Post-electoral promises, performance and trust in government

Evidence from a Survey Experiment*

Susana Otálvaro-Ramírez[1], Carlos Scartascini[1¶], Jorge Streb[2]

[1] Inter-American Development Bank. 1300 New York Ave, NW, Washington, DC 20577
[2] Universidad del CEMA. Av. Córdoba 374, Autonomous City of Buenos Aires, Argentina.
[¶] Corresponding author. E-mail. carlossc@iadb.org.

Abstract

Trust affects the level and composition of government spending and growth. Increasing information and empowering citizens are well-known tools to increase trust. However, these initiatives usually depend on having a government official willing to introduce them. In this paper, we show theoretically that there are conditions under which transparency initiatives, in the form of post-electoral commitments, are incentive-compatible for politicians. We also seek to understand how citizens react to various forms of information and whether releasing that information increases trust by relying on a randomized survey experiment conducted in the City of Buenos Aires, Argentina. The city government made a set of post-electoral promises, which can be interpreted as an additional signal to government performance in an incomplete information setup. We expect the effect of transparency to depend on government performance. Results show that post-electoral promises matter for shaping citizens' perceptions about trustworthiness of the government. The content of the information also matters: those who received a treatment showing that the government was fulfilling its promises trusted the government more than those who received a treatment in which they only learned that the government had made the promises. There was strong heterogeneity among citizens. One group, who did not know of the policy and had the lowest trust priors, was impervious to treatment: they seem to react to deeds, not words. The other group, which reacts positively to the promises, can be divided in two: those who already knew about the promises had the highest trust priors, but the treatment effects are entirely through the group that had heard of the promises, and had intermediate trust priors; the treatment effects close the gap in trust with the group that already knew about the promises.

Keywords: Trust, Political Economy, Development, Survey Experiment.
JEL codes: D02, D7, D82, H41

*The information and opinions presented herein are entirely those of the authors, and no endorsement by the Inter-American Development Bank, its Board of Executive Directors, or the countries they represent is expressed or implied.

1 Introduction

Public expenditure in developing countries is inefficient and biased towards current spending. Latin America and the Caribbean is the region with the highest gap in long-term public investment, and high transfers ans subsidies that redistribute little (Izquierdo et al., 2018). Mistrust, both political and interpersonal, plays a significant role in shaping public policy preferences for public spending (Keefer and Scartascini, 2022). Citizens who show low trust prefer politicians who favor transfers over public goods, and who promise immediate and certain benefits, even if they don't yield long-term growth (Yamamura, 2014; Anderson, 2017; Keefer et al., 2018; Cai et al., 2020; Keefer et al., 2020, 2022).

Trust can be defined as the belief that others will not act opportunistically given the opportunity. The ability of people, firms, or governments to take advantage over others is more prevalent when there are information and power asymmetries (Keefer and Scartascini, 2022). Therefore, reforms aimed at fostering trust should inform and empower. For example, enhancing communication and information strategies about policy outcomes can increase trust (Grimmelikhuijsen, 2012; Alessandro et al., 2021; Butler et al., 2021). Voters' access to information can also empower citizens by bolstering their power to reward or censure elected officials for their performance in office (Ferraz and Finan, 2008; Kendall et al., 2015). Reliable information that has a direct connection to citizens' well-being is more likely to increase their willingness to hold the government accountable (Khemani et al., 2016). That is why transparency initiatives have been a front-line step to improve individuals' perception of a trustworthy government.

While there are examples of successful initiatives that have helped increase trust by increasing information (Alessandro et al., 2021; Ardanaz et al., 2022; Keefer and Scartascini, 2022), evidence on the particular conditions under which transparency fosters trust is still scant, especially in the context of developing countries and young democracies (Blanco and Ruiz, 2013; Evans et al., 2019). Moreover, little is known about the conditions under which introducing trust-building initiatives is incentive-compatible for policymakers.

The purpose of this study is to identify the conditions under which politicians could have an incentive to commit to transparency initiatives in the post-electoral period (theoretically), and the impact of such an initiative in trust in the government (with the use of a survey experiment). Theoretically, we analyze the role of different types of information when the respondents face

incomplete information about the characteristics of the government. We set up a two-period model in which having positive reputation boosts the likelihood of reelection; as trust is a valence issue, incumbents that are seen as more trustworthy will have an advantage in relation to challengers (see Ansolabehere and Snyder, 2000). In a first stage, we develop a standard model where the signal the politician emit is the provision of public goods. In the model, a government that "steals less" can deliver more public goods. In a second stage, we add the possibility of a second signal, a post-electoral commitment, and reports on the results of its plans. We find that this can work as a signal of trustworthiness if it is harder for dishonest, non-benevolent governments to provide this information because of cover-up costs. The higher the cover-up costs, the more likely it is that there is going to be a separating equilibria where the benevolent government implements the transparency policy and the non-benevolent does not.

Based on Bayesian updating, these models allow us to understand the two-way relationship between trust and transparency that we then take to the data. When voters have heterogeneous beliefs, those who have no trust in the government will be unaffected by the transparency initiative. The impact of transparency will thus depend on priors, i.e., initial trust. We expect the most substantial effect of the transparency treatment to be on those with intermediate priors of trust, because those with highest trust face a ceiling effect. We look at this prediction of the model in the data in two ways. First, looking at how trust changes along the full spectrum of priors. Second, looking at the effect of transparency conditional on perceived quality of government, a variable that is very closely related to initial trust. People with intermediate beliefs are indeed more responsive to information about the commitments and their fulfillment than those at the extremes. Additionally, the model predicts that the effect of the signal is conditional on an individual's starting knowledge of municipal government plans and actions: since the information set of the group already aware of the government's plans does not change with the treatment, information should only affect individuals unfamiliar with the commitments.

The empirical analysis is based on a survey experiment conducted in the City of Buenos Aires City where we provide information about government commitments and results. Providing access to information may allow citizens to update their expectations about the government (Grimmelikhuijsen, 2012). Still, as described in the theoretical model, increasing access to information may not be enough to change citizens' mistrust, because the impact of transparency

depends on initial trust in government: if people think they are being provided fake news by an untrustworthy government, the impact of information will be nil. Furthermore, since the effect of transparency on trust depends on whether the government is successful or not in its plans (Sances, 2021), it may be necessary not only to open up, but to make strong commitments that citizens can observe and use as a yardstick to evaluate the government's actions and keep it accountable.

We conducted a survey experiment on a random sample of 2,375 individuals in the city of Buenos Aires, Argentina (henceforth CABA, for its acronym in Spanish) to examine an actual policy setting in a developing country.We use a simple design, following the recommendations in Bouwman and Grimmelikhuijsen (2016), to evaluate the importance of information regarding commitments and their impact on trust.

Participants were assigned randomly to three different informational treatments and a control group. Each treatment arm provides information about a series of commitments made by the Mayor of Buenos Aires at the beginning of the government period (treatment 1). Treatments 2 and 3 provide information about compliance with the commitments.[1] Those two treatments vary in the level at which fulfillment of promises is presented: aggregate (city), or local level (commune -the city is divided into 15 communes).[2] The treated and control individuals were asked about their perceptions regarding the city government's degree of trustworthiness using a multidimensional approach that includes the components of trust listed by Grimmelikhuijsen (2012): competence, benevolence and honesty, and questions regarding trust in government members that we developed for previous work (Keefer et al., 2018, 2020).

Since nearly 60 percent of people from the control group consider the government transparent when asked directly about this feature of local government, we expect that treatments will have a positive and significant effect on trust ($T^1 > 0$, $T^2 > 0$, $T^3 > 0$), and that informing about the compliance with commitments should have a more substantial effect than informing only about the commitments ($T^2 \geq T^1$, $T^3 \geq T^1$). Results show that disclosing information (T^1) indeed increases the perception of trust in government by about 0.1 standard deviations (SD).

