by the audited program (MFA or SISBEÉN). The base category is MFA. The program indicator is interacted with all factors in the experimental design. "Conditional effect" refers to the conditional effect of lower-middle class. Heteroskedasticity-robust standard errors in parentheses.

A17.2 Isolating Complaint-Driven Bias

Table A21 examines the conditional effect of class by level of municipal poverty. Panels A and B show the estimates using both of the main estimators in the manuscript. Panels C and D demonstrate that these findings are robust to the use of flexible, interactive controls for municipal population.

One alternative explanation for the positive association between class-bias and poverty rate is clientelism. I conduct three tests to show that the evidence is inconsistent with this explanation. First, I show that levels of service given to the lower middle class do not vary in municipal poverty rates. If a politician were simply co-opting a social program to devote services clentelistically as in Weitz-Shapiro (2012), we would expect lower levels of service by bureaucrats across the board (e.g. even for the lower middle class). If this happens disproportionately in poor places, then there should be a negative association between municipal poverty rates and service outcomes for the lower middle class. Table A22 indicates that this is not the case. There is little evidence of a correlation, and if anything, there is weak evidence of a a *positive* correlation between municipal poverty and information provision. This correlation is not robust to alternate functional forms (Panel B) or to dropping Bogotá (Panels C-D), which accounts for a disproportionate share of observations. This provides no evidence in favor of a clientelism explanation for observed findings.

	Complete	Incomplete	Any Info.	Alcaldía Only	Red Tape
	(1)	(2)	(3)	(4)	(5)
PANEL A: CONDITIONAL AMCE BY TERCILE OF POVERTY RATE; IPW ESTIMATES					
Lower-Middle Class	0.013	−0.024	−0.011	−0.027	−0.007
	(0.029)	(0.046)	(0.051)	(0.028)	(0.045)
Medium Poverty: Lower-Middle Class	0.019	0.085	0.105	−0.025	0.062
	(0.042)	(0.067)	(0.070)	(0.043)	(0.062)
High Poverty: Lower-Middle Class	0.001	0.139**	0.139**	−0.012	−0.032
	(0.045)	(0.066)	(0.070)	(0.040)	(0.060)
Conditional Effect in M	0.032	0.061*	0.094**	-0.051*	0.055*
	(0.028)	(0.047)	(0.046)	(0.034)	(0.039)
Conditional Effect in H	0.014	0.115***	0.128***	-0.038*	-0.039
	(0.035)	(0.045)	(0.044)	(0.026)	(0.040)
PANEL B: CONDITIONAL AMCE BY TERCILE OF POVERTY RATE; ENTITY + ENUMERATOR FE					
Lower-Middle Class	0.019	−0.024	−0.005	−0.020	−0.013
	(0.036)	(0.050)	(0.053)	(0.030)	(0.054)
Medium Poverty: Lower-Middle Class	0.025	0.091	0.117	−0.039	0.080
	(0.051)	(0.077)	(0.077)	(0.050)	(0.074)
High Poverty: Lower-Middle Class	−0.001	0.151**	0.150**	−0.030	−0.033
	(0.057)	(0.077)	(0.076)	(0.048)	(0.075)
Conditional Effect in M	0.044	0.067	0.111**	-0.058*	0.067*
	(0.037)	(0.061)	(0.060)	(0.042)	(0.054)
Conditional Effect in H	0.018	0.127**	0.145***	-0.050*	-0.047
	(0.048)	(0.060)	(0.059)	(0.039)	(0.056)
PANEL C: CAMCE BY TERCILE OF POVERTY RATE WITH INTERACTIVE POPULATION CONTROL; IPW ESTIMATES					
Lower-Middle Class	0.010	0.019	0.029	−0.033	0.004
	(0.036)	(0.056)	(0.056)	(0.036)	(0.058)
Medium Poverty: Lower-Middle Class	0.028	0.037	0.065	−0.012	0.054
	(0.052)	(0.080)	(0.081)	(0.056)	(0.077)
High Poverty: Lower-Middle Class	−0.009	0.111	0.102	−0.005	−0.034
	(0.053)	(0.082)	(0.082)	(0.057)	(0.083)
Conditional Effect in M	0.039*	0.056	0.095**	-0.045*	0.057*
	(0.031)	(0.050)	(0.047)	(0.036)	(0.040)
Conditional Effect in H	0.001	0.130***	0.131***	-0.038	-0.030
	(0.036)	(0.048)	(0.049)	(0.032)	(0.048)
Interactive Population Decile Bins	✓	✓	✓	✓	✓
PANEL D: CAMCE BY TERCILE OF POVERTY RATE WITH INTERACTIVE POPULATION CONTROL; ENTITY + ENUMERATOR FE					
Lower-Middle Class	0.023	−0.003	−0.011	−0.030	−0.020
	(0.045)	(0.069)	(0.062)	(0.041)	(0.072)
Medium Poverty: Lower-Middle Class	0.028	0.054	0.113	−0.021	0.087
	(0.065)	(0.096)	(0.083)	(0.062)	(0.094)
High Poverty: Lower-Middle Class	−0.028	0.146	0.159*	−0.017	−0.021
	(0.071)	(0.105)	(0.093)	(0.067)	(0.107)
Conditional Effect in M	0.051*	0.051	0.101**	-0.051	0.067*
	(0.040)	(0.065)	(0.060)	(0.045)	(0.053)
Conditional Effect in H	-0.005	0.144**	0.147**	-0.048	-0.041
	(0.049)	(0.065)	(0.063)	(0.044)	(0.065)
Interactive Population Decile Bins	✓	✓	✓	✓	✓
Observations	1,194	1,194	1,194	1,194	1,194
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A21: Estimates of the conditional AMCE by tercile of poverty. The base category is the first tercile (lowest rate of poverty). All covariates and moderators (poverty tercile indicator and demeaned poverty decile bins) are interacted across all factors in the design. Standard errors are clustered at the municipality level since this is the level of measurement of the poverty moderator.