[1] The commitments and their fulfillment are publicly available on the municipal government's website https://www.buenosaires.gob.ar/compromisos.

[2] They also differ in more subtle ways such as the amount of information provided –for design reasons, treatment 2 provides information about four commitments while treatment 3, three commitments– and the type of commitments –the set of commitments at the city level is broader than the set of commitments at the commune level.

Effects tend to be relatively uniform across the different subcomponents of our main index. It raises the perceptions of competence by 0.08 SD, and our measure of honesty and benevolence by 0.10 SD, and a direct measure of trustworthiness by 0.13 SD. The coefficients for T^2 and T^3 are positive and statistically significant, and they tend to be larger than those of T^1 but are not statistically different from it.

While the data confirm the model's prediction that those who already knew about the commitments would be unaffected by the treatment, the model did not predict the existence of two distinct subgroups among people unfamiliar with the government's ex ante plans. A subgroup vaguely familiar with the post-electoral promises, i.e., those that had heard about it, reacted to the treatment assignment. A second subgroup that was completely unaware of the commitments, showed no response to the informational treatment. One interpretation of this second subgroup response is that these individuals acquire knowledge through direct observation rather than through third-party information (Hertz et al., 2021). While we cannot prove causality in our setting, there is suggestive evidence that this group of people does not rely on others (family, social, or traditional media) and even less on government to acquire information about the government performance. The model can accommodate this dynamics by differentiating among individuals according to their preferred learning modes.

Finally, the treatment effects on people who had only heard of the program, but were not well familiarized with it, are of the same magnitude as the initial gap in trust between people aware of the commitments and people who had only heard of them. This suggests that the impact of the vignettes is not only short term: the effects are between 0.14 and 0.19 SD . Hence, unlike Marvel (2016), the effects of the treatment (here, post-electoral promises) seem to be long-lasting, perhaps because they are backed by government performance that is seen as good by most respondents.

These results have important implications. First, they emphasize the importance of actively delivering information to most citizens to enhance trust. Second, they elicit that individuals may be unconcerned about the aggregation level at which the information is provided. Although the effect of commitments' compliance is consistently higher than the effect of general information about them, citizens equally change their trust levels when the government informs fulfillment of their promises at the city or the commune level. Moreover, all informational treatments elicited a feeling of caring and concern for the population by the city government. Consequently, these

results should encourage to follow these types of policies: making and keeping promises appears to be rewarded with greater trust, in a context where a great deal of citizens already have substantial initial trust in government and perceive on average that government quality is high. Third, the paper shows that there are conditions under which transparency initiatives can be incentive-compatible for politicians. Some interventions that increase what we call "coverage costs" could have substantial effects in increasing transparency and trust in the public sector.

The paper proceeds as follows. Section 2 introduces a model that examines the role of transparency reforms when voters face incomplete information scenarios. In Section 3 we describe the survey experiment that explores the effect of a transparency initiative on political trust. Section 4 describes the data and empirical strategy. Finally, we present a discussion on the results in Section 5, and section 6 concludes.

2 Performance and post-electoral commitments as signals

We set up a work-horse model to understand how post-electoral commitments can affect trust in government, beyond the information provided by government performance. According to the model, only new information can affect beliefs.[3]

To capture why governments may be interested in showing that they are trustworthy, we use a two-period model in which a good reputation increases the chances of reelection.[4] The literature on rational retrospective voting that starts with Rogoff and Sibert (1988) already shows that information on government performance matter.[5] Why may post-electoral commitments matter too? Our explanation is that they may act as an additional signal about the government type by increasing transparency .

In the specific context of this study, we define political trust as a multidimensional characteristic that is influenced by government benevolence, honesty, and competence.[6] These traits go together naturally: a benevolent government does not divert resources to its own pockets, so it has nothing to hide, and hence can be open and honest about what it does; furthermore, it

[3] In the experiment below, we control for the previous knowledge of the survey participants, since we only expect the treatment to affect those who did not already know the commitments.

[4] The logic is similar in an infinite-horizon model where the incumbent faces a two-term limit.

[5] Ashworth (2012) summarizes theoretical and empirical literature on this.

[6] Regarding the influence of government performance on beliefs in our sample, the respondents in the control group already associate the direct measure of trustworthiness and its three components, benevolence, honesty, and competence, closely. According to parallel and factor analyses a single factor drives the responses.

can provide more public goods with a given budget, so citizens will perceive it as more competent. We formalize this in a standard signaling model where governments differ in benevolence, and the provision of public goods works as a signal of government type. If a better reputation improves reelection chances, non-benevolent incumbents are tempted to send the same signal as benevolent governments in the first period.

We then develop an alternative signaling model to incorporate the impact of post-electoral commitments on respondents in the treatment group, who trust the government more, and see it as more benevolent, honest, and competent. We model post-electoral commitments as a costly signal linked to these traits: post-electoral commitments increase transparency, revealing information about government goals and performance, because they require cover-up costs for dishonest government types.[7]

Voters

We divide public goods into more and less visible classes. While visible public goods are immediately observed by the voter, non-visible public goods are observed later. Visible public goods g_v go from varieties 0 to υ, while non-visible public goods g_{nv} go from υ to 1. The utility is the same for each variety, so governments will want to provide the same quantity within each group:

$$u(g_{vt}, g_{nvt}) = \upsilon u(g_{vt}) + (1 - \upsilon)u(g_{nvt}).$$

The per-period utility u is concave in the consumption of public goods. We further assume that utility is logarithmic in consumption, $u(.) = \ln(.)$, leading to an explicit analytical solution in the standard and alternative models.

Each individual voter i is subject to a political shock σ_i that has an idiosyncratic component identically and independently distributed over time. It represents the relative preferences for the opposition party in relation to the incumbent party.

$$u_{it} = u(g_{vt}, g_{nvt}) + \sigma_{it}.$$

[7]Since post-electoral commitments involve an administrative reform that can improve the monitoring of the public administration, this might enhance the provision of public goods. We abstract from this issue in what follows.

We take this shock to reflect differences in the utility of visible public goods due to differences in the perceptions of their provision. In the model, differences in the provision of public goods turn out to reflect differences in benevolence, so these shocks will affect how trustworthy the government is seen to be. From the data, we know that perceptions of trustworthiness and its components in the control group are heterogeneous, so this additive shock is a way to capture the different priors in the population. We assume the shock is uniformly distributed around zero, so the median voter $i = m$ is not affected by the idiosyncratic shock.[8] A voter's expected utility is given by the discounted sum $\mathbb{E}_t[U_{it}] = \mathbb{E}_t[\sum_{t=1}^{2} \delta^{t-1} u_{it}]$.

The government

Incumbents differ in benevolence. Differences in benevolence are captured by whether the per-period utility u of the incumbent is only determined by what is best for voters, or it also depends on an additional term r of personal rents:

$$u_{jt} = u(g_{vt}, g_{nvt}) + \alpha_j r_t,$$

where a benevolent government, $j = b$, has $\alpha_b = 0$, while a non-benevolent government, $j = nb$, has $\alpha_{nb} = \alpha > \underline{\alpha}$, where $\underline{\alpha} > 0$ is the threshold level beyond which personal rents are positive, as shown below. The expected utility of the incumbent is given by $\mathbb{E}_t[U_{jt}] = \mathbb{E}_t[\sum_{t=1}^{2} \delta^{t-1} u_{jt}]$.

By the per-period budget constraint, which is in per-capita terms, government expenditures γ_s, for $s = v, nv$, plus rents r appropriated by the incumbent, equal tax revenues τ:

$$\upsilon\gamma_{vt} + (1 - \upsilon)\gamma_{nvt} + r_t = \tau.$$

In the standard model, public expenditure determines the provision of public goods $s = v, nv$:

$$g_{st} = \gamma_{st}.$$

The alternative model introduces post-electoral commitments that increase transparency. If the incumbent implements a reform, the production functions for $s = v, nv$ no longer depend on

[8] This makes voting deterministic, because candidates have complete information about the median voter that is decisive in choosing between the incumbent and the challenger. If there were also a common component in the political shock, it would make voting probabilistic.

public expenditure γ alone, but also on whether the incumbent acts honestly or not, $l = h, nh$:

$$g_{st} = \gamma_{st} \frac{1}{1 + \chi_{lt}}.$$

If incumbents act honestly, $l = h$, the administrative reform imposes no costs: χ_h=0. On the other hand, if they act dishonestly, $l = nh$, while a proportion q_{nc} faces no costs, χ_{nbnc}=0, a proportion $q_c = 1 - q_{nc}$ faces a cover-up cost of $\chi_{nbc} = \kappa > 0$, because a more transparent public administration system makes it harder for them to hide the diversion of resources for personal rents.

2.1 Standard model: visible public goods as a signal

In the standard model, the signal to voters is the provision of visible public goods. This setup characterizes the control group that has not been exposed to the post-electoral commitments.

Incumbents differ in the degree of benevolence. The priors are that there is a proportion β of benevolent incumbents and a proportion $1 - \beta$ of non-benevolent ones. Benevolence turns out to be closely tied to perceptions of honesty and competence. While benevolent governments share the objectives of voters, so that they can be honest, non-benevolent incumbents cannot be honest once they deviate resources from the public treasure to their own pockets. Though incumbents do not differ in intrinsic competence, voters will perceive benevolent governments as more competent because they provide more public goods for a given level of tax revenues.

In the second period there are no reputational concerns, so each type picks its preferred level of public goods provision. Benevolent types pick $\left(\gamma^b_{vt+1}, \gamma^b_{nvt+1}\right) = \left(\frac{\tau}{2}, \frac{\tau}{2}\right)$, since they do not appropriate any personal rents. Non-benevolent types pick instead $\left(\gamma^{nb}_{vt+1}, \gamma^{nb}_{nvt+1}\right) = \left(\frac{1}{\alpha}, \frac{1}{\alpha}\right)$, given that α is the constant marginal utility of private rents for non-benevolent incumbents, and $\frac{1}{\alpha} = u^{-1}_{g_{st+1}}(\alpha)$ for $s = v, nv$.[9]

In the first period, voters will want to reelect a benevolent incumbent and replace a non-benevolent one. This introduces reputational concerns in the model, since a good reputation is important to get reelected. In this setup, announcements are cheap talk. Though a benevolent government, which shares the objectives of voters, has no problem in announcing what it actually plans to do, because it has nothing to hide, a non-benevolent government must be willing to lie

[9] Above the threshold $\underline{\alpha}$, the higher α is, the lower the provision of public goods. With log utility, the threshold is determined by $\tau : \underline{\alpha} = \frac{2}{\tau}$; below that threshold, rents are null.

if it deviates resources from the public treasure to its own pockets. Hence, only the provision of visible public goods counts. The equilibrium can be either pooling or separating.

If there is a pooling equilibrium, the non-benevolent type mimics the provision of visible public goods by the benevolent type, while all the rents are extracted from the under-provision of non-visible public goods: $\left(\gamma_{vt}^{b}, \gamma_{nvt}^{b}\right) = \left(\frac{\tau}{2}, \frac{\tau}{2}\right)$, $\left(\gamma_{vt}^{nb}, \gamma_{nvt}^{nb}\right) = \left(\frac{\tau}{2}, \frac{1}{\alpha_{nb}}\right)$. Plugging the optimal $t+1$ solutions derived above in the utility function $U(.)$, in a pooling equilibrium the expected indirect utility of the non-benevolent type has to be larger if it mimics the benevolent type:

$$\mathbb{E}\left[V^{nb}(\gamma_{vt}^{b}, \gamma_{nvt}^{nb})\right] > \mathbb{E}\left[V^{nb}(\gamma_{vt}^{nb}, \gamma_{nvt}^{nb})\right]. \tag{1}$$

If the non-benevolent type mimics the benevolent type, it will have a positive probability of reelection $P(\gamma_{vt}^{b})$:

$$\begin{aligned}\mathbb{E}\left[V^{nb}(\gamma_{vt}^{b}, \gamma_{nvt}^{nb})\right] =& u(\gamma_{vt}^{b}, \gamma_{nvt}^{nb}) + \alpha(\tau - \upsilon\gamma_{vt}^{b} - (1-\upsilon)\gamma_{nvt}^{nb}) \\ &+ \delta P(\gamma_{vt}^{b})\left[u(\gamma_{vt+1}^{nb}, \gamma_{nvt+1}^{nb}) + \alpha(\tau - \upsilon\gamma_{vt+1}^{nb} - (1-\upsilon)\gamma_{nvt+1}^{nb})\right] \\ &+ \delta(1 - P(\gamma_{vt}^{b}))\left[\beta u(\gamma_{vt+1}^{b}, \gamma_{nvt+1}^{b}) + (1-\beta)u(\gamma_{vt+1}^{nb}, \gamma_{nvt+1}^{nb})\right].\end{aligned}$$

If it instead separates out,

$$\begin{aligned}\mathbb{E}\left[V^{nb}(\gamma_{vt}^{nb}, \gamma_{nvt}^{nb})\right] =& u(\gamma_{vt}^{nb}, \gamma_{nvt}^{nb}) + \alpha(\tau - \upsilon\gamma_{vt}^{nb} - (1-\upsilon)\gamma_{nvt}^{nb}) \\ &+ \delta\left[\beta u(\gamma_{vt+1}^{b}, \gamma_{nvt+1}^{b}) + (1-\beta)u(\gamma_{vt+1}^{nb}, \gamma_{nvt+1}^{nb})\right].\end{aligned}$$

Rearranging terms, a non-benevolent incumbent resigns some utility from rents today (net of utility gain from more visible public good provision now) if this opens the door to even more utility from rents in the future (net of utility loss from lower future provision of both types of public goods):

$$\begin{aligned}\alpha\upsilon(\gamma_{vt}^{b} - \gamma_{vt}^{nb}) - \upsilon\left[u(\gamma_{vt}^{b}) - u(\gamma_{vt}^{nb})\right] <& \delta P(\gamma_{vt}^{b})\alpha\left(\tau - \upsilon\gamma_{vt+1}^{nb} - (1-\upsilon)\gamma_{nvt+1}^{nb}\right) \\ &- \delta P(\gamma_{vt}^{b})\beta\left[u(\gamma_{vt+1}^{b}, \gamma_{nvt+1}^{b}) - u(\gamma_{vt+1}^{nb}, .\gamma_{nvt+1}^{nb})\right]\end{aligned}$$

The median voter will be indifferent between the incumbent and the challenger in a pooling equilibrium. Expected utility in the second period is the same with either candidate: there will

be a proportion β of benevolent incumbents and a proportion $1-\beta$ of non-benevolent ones. Hence, the probability of reelection $P(\gamma_{vt}^{b}) \in [0,1]$.