Second, I draw on documentation of clientelism in Colombia to identify variation in the presumed intensity of clientelism across municipalities. I show that the positive association between class-bias and poverty rate is robust to controlling interactively for these measures. The tests are as follows. First, I show that there is no association between documented threats to electoral integrity. These threats include clientelism, fraud, intimidation, and electoral violence. To the extent that an emerging literature on clientelism suggests that clientelism includes both "carrots" and "sticks," this provides a measure of both instruments (Mares and Young, 2018). Using data from the Misión de Observacion Electoral in Colombia, I control for a binary indicator for a general predicted threat to the 2018 national elections (Misión de Observación Electoral, 2018). Where a threat is identified, I code this variable as a "1." This creates two categories, which I demean and interact across all factors and the program indicator. Panels A and B of Table A23 suggest that results are not sensitive to estimating effects within levels of electoral threat.

Writing on clientelism in Colombia suggest that clientelism is practiced in distinct patterns in different regions (Ocampo, 2014). To account for these patterns, I include interactive department ($n = 30$) fixed effects in Panels C and D of Table A23. Note that some departments have few municipalities and few calls, so Panel D, in particular represents a subset of the sample in departments where there is variation in both population category within department. Nevertheless, results are consistent with the broader patterns documented in the main text and in Table A21. These analyses provide no evidence that clientelism is driving the observed association between poverty and bias.

Figure A16 shows no evident variation in class-based bias as a function of local political competition. The three columns use three alternate measures of local political competition. The left column uses the ratio of unique councilors to total councilors (1997-2015); the middle column uses the effective number of mayoral candidates in the last 3 elections; and the last column uses an inverse covariance weighted index of the first two plus unique last names (*apellidos*) over council elections from 1997-2015.

	Complete	Incomplete	Any Info.	Alcaldía Only	Red Tape
	(1)	(2)	(3)	(4)	(5)
Panel A: Linear Association between Poverty and Outcomes, Lower-Middle Class Petitioners Only					
Poverty rate	−0.011	0.268***	0.255**	−0.017	−0.098
	(0.060)	(0.104)	(0.105)	(0.043)	(0.080)
Observations	600	600	600	600	600
Panel B: Quadratic Association between Poverty and Outcomes, Lower-Middle Class Petitioners Only					
Poverty rate	0.195	0.775	0.970	0.445	−0.091
	(0.360)	(0.616)	(0.593)	(0.283)	(0.499)
Poverty rate^2	−0.182	−0.449	−0.631	−0.409*	−0.006
	(0.311)	(0.535)	(0.508)	(0.238)	(0.428)
Observations	600	600	600	600	600
Panel C: Linear Association between Poverty and Outcomes, Lower-Middle Class Outside of Bogotá					
Poverty rate	−0.008	0.053	0.044	−0.028	−0.136
	(0.066)	(0.107)	(0.593)	(0.047)	(0.091)
Observations	559	559	559	559	559
Panel D: Quadratic Association between Poverty and Outcomes, Lower-Middle Class Outside of Bogotá					
Poverty rate	0.264	−0.747	−0.483	0.477	−0.379
	(0.419)	(0.645)	(0.591)	(0.311)	(0.586)
Poverty rate^2	−0.233	0.685	0.452	−0.432*	0.208
	(0.350)	(0.554)	(0.505)	(0.255)	(0.484)
Observations	559	559	559	559	559
Factors (not Class)	✓	✓	✓	✓	✓
Program Indicator	✓	✓	✓	✓	✓
Estimator	IPW	IPW	IPW	IPW	IPW
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A22: Estimates of the association between municipal rates and service devoted to the lower middle class. The sample includes only calls made by lower-middle class petitioners, and in Panels C and D, only calls made outside of Bogotá. All covariates and and a program indicator are interacted across all factors in the design. Standard errors are clustered at the municipality level since this is the level of measurement of poverty rates.