We assume that $P(\gamma_{vt}^{b}) = 1$ if the median voter is indifferent, so the government is reelected for sure. This drastically simplifies the analysis, because there is a unique equilibrium: if condition 1 holds with $P(\gamma_{vt}^{b}) = 1$, the equilibrium is indeed pooling, because the non-benevolent type wants to mimic the provision of visible public goods by the benevolent type; if not, so a non-benevolent type does not want to pool even with the best possible future payoffs, there is a separating equilibrium where a non-benevolent incumbent chooses a lower provision of visible public goods, $\gamma_{nvt}^{nb} = \frac{1}{\alpha}$, which acts as a separating signal.

If we assumed instead that $P(\gamma_{vt}^{b}) = \frac{1}{2}$ when the median voter is indifferent, multiple equilibria arise. Besides the equilibrium where the pooling signal is $\gamma_{vt} = \tau/2$, there can also be an equilibrium with a separating signal $\gamma_{vt} > \tau/2$, in which the benevolent incumbent provides an extraordinary amount of visible public goods. The drawback for benevolent governments is that this signal has a welfare cost, because it strongly distorts the optimal provision of public goods. We explore below an alternative way for benevolent governments to separate out by using a novel signal: post-electoral commitments.

In this standard setup, the pooling equilibrium is consistent with the heterogeneous priors of the control group, as we find in the data. A separating equilibrium would instead point to either complete trust or distrust in the government, except for individuals with extreme priors of either zero (no trust whatsoever in the government) or one (complete trust in the government), who are not affected by empirical evidence. We return to this heterogeneity issue below.

2.2 Alternative model: post-electoral commitments as a signal

The setup changes as follows in the alternative model: after the government decides its budget, but before the voters observe government performance, the government can announce an administrative reform that increases transparency, by publicly announcing specific goals and promising to report about their fulfillment. This characterizes the scenario faced by individuals assigned to the treatment group in our experimental design.

These post-electoral commitments can be formalized as a signal that has differential costs for benevolent and non-benevolent governments. Post-electoral commitments allow honesty to have a bite, insofar as this signal is costlier if a government must cover up what it is actually doing.

Honest governments face no costs in this regard, since they do not have to cover up anything. Hence, they will not have problems making commitments. Non-benevolent governments do face a cost of being dishonest, so these announcements are not cheap talk: cover-up costs negatively affect the provision of public goods. This reform thus works as a potential signal of trustworthiness.

With log-utility, the benefits of adopting a reform are additively separable from those of the provision of public goods. In the second period, there are no reputational concerns, so each type picks its preferred level of public goods provision. As with the standard model, benevolent types pick $\left(\gamma^{b}_{vt+1}, \gamma^{b}_{nvt+1}\right) = \left(\frac{\tau}{2}, \frac{\tau}{2}\right)$, since they do not appropriate any personal rents, while non-benevolent types pick $\left(\gamma^{nb}_{vt+1}, \gamma^{nb}_{nvt+1}\right) = \left(\frac{1}{\alpha_{nb}}, \frac{1}{\alpha_{nb}}\right)$.

A benevolent type assigns the full budget in the first period to the provision of public goods, $\left(\gamma^{b}_{vt+1}, \gamma^{b}_{nvt+1}\right) = \left(\frac{\tau}{2}, \frac{\tau}{2}\right)$. A benevolent incumbent will always be willing to launch the reform if it increases its reputation of trustworthiness, because the reform has no current costs but future benefits, raising its probability of reelection.

While benevolent types will always want to implement the reform, the choice of non-benevolent types depends on their cover-up costs if they act dishonestly. If a non-benevolent type nb mimics the provision of visible public goods undertaken by a benevolent type, the benefits of a reform for types who act dishonestly and have type $k = nc$ are not affected, but types $k = c$ are negatively affected by the cover-up costs $\kappa > 0$ of the rents extracted from the under provision of non-visible public goods in the first period:

$$B^{nbc} = -(1 - \upsilon)\ln(1 + \kappa)$$

The reform can work as a semi-separating signal if it is only adopted by benevolent types and by non-benevolent types that face no cover-up costs. Using the indirect utility function $V(.)$, the expected utility for the non-benevolent type that face cover-up costs by mimicking the visible expenditure of benevolent types and adopting the reform, Π has to be smaller than if it does not, $\sim \Pi$,

$$\mathbb{E}\left[V^{nbc}\left(\gamma^{b}_{vt}, \gamma^{nb}_{nvt}, \Pi\right)\right] \leq \mathbb{E}\left[V^{nbc}\left(\gamma^{nb}_{vt}, \gamma^{nb}_{nvt}, \sim \Pi\right)\right] \tag{2}$$

Combining the results from the standard model with the additional term that captures the

impact of the reform, these conditions imply that

$$\mathbb{E}\left[V^{nbc}\left(\gamma_{vt}^{b},\gamma_{nvt}^{nb},\Pi\right)\right]-\mathbb{E}\left[V^{nbc}\left(\gamma_{vt}^{nb},\gamma_{nvt}^{nb},\sim\Pi\right)\right]=\mathbb{E}\left[V^{nb}\left(\gamma_{vt}^{b},\gamma_{nvt}^{nb}\right)\right]-\mathbb{E}\left[V^{nb}\left(\gamma_{vt}^{nb},\gamma_{nvt}^{nb}\right)\right]+B^{nbc}\leq 0$$

We know that $B^{nbc}<0$. If there is a pooling equilibrium in the standard model, there will be a semi-separating equilibrium if B^{nbc} is sufficiently large in absolute value, so that non-benevolent incumbents that face cover-up costs κ do not want to implement the reform.

If the equilibrium were pooling, the treatment group would have the same posteriors as the control group, contrary to the experimental results we show below. If the equilibrium were instead separating, then the treatment group would have completely updated its beliefs, except for respondents with priors that either show no trust or complete trust in the government. A semi-separating equilibrium falls in between both these cases: the posteriors of respondents will be affected partially. The evidence below indeed shows that the posteriors regarding trust and its components improve, but they are still heterogeneous, pointing to a semi-separating equilibrium that resembles the outcome in Alessandro et al. (2021).

If a reform is carried out, Bayes' law gives the following posterior for the trustworthiness of the incumbent, i.e., the belief that the incumbent is benevolent:

$$\mu(b|\Pi)=\frac{p_b}{p_b+(1-q_c)(1-p_b)}>p_b, \text{if } q_c>0.$$

According to this equation, the impact of the treatment on beliefs will depend on the priors and the parameter q_c, which gives the probability that a non-benevolent government will not launch the administrative reform. We later calibrate the model to the data in the survey.

3 Survey Experiment

We used a random sample of 2,375 people from CABA to study an actual transparency initiative to assess the value of information about post-electoral commitments on political trust.[10] Horacio Rodríguez Larreta, the mayor of Buenos Aires, made a series of promises to residents upon

[10]The City of Buenos Aires is Argentina's capital and most populous city. It is subdivided in 15 *comunas* which work as territorial, administrative and political entities, and include 48 neighborhoods -see Figure B15 for details.

taking office as part of his transparency promotion strategy, complementing the city's ongoing effort.[11] These commitments are specific and quantifiable objectives that span government sectors and are based on citizens' interests and the United Nations' Sustainable Development Goals (UN). Over 50 goals are outlined and their progress toward compliance is reported on the local government's website.[12] We use these commitments to examine the influence on trust of learning about post-electoral promises.

3.1 About the Survey

We designed an online questionnaire conducted in December 2019 to elicit information about trust in the government. It includes questions that attempt to capture individuals' perceptions about the competence, benevolence, and honesty of the government, and their confidence in politicians and public servants, following Grimmelikhuijsen (2012) and Keefer et al. (2018, 2020) – questions are available in the Supplemental Material B.[13] The sample was stratified with quotas by gender, age group (18 to 60 years old), and socioeconomic status. Within strata, individuals were assigned at random in one of four possible treatment categories: three informational vignettes, or a control group. Treated individuals answered the battery of questions on trust in the government after receiving the informational pieces. A total of 2,375 complete interviews were carried out by a company that specializes in collecting online survey data.[14] Figure A1 presents a description of the timeline of the survey experiment.

Each treatment presents information about the commitments made by the government. Treatment 1 mentions the existence of the mayoral promises, their relevance, and four examples without giving specific details or level of compliance (see Figure B12). Treatment 2 provides the same information as Treatment 1 and it also shows the government's performance fulfilling the promises it had made, with examples at the city level (see Figure B13). Treatment 3 provides the same general information about the commitments and presents a map with dots for all the achievements at the *comuna* level, highlighting three specific examples of promises fulfilled

[11] The city has been steadily increasing transparency over the last two decades (Alessandro et al., 2021). All the city government initiatives for increasing transparency are listed here https://www.buenosaires.gob.ar/ministerio-de-gobierno/transparencia.

[12] Government commitments are available at http://www.buenosaires.gob.ar/compromisos. For example, Rodríguez Larreta in public statement after election as Mayor of the city indicated: "We are committed to ensuring that, during this term of office, 20,000 families will be able to fulfill their dream of owning their home". Another example was building 8 educational centers, mainly located in vulnerable neighborhoods of the city.

[13] See Alessandro et al. (2021) for additional details. The full questionnaire in Spanish is available upon request.

[14] For details about the company, visit http://www.isonomia.com.ar/en/

at such level. It is important to note that individuals in Treatment 3 received an infographic designed for the specific *comuna* they report living in (see Figure B14) – individuals were asked in the characterization module about the *comuna* they live in; then, the survey program selected the infographic that matched it. Unfortunately, by the nature of the implementation of the experiment, there are differences between treatments 2 and 3 beyond the fact that the first provides examples of fulfillment at the city level and the latter at the *comuna* the respondent lives in level. On the one hand, the number of commitments displayed differs, 4 versus 3 promises are displayed, respectively. On the other, the type of commitments is also different and does not necessarily align with the policy area disclosed in treatment 2 (infrastructure investments, mainly).

4 Data and Empirical Analysis

4.1 Data Description

Table 2 presents descriptive statistics for the main observable characteristics of the respondents and balance on covariates measured before presenting the experimental vignette to participants. The first column shows the sample average and the standard deviation for the control group. Note that the average respondent is female (57%), completed secondary education (nearly 85% of surveyed individuals have completed high school), and is employed (52%).[15] Despite the existence of a dedicated website and the public announcement of commitments, only 22% of the sample knew about them before the intervention took place (see Figure A2).

The groups are well balanced; only 4 out of 45 differences are statistically significant at the 10% level, consistent with chance. Based on the balance on observable characteristics, we consider the randomization was successful.[16] Additionally, p-values of tests of equality of coefficients identify there are no systematic differences in sample composition across treatment groups. There are minor disparities between people assigned to the first and second treatment

[15]The sample is not necessarily representative of the overall population of the city. In particular, it is slightly more educated, older (average age is about 40 years old), and has a higher share of women than the city's population –which is about 53% according to the 2010 census data.

[16]We estimate an OLS regression with clustered errors at the *comuna* level to assess balance across treatment assignment. In this specification, observable characteristics act as dependent variables and treatment variables as independent. We also conducted randomization inference procedures over the difference in means between treated and control units to assess balance. Adjusted p-values from a thousand Montecarlo simulations of the OLS regressions provide similar conclusions.

groups regarding gender and educational attainment, with more women in the first treatment arm and more educated individuals in the second one.

Dependent variables are classified into two encompassing groups: (1) Trust in the Government and (2) Trust in Actions made by politicians and public servants.[17] For the first set of questions that evaluate trust in the government, we work with the individual responses to a series of 11 questions that inquired the participant to show her position on different statements about the Government of the City of Buenos Aires, ranging from *Strongly disagree* to *Strongly agree*. Those questions attempt to identify how the respondent feels about the competence (is capable? does what is best for the city? is innovative? thinks in the long term? plans and informs its plans?), benevolence (acts in the interests of residents? helps those in need? pursues policies my family cares about?), and honesty (is sincere? is transparent? fulfills its promises?) of the city government. To reduce the dimensionality of the information provided by the 11 questions, we construct summary indexes.[18] Although we have information on the trustworthiness perception, we use it as a direct measure of trust and do not include it in the estimation of the indexes.

To construct the indexes, we exploit a principal component analysis methodology (PCA henceforth) in which the first component explains at least 80 percent of the variance regardless of the index (see Table A2 in the SM).[19] We build three intermediate indexes (Competence, Benevolence, and Honesty) and one global index that summarizes all questions on perceptions.[20]

Note that citizens have a relatively high trust in the city government to begin with. Figure 1 provides a first glimpse at the perceptions of the respondents. It looks at the control group's responses on each aspect of government competence, benevolence and honesty. In general, there is a positive assessment of the government. Respondents consider the government to be capable and innovative, among the top attributes. However, they grade the government lower regarding how much it helps those in need and pursues programs that are beneficial to the families. We also observe heterogeneity in the initial distribution of perceptions about the government,

[17] Descriptive statistics on dependent variables are shown in Table A1 in the Supplementary Material (SM).

[18] The aggregation improves statistical power to detect effects that are consistent across specific outcomes when these specific outcomes also have idiosyncratic variation.

[19] We also applied a factor analysis methodology and consistently found that the first factor explains a significant proportion of the variation. Furthermore, in a parallel analysis, we see that we should stay with a single factor under a decision rule of thumb of one. However, we report results for indices constructed for each dimension of trust, competence, benevolence, and honesty for interpretation purposes.

[20] Robustness exercises include running the regressions with the individual questions. Conclusions on statistical inference remain the same when we correct p-values using the Westfall and Young procedure (Jones et al., 2019).

especially among people who are aware of the commitment policy (see Figure A3 in the SM). The remaining dependent variables will be explored in the discussion section.

4.2 Empirical Strategy

We first estimate the following model to understand the effect of providing information about government commitments and its fulfillment on government trust:

$$Y_{ic} = \alpha + \beta_1 T_i^1 + \beta_2 T_i^2 + \beta_3 T_i^3 + \lambda X_{ic} + \epsilon_c, \tag{3}$$

where T^n is the treatment assignment, $n = 1, 2, 3$, depending on the treatment individual i from *comuna* c was exposed to. The treatment arms are as follows: (1) Commitments, (2) Commitments and their fulfillment at the city level, and (3) Commitment and their fulfillment at the *comuna* level. A respondent was exposed to one treatment arm only, and individuals in each treatment arm were compared against those who did not receive any information. X_{ic} is a vector of controls that includes all observable characteristics available from the survey: age, gender, education (i.e., college or high school degree), labor status (i.e., employed or unemployed), socioeconomic level (ABC1 describes the group with the highest income distribution), revealed preferences for public budget allocation between education and infrastructure, and pre-treatment beliefs on government quality. We also include *comuna* fixed effects (i.e., geographical and political divisions within the city).

Given that 60 percent of respondents in control group consider that the government is transparent, so it says the truth, our expectation is that providing information about commitments and their fulfillment matters: $\beta_n > 0$. If trust depends on the expectations that other people will act in good faith and comply with what they promise, providing information about what the government is doing and its compliance should increase trust. Promising, if the government is not expected to comply with those promises, is not good policy in equilibrium; as such, promises by themselves already carry a load of information (Alessandro et al., 2021).[21] Still, because individuals may feel more comfortable in updating their perceptions if they are shown fulfillment, we expect $\beta_2 \geq \beta_1$.