	Complete	Incomplete	Any Info.	Alcaldía Only	Red Tape
	(1)	(2)	(3)	(4)	(5)
PANEL A: CONDITIONAL AMCE BY TERCILE OF POVERTY WITH INTERACTIVE ELECTORAL THREAT CONTROLS					
Lower-Middle Class	0.010	−0.023	−0.013	−0.017	−0.017
	(0.029)	(0.050)	(0.046)	(0.030)	(0.047)
Medium Poverty: Lower-Middle Class	0.026	0.090	0.116*	−0.041	0.081
	(0.044)	(0.074)	(0.068)	(0.044)	(0.064)
High Poverty: Lower-Middle Class	−0.007	0.147**	0.140**	−0.014	−0.028
	(0.044)	(0.070)	(0.068)	(0.040)	(0.061)
Conditional Effect in M	0.036*	0.067*	0.103**	-0.059**	0.064*
	(0.029)	(0.048)	(0.047)	(0.033)	(0.041)
Conditional Effect in H	0.003	0.124***	0.128***	-0.032*	-0.045
	(0.035)	(0.046)	(0.046)	(0.026)	(0.041)
PANEL B: CONDITIONAL AMCE BY TERCILE OF POVERTYWITH INTERACTIVE ELECTORAL THREAT, POPULATION CONTROLS					
Lower-Middle Class	0.015	0.020	0.034	−0.041	0.012
	(0.035)	(0.059)	(0.054)	(0.037)	(0.059)
Medium Poverty: Lower-Middle Class	0.024	0.042	0.066	−0.005	0.048
	(0.050)	(0.083)	(0.080)	(0.054)	(0.075)
High Poverty: Lower-Middle Class	−0.033	0.123	0.089	0.018	−0.060
	(0.052)	(0.087)	(0.087)	(0.057)	(0.087)
Conditional Effect in M	0.039*	0.061	0.101**	-0.046*	0.060*
	(0.031)	(0.050)	(0.048)	(0.036)	(0.041)
Conditional Effect in H	-0.018	0.142***	0.124***	-0.023	-0.048
	(0.036)	(0.050)	(0.050)	(0.034)	(0.048)
Interactive Population Decile Bins	✓	✓	✓	✓	✓
PANEL C: CAMCE BY TERCILE OF POVERTY RATE WITH INTERACTIVE DEPARTMENT FIXED EFFECTS					
Lower-Middle Class	−0.034	0.032	−0.003	−0.035	0.010
	(0.041)	(0.061)	(0.056)	(0.036)	(0.058)
Medium Poverty: Lower-Middle Class	0.072	0.059	0.131	−0.028	0.077
	(0.054)	(0.080)	(0.080)	(0.054)	(0.077)
High Poverty: Lower-Middle Class	0.072	0.097	0.169*	−0.042	−0.006
	(0.066)	(0.098)	(0.091)	(0.062)	(0.084)
Conditional Effect in M	0.038	0.091**	0.129***	-0.063**	0.088**
	(0.033)	(0.051)	(0.052)	(0.038)	(0.045)
Conditional Effect in H	0.038	0.129**	0.166***	-0.077**	0.005
	(0.040)	(0.059)	(0.055)	(0.037)	(0.049)
PANEL D: CAMCE BY TERCILE OF POVERTY RATE WITH INTERACTIVE DEPARTMENT FIXED EFFECTS, POPULATION CONTROLS					
Lower-Middle Class	−0.068	0.160**	0.091	−0.041	0.027
	(0.055)	(0.077)	(0.073)	(0.037)	(0.072)
Medium Poverty: Lower-Middle Class	0.116	−0.077	0.039	−0.005	0.066
	(0.071)	(0.095)	(0.096)	(0.054)	(0.091)
High Poverty: Lower-Middle Class	0.114	−0.105	0.009	0.018	−0.036
	(0.092)	(0.129)	(0.126)	(0.057)	(0.115)
Conditional Effect in M	0.048*	0.083*	0.130***	-0.046*	0.094**
	(0.036)	(0.054)	(0.053)	(0.036)	(0.046)
Conditional Effect in H	0.046	0.055	0.101*	-0.023	-0.008
	(0.046)	(0.071)	(0.069)	(0.034)	(0.061)
Interactive Population Decile Bins	✓	✓	✓	✓	✓
Estimator	IPW	IPW	IPW	IPW	IPW
Observations	1,194	1,194	1,194	1,194	1,194
Note:					*p<0.1; **p<0.05; ***p<0.01