[21] This project is carried out with political will and in conjunction with the City of Buenos Aires' local government. The government is aware that having made promises and reminding citizens of them has a meaningful information load.

Although showing information on the achievement of the goals at a more local level could be more informative than presenting the commitments alone, we do not have a strong prior on the differentiated effect of T^3 compared to T^1. First, because T^3 shows fewer commitments than those presented in T^1 due to logistical issues in the implementation phase. Second, T^3 does not consistently present commitments in policy areas comparable to those shown in treatments 1 and 2. While commitments in T^1 and T^2 mainly concentrate on urban mobility and infrastructure issues, the ones presented in T^3 include health and education, in addition to infrastructure projects. And third, some of the displayed vignettes in T^3 contain information that may not be informative for all citizens from those *comunas*. Imagine the case of information about a public school extended schedule; this exclusively affects those individuals with school-aged children who could be potential beneficiaries of such improvement. However, the overall impact might not be different from the one obtained from giving general information about commitments for the rest of the treated individuals. Our hypothesis is, therefore, that $\beta_3 \geq \beta_1 > 0$. We do not have strong priors regarding the relationship between β_3 and β_2.[22]

On the one hand, the targeted information could have added value for the recipient ("the government is not only doing what they promise but they are doing it in my *comuna*").[23] On the other, the information the individual receives will be automatically translated by the priors the individual has regarding the relative status of the *comuna* vis-à-vis other *comunas*. As such, the information may improve or worsen the priors the respondent has.

4.3 Results

The survey experiment aims to study how the provision of information to citizens shapes their perceptions about the government. Information should allow them to update their beliefs about

[22]Table A9 shows the results by comuna, considering the third treatment arm nature. Although individuals assigned to this arm were informed about the projects carried out in their commune and could compare the information received with their personal experience, we do not see significantly greater responses to informational vignettes than those who received information at the city level. There is no regular pattern in the relationship between treatments 2 and 3. As expected, the influence of tailored information on individuals' beliefs about the government is systematically stronger than generic information about the commitments. However, we do not observe statistically significant changes in treatment allocations (except for comuna 5 and 14). Because mixed results within communes could be explained by the fact that families living in different communes may have characteristics that cause them to respond differently to treatment, we offer in Table A10 a selection analysis that evaluates whether socio-demographic traits and ex-ante perceptions of the government are more widespread in particular communes than others.

[23]There is evidence that individuals update their priors more when the information they receive refers to a group closer or more homogeneous. See, for example, Miranda et al. (2020) for the case of water consumption. Still, that evidence may not travel well because the information is not about individuals in the *comuna* but government action that affected the *comuna*.

the government's reliability in developing its tasks, responsiveness to voters' needs, and fairness. We start by evaluating the effect of the different vignettes on indices that approximate the trust components we attempt to explain: a general composite index of trust in government and three sub-indexes that capture its perceived competence, benevolence, and honesty. Then, we estimate the effect on a direct measure of trustworthiness.

Figure 2 presents a graphical first approximation to the overall results of the intervention. Differences between the control and treated groups (pooled) are of about 0.10 standard deviations (SD) for the composite index and its sub-indices (referred herein as dimensions). Table 3 presents the compound index results, and each of the dimensions of trust identified by Grimmelikhuijsen (2012), looking at the disaggregated treatment arms level.[24] The first three columns suggest that information increases the perception of trust in the government, in general. Providing information about fulfillment at the city or *comuna*-level increases the perception of trust in 0.12 SD, on average, in a specification without controls. Once we include socio-demographic characteristics and *comuna* fixed effects, the effect of information on performance at the *comuna* level slightly decreases (0.104 SD). In contrast, the impact of information on performance at the city level marginally increases (0.125 SD).

We do not observe significant differences between providing performance information or just informing about the commitments.[25] Further, when we compare people who received information on performance at the aggregate -city- versus the local -*comuna*- level we do not observe differences in trust perceptions. However, the effect of Treatment 2 is consistently higher than the remaining two treatment arms. Learning about commitments is important to shape individuals' perceptions about the city government's trustworthiness, but providing detailed information on fulfillment at the city or *comuna*-level may not add much value.

Results on each dimension of the index are very similar to those of the composite index. Providing information about the commitments increases the perceptions of competence, benevolence, and honesty between 0.083 and 0.125 SD. The most remarkable effects are found in the competence dimension. Informing about the government commitments and its performance at the city level can increase the perception of competence in 0.125 SD. We see slightly lower impacts when we provide general information or locally targeted data, 0.083 and 0.107 SD,

[24] The table is constructed progressively, including control variables and *comuna* fixed effects.

[25] We conducted Wald tests of equality of coefficients in each estimation. We did not find any statistically different results with none of the informational treatments.

respectively. However, these minor differences among informational treatments are not statistically significant. General information about the commitments also increased the perception of a government that is benevolent and honest in 0.10 SD, approximately. Although there was more room for improvement in the honesty dimension (see Figure 1), informational treatments that provided more detailed data on performance did not manage to modify honesty perceptions in a greater magnitude than they did in the competence and benevolence dimensions.

While providing the answers to the different components that sum up the indexes previously addressed, people indicated the degree to which they considered the government trustworthy. The last column of Table 3 depicts the results of providing information on the direct measure of trustworthiness. Receiving both general information and performance data is positively associated with a more favorable assessment of the government's reliability, 0.11 SD higher than for those in the control group, on average. Interestingly, the dynamics of how information affects trust in the government do not hold. Although differences between treatment assignments are not statistically significant, people who receive general information about commitments increase their perceived trustworthiness more (in 0.127 SD) than those who received detailed information at the city level (0.114 SD). Information about the commitments and government performance at the *comuna* level increases to a lesser extent the perception of trustworthiness of the city government (0.081 SD).

Table 4 shows the results for each component of the dimensions. While conclusions prevail as in Table 3, there are some interesting findings. In the competence dimension, we detect a marginal increase in the perceived capacity and a sense of what is best for the city (between 0.06 and 0.11 SD) when people receive information about the commitments compared to those with no information. We observe larger gains in the perception of a government that thinks in the long-term (around 0.13 SD) and plans and informs its plans (between 0.10 and 0.18 SD). In general, information about 'Compromisos' and government performance at the city level lead to higher increases in perceptions, with no statistically differences within treatment arms, except for the case of the perception of a government that plans and informs its plans, where information on fulfillment at the city level increases the perception of competence to a greater extent than commitments themselves.

On the benevolence side, we find that providing information on performance translates into a greater perception of a government that acts in the interest of residents (0.08 SD). Receiving

information on commitments positively associates with a perception of the government that helps the people in need (around 0.13 SD) and pursues policies and projects that are beneficial for the families (0.11 SD). When detailed performance information is provided, the perception of benevolence increases to a greater, although non-statistically different, extent. In the honesty dimension, there are important improvements. There is an increase in the perception that the government is sincere and transparent between 0.07 and 0.10 SD for those who received information about commitments plus the ones who also received data on fulfillment. There are significant gains in the perception of a government that fulfills its promises, especially when people received information about the government's performance at the city level (0.17 SD).

From a Generalized Ordered Logit specification that considers all categories of the agreement scale for each component of the trust dimensions (see Table 5 and Figures A4 to A6 in the SM), we observe that greater levels of trust are led by changes in the middle part of the distribution. On average, people who were firmly in disagreement to somewhat in agreement move to higher levels of the trust distribution. When people are asked about the government's capacity to pursue projects that benefit their families, they tend to move from middle assessments to higher ones once presented with information about the "Compromisos" policy. Similar shifts take place when they think of long-run vision, competence, and the government's fulfillment of promises.

These results indicate that: (i) providing information about commitments is valuable for increasing trust, i.e., $\beta_n > 0$ in terms of equation 3; (ii) providing information about the fulfillment of those commitments seem to add some but little additional information $\beta_2 \geq \beta_1$; (iii) providing information at the *comuna* level does not increase trust more than providing information at the city level but we can not interpret these results given the caveats about the experiment design discussed earlier.

In our sample, supplementing information about fulfillment either at the city or comuna level does not add significantly more to trust than simply providing information about commitments. This is perhaps due to the fact that, besides the government transparency initiative, there is another signal at work: many of the respondents have already acquired much of that information first hand, by personal experience in the city streets, through visible public good provision, a potent, and classic, signal of government quality. Therefore, ex-ante information may generate cognitive dissonance with the information provided in the survey experiment. It can also happen that individuals know the developments in their commune but do not have much information

about the whole city, so they respond to this information to a greater extent.

The Generalized Ordered Logit model results indicate that political trust gains are mainly obtained through belief updating of individuals who have intermediate priors. We calibrate the model to see if it can replicate this pattern in the sample. Using the assumption that the responses are uniformly distributed within each category, we find that the model can indeed explain this pattern.

Table 1: Direct Measure of Trustworthiness: Calibration of Treatment Effect on Distribution

Category	Control Group	Treatment Group	Prediction with $q_c =$ 0.050	0.100	0.142	0.150	0.159	0.200
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	0.149	0.129	0.143	0.136	0.131	0.130	0.128	0.123
2	0.068	0.058	0.069	0.071	0.072	0.072	0.073	0.074
3	0.075	0.070	0.074	0.072	0.070	0.070	0.070	0.068
4	0.131	0.126	0.126	0.121	0.116	0.115	0.114	0.109
5	0.143	0.145	0.144	0.145	0.146	0.146	0.146	0.146
6	0.179	0.189	0.182	0.185	0.187	0.187	0.188	0.190
7	0.253	0.284	0.261	0.270	0.278	0.280	0.282	0.291
Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Proportion absolute deviations explained			0.277	0.459	0.582	0.608	0.623	0.407
Proportion squared deviations explained			0.437	0.712	0.791	0.788	0.778	0.621

Notes: There are seven categories of trust that go from 1 (low) to 7 (high). The last two lines present the proportion of the sum of the absolute differences between the control and treatment group, and the sum of their squared deviations, that can be explained by different values of parameter q_c.

When q_c is close to 0.15, the model correctly predicts that categories 5 to 7 increase, and categories 1 to 4 decrease, with one exception: the miss is that it predicts that category 2 will increase, when in fact it decreases. The model implies that the impact of the signal is greatest for the intermediate categories, especially categories 3 and 4: the percentage that shifts to the next category, for $q_c = 0.142$, is $0.124, 0.211, 0.259, 0.265, 0.226, 0.139, 0$, while for $q_c = 0.159$, it is $0.139, 0.238, 0.293, 0.300, 0.256, 0.158, 0$. These values of q_c suggest that respondents in the treatment group do not consider the administrative reform overly informative, but the effect is large enough to generate a significant difference between the control and treatment groups. With these parameter values, we can explain around 60 percent of the absolute differences between the control and treatment groups, and almost 80 percent of the squared differences, in the sample. However, this calibration has the limitation that it does not distinguish between

individuals who already know the commitments (which we call "Know"), and hence should not be affected by the treatment.

While these findings indicate that treatment has a considerable effect, we may be overlooking a high degree of heterogeneity among the respondents. According to the Bayesian updating process in the signaling model, the informational treatment should have an effect only on individuals who are unfamiliar with the government's pledges, while having no effect on those who are already familiar with them. We turn to this now.

4.4 Heterogeneous effects

To explore the affective dimension emphasized by Grimmelikhuijsen, 2012, pp. 56-7, we first look at how trust behaves for different priors on government quality. For instance, positive affective cues may lead to more favorable initial beliefs about the government's quality and trustworthiness among citizens who voted for the incumbent.

We then explore the model's prediction that individuals who have ex-ante information about commitments will respond very differently to the informational treatments than those that do not. We separate two categories of those unfamiliar with the pledges: those who are completely unaware with the post-election promises ("Do Not Know") and those who are only slightly familiar with the transparency program ("Heard of"). Finally, we examine the interaction of prior knowledge and perceived quality of government, attempting to disentangle the impacts of distinct information from those of various affect states.

Perceived Quality of the Government

Several studies have looked at how citizens' responses to information are influenced by their prior expectations and attitudes toward the government (Baekgaard, 2015; Marvel, 2016; Sances, 2021). These priors about the performance of the government may be dominant in the formation of citizens' preferences (Porumbescu et al., 2021; Barrows et al., 2016). As a result, people with different priors may respond differently to the informational treatment vignettes. Perceived quality could be understood as a proxy for political ideology. Individuals who gave the government a low rating are more likely to be opposed to it, while those who gave it a high rating are more likely to be supporters, therefore both ends could behave as poles.

Individuals can base their perceptions on their partisan commitments employing motivated reasoning to ensure that the preferences and perceptions correspond to their priors and partisanships (Bolsen et al., 2014; Slothuus and De Vreese, 2010). They can update their beliefs about the government's competence and benevolence by recalling their prior trust level, i.e., responding with an affect mechanism. People with extreme priors about the quality of the government, i.e., those with polarized political views, may rely on this mechanism to evaluate the city government after receiving information (Butler et al., 2021); it facilitates the cognitive process of analyzing acquired information.

Individuals in our setup who rated the quality of city government management as low are unlikely to incorporate new information into their perceptions about the city government and keep their previous level of confidence (see Figure A8 in the SM). Furthermore, those who have a very high initial assessment of the city administration may be unable to enhance their evaluation due to a ceiling effect.

Participants in the middle range of the distribution can benefit from the informational vignettes. They may rely on a learning mechanism to reevaluate their politicians' trustworthiness, building upon previously acquired information and beliefs about the governmental performance regarding the commitments, i.e., the effect arises as a result of changes in the information set available to respondents (Barrows et al., 2016). They are more likely to have heard of the program (see Figure A10), implying that they are aware of the pledges but may lack complete information about the government's effectiveness in reaching these objectives, and are less likely to be politically radical. As stated in Butler et al. (2021), those with a moderate view of politics are pivotal; they are less concerned with politicians' policy positions and more concerned with problem-solving. Moreover, as shown in Figure A8, the quality perception goes hand in hand with the level of trust individuals have in their government to begin with. Thus, the informational vignette may be beneficial in updating beliefs.

Our sample thinks highly of the city government's quality. People in the control group gave an average rating of 7.2 points on a scale of 1 to 10, with 1 being the lowest perceived quality and 10 being the highest. However, people appear to respond differently to the informational treats across the perceived quality distribution. Figure 3 shows the heterogeneous effects of perceived quality of the government on trust (see Table A5 in the SM for further details in a simplified version considering a general treatment condition).

Considering the sample sizes in each of the initial variable's categories and for interpretation purposes, we have narrowed the spectrum of perception of government quality to three categories to evaluate heterogeneous effects: low, medium, and high quality. Categories were constructed ad hoc to reflect their definition. Those who ranked the government quality between 1 and 3 were classified as low, 4 to 7 as intermediate, and 8 to 10 as high.