Table A23: Estimates of the conditional AMCE by tercile of poverty, controlling for electoral threats, department, and poverty. The base category is the first tercile (lowest rate of poverty). All covariates and moderators (poverty tercile indicator, demeaned, poverty decile bins, demeaned electoral threat indicators, and demeaned department indicators) are interacted across all factors in the design. All estimates use the IPW estimator. Standard errors are clustered at the municipality level since this is the level of measurement of the poverty moderator.

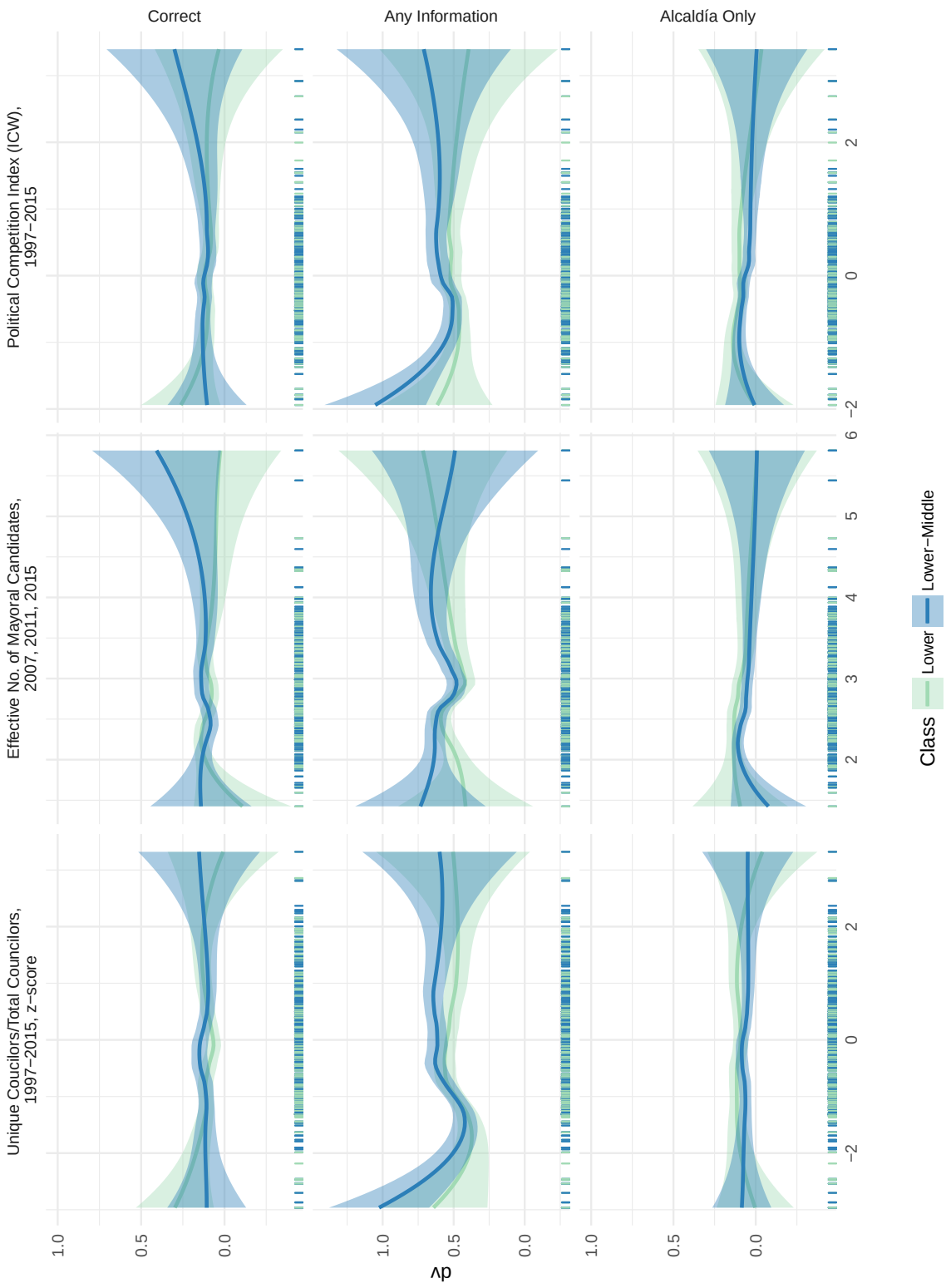


Figure A16: There is no evidence that class-based bias varies in measures of local political competition. Lines estimated by local polynomial regression (Loess) with 95% confidence intervals.

A18 Robustness of Link to Administrative SISBÉN Enrollment

For the analysis of SISBÉN enrollment and bias in Section 7, I examine the robustness of the classification of "plausibly intended enrollment." In the main text, this category encompasses any municipality for which enrollment falls between the number of individuals in poverty and the population. However, the "plausible" category could also include places with substantial over-enrollment. I examine the robustness of the finding to redefinition of this category. Specifically, I define this category as:

$$\text{Plausible} \in [\text{Poverty Rate}, \min\{j + \text{Poverty Rate}, 1\}] \tag{13}$$

for $j \in [0.4, 1]$. Note that the main definition assumes that $j = 1$. The revised scatter plot illustrating this coding is graphed in Figure A17.

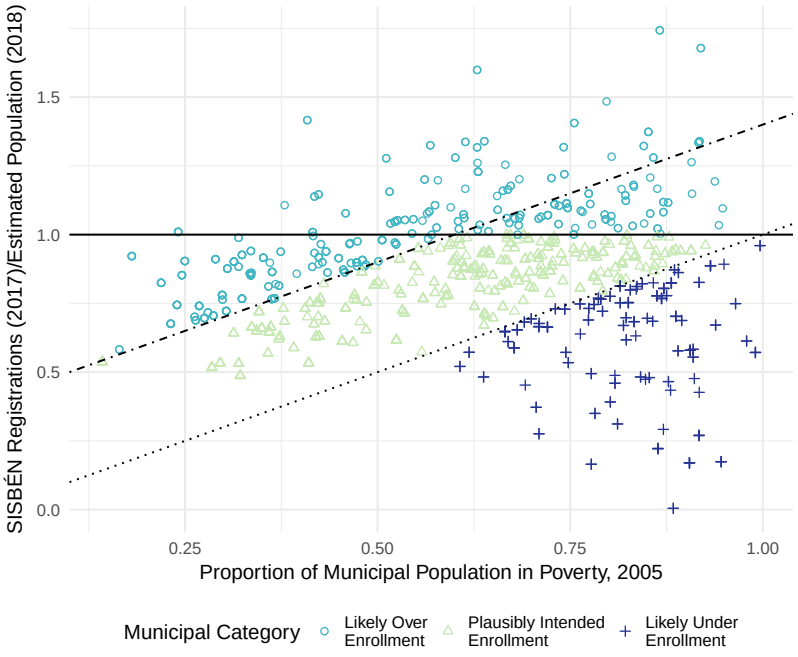


Figure A17: Visualization of the redefinition of "plausible enrollment" for $j = 0.4$.

Re-estimating Panel A of Table 5 with this specification, Figure A18 indicates the the point estimates for Plausible Enrollment and the difference (lower panel) are remarkably similar

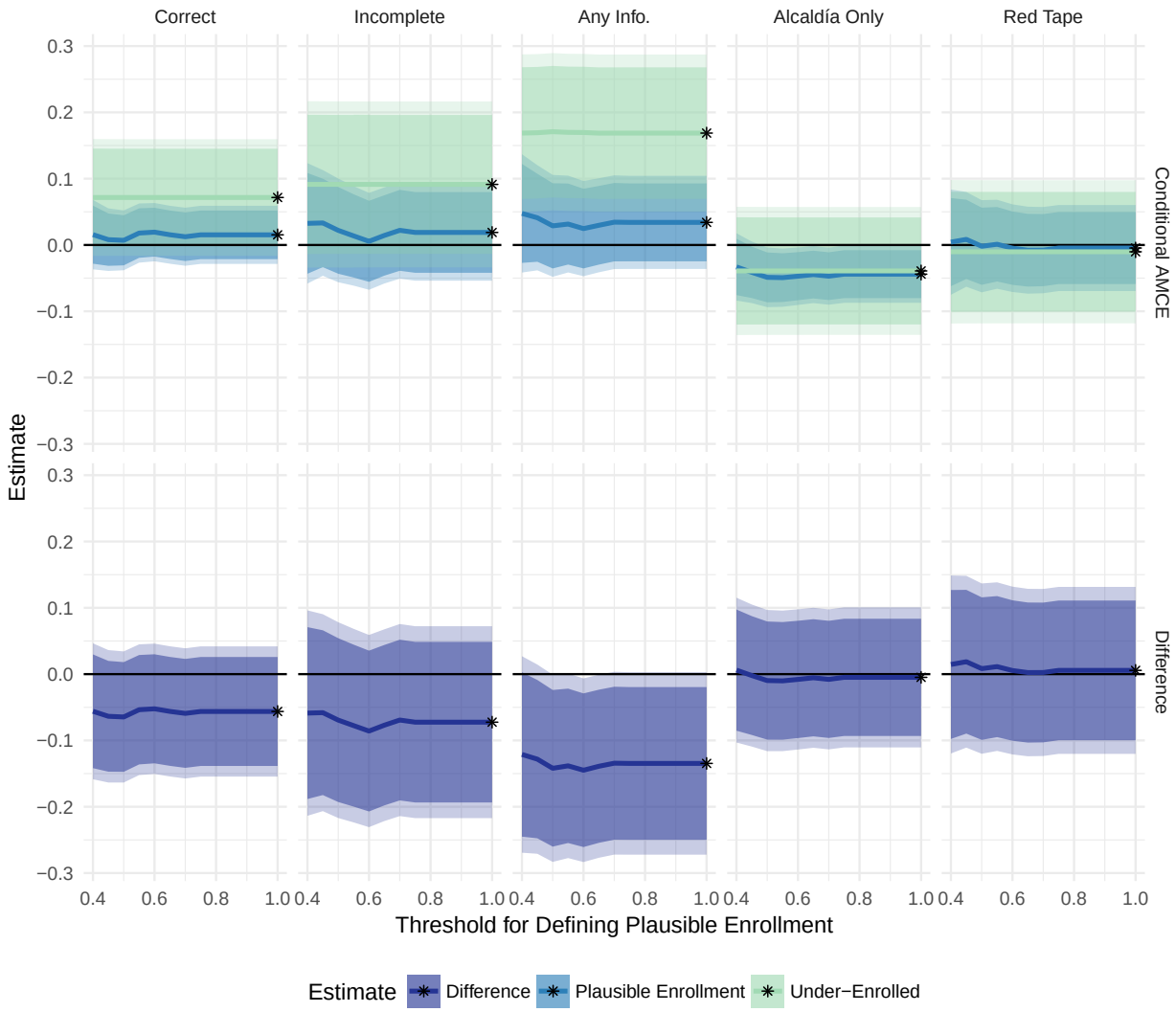


Figure A18: Robustness of the results in Table 5, Panel A to redefining "plausible enrollment." The x-axis corresponds to j in Equation 13. The stars represent the estimates reported (or implied) by Table 5 Panel A. 90 and 95% confidence intervals calculated on the basis of cluster robust standard errors.

Supplementary Appendix: References

Eframidis, Pavlos S., and Paul G. Spirakis. 2006. "Weighted Random Sampling with a Reservoir." *Information Processing Letters* 97 (5): 181–185.

Mares, Isabela, and Lauren Young. 2018. "The Core Voter's Curse: Clientelistic Threats and Promises in Hungarian Elections." *Comparative Political Studies* 51 (11): 1441–1471.

Misión de Observación Electoral. 2018. Mapas y factores de resgo electoral: Elecciones nacionales Colombia 2018. Technical report Misión de Observación Electoral Bogotá, D.C.: .

Ocampo, Gloria Isabel. 2014. *Poders Regionales, Clientelismo, y Estado: Etnografía del Poder y La Política en Córdoba, Colombia*. Bogotá, D.C.: ODECOFI-CINEP.

Weitz-Shapiro, Rebecca. 2012. "What Wins Votes: Why Some Politicians Opt Out of Clientelism." *American Journal of Political Science* 56 (3): 568–583.