People who had a very high assessment of the government quality ex-ante do not significantly respond to the information provided through the vignettes, which is consistent with a ceiling effect or an affection mechanism. Participants who had an intermediate evaluation of the government's quality, on the other hand, were highly receptive to information. Individuals with the lowest assessments of the government quality have wide confidence intervals that do not reject the null hypothesis of no effect; however, they also show a positive response to information, although lower than those who have an intermediate evaluation of the government quality. As previously stated, those in the middle range of the trust and quality perception distribution can improve their assessments of the city government to a greater extent. These results go in line with the predictions of the model: the most remarkable trust changes happen among those with intermediate priors.

Prior knowledge

Our signaling model leads us to anticipate that prior knowledge of the promises can strongly affect the response to informational treatments. As pointed out previously, 22% of the participants knew about the commitments before implementing the survey experiment despite the existence of a dedicated website and its public announcement. Moreover, people who knew about the promises before taking the survey had different opinions of the local government than those who did not. Although we observe balance across treatment assignments among people with diverse levels of knowledge of the policy, there is huge heterogeneity in the assertions of the government's competence, benevolence and honesty among them. Figure A3 shows a positive gradient in the relationship between prior knowledge of the policy and assessments of the city government. Individuals within the control group who are unaware of or have never heard of the policy start with a lower level of trust in government (and each of its components). This level increases as people acquire information about the commitments.

Prior knowledge of the policy was not randomly assigned, yet from Table 2 we observe that

the proportion of people who have previously known about the policy is not statistically different among treatment status (which was randomly assigned). People who received information about commitments and their fulfillment at the city level are marginally more aware of the policy's existence. Considering that treatment arms have roughly the same proportion of people who know the policy ex-ante, and they are assigned at random, we explore how differing levels of previous knowledge of the policy modify trust in the government. We assess whether there were heterogeneous responses to the treatment among people who did not know the commitments previously, people who had heard of it, and those who did know it.

Suppose someone already had information about the commitments and the government's performance in fulfilling them in their information set before taking the survey $I(t = 0)$. Once we provide information about the policy, the information set remains unchanged, i.e., vignettes may provide redundant or irrelevant information. We observe that people who previously knew the commitments already have high assessments of the city government's competence, benevolence, and honesty; thus, they may have already incorporated this information into their trust perceptions (see Figures 4 and A3). On the other hand, those who have a vague idea of 'Compromisos' or do not know anything about it could modify their information set with the data we provide in the vignette, as in Butler et al. (2021). Compared to those who knew 'Compromisos', we expect that people who did not know the policy and those who had heard about it but did not know it for sure respond to the informational treatment to a greater extent.

Figure 5 shows the heterogeneous effects of previous knowledge of the governmental promises (Table A3 in the SM provides the regression findings of a simplified version interacting previous knowledge with ever been treated, regardless of the information received). As before, we observe that the second treatment arm had a slightly higher impact on trust than T1 and T3, although the difference is not statistically significant. However, this result is heterogeneous depending on the initial information set.

Empirical evidence shows that transparency does not always have positive impacts on governmental perceptions; in some cases, it can even have detrimental effects (Piotrowski et al., 2019). In particular, citizens who actively engage with politicians and the political process might be more critical when presented with information about the government's achievements. In our case, it is just the opposite: the people with a more extensive set of information ex-ante, who are not responsive at all to the vignettes, are also the people who seem to inform themselves

actively and have higher levels of trust than the rest. Hence, being more informed does not necessarily lead to lower trust if deeds back the government's words. The fact that people who previously knew about the 'Compromisos' policy do not significantly respond to the information provided through the vignettes confirms the model's prediction that those who already knew about the commitments would be unaffected by the treatment.

Participants who had only heard about the policy, on the other hand, were much more receptive to information than those who were exposed for the first time. The model's initial prediction that the information would allow updating the beliefs of those with little or no information about the commitments is not so straightforward. Those who have heard about the policy, rather than those who know about it for the first time through the experiment, are the ones who move the effect of information on trust, regardless of the content of the informational vignette. In fact, in Table A3 we do not reject the null hypothesis that all treatments close the gap in the initial perception of trust between those that knew about the commitments for sure and those that had only heard something about them. On the contrary, for individuals who are first exposed to the commitments through the survey, we reject the null hypothesis that the treatment closes the gap with the trust perceptions of those that knew about the policy. Thus, the treatment effect, while reducing the negative perception of the government, fails to offset the initial view held by individuals in the control group who were unaware of the commitments.[26]

The differentiation in the categories of the knowledge of 'Compromisos' variable indirectly considers the sources through which individuals gather information. People who knew the commitments and those who had heard about them obtained information about the city government management differently. Those who were more knowledgeable about the commitments at the beginning of the survey obtained information primarily from official sources. They were 9.3 pp more likely to have known them by attending a mayor's presentation, 8.1 pp by visiting the city's website, and 3.6 pp by reading CABA articles or notes. Around 60% of the people who reported knowing about 'Compromisos' had visited its website, and an additional 15% were aware of its existence. People who had heard of the policy but did not know what it was about obtained information through social media and pamphlets, essentially (see Table A4), with only a small percentage visiting the policy website. Since people who were barely aware of the policy

[26]Figure A9 in the SM similarly shows how initial levels of confidence are distributed differently according to the knowledge of the policy rank, and how the information has a relevant effect for the group that is vaguely knowledgeable about the policy while it is irrelevant for the extremes.

had learned about it through non-traditional channels, the provision of verified/truthful information may have helped them update their attitudes about the city government management more than those who knew nothing. In an instance in which a person initially judges the government based on inaccurate perceptions of its performance, she can adjust her judgment when she is provided accurate information. Those who had heard about the policy may update their "wrong" beliefs since they were primarily informed by social media, which may be subject to incomplete, unverified or biased information (Aruguete et al., 2021; Lupu et al., 2020). Thus, learning effects may occur in a contexts in which citizens have access to multiple sources of information from their own and others' experiences that may involve people forming beliefs based on untruthful aspects of government management - "wrong" beliefs- (Barrows et al., 2016).

The effects vanish when we analyze the extreme pole of no prior knowledge of the commitments. The null effect among the individuals that had no previous knowledge of the commitments might be related to the way in which they gather information. These citizens are more likely to acquire and trust knowledge that comes from direct observation rather than third-party information (Hertz et al., 2021). Figure A11 suggests that this group of people does not trust information regarding the government performance coming from most sources (family, social, or traditional media) and even less issued by the government itself. It also suggests that the theoretical model has to be modified, to be able to cope with individuals with different learning styles: some who learn solely through direct observation, like walking the streets of the city, while others benefit from prompted information such as the vignettes in our treatments.

Joint effect of previous knowledge and perceived government quality

Those who respond the most to the informational treatment are the ones that can update their beliefs, either because they already have basic knowledge about the commitments, but the treatment provides them with relevant data or because they are not at the extremes of the initial perception of the government and are therefore willing to incorporate new information into their beliefs. Notice that the perceived quality of the government and previous knowledge of the policy go along. Figure A10 depicts the distribution of the perceived quality scores by each level of knowledge. We observe that the distribution of people who previously knew about the commitments is skewed right, while those who did not know about the commitments

are marginally shifted to the left.[27] Thus, the interaction of initial beliefs with the level of knowledge plays an important role in how information affects trust perceptions.

When one looks at the treatment effects of the experiment on trust, conditional on the priors of perceived quality of government, the only distribution that shifts to the right is that of the intermediate group that had only heard of the policy. This confirms that this is the only group receptive to the treatment, and that the key variable is type of prior knowledge, rather than priors on quality of government, that at most are important for individuals at the polar extremes.

In this same line, Table A6 presents a triple interaction identification strategy that incorporates our observation about the relationship between prior knowledge and perceived quality of government. It presents a version in which the knowledge and quality variables are binary, but it also depicts all the possible combinations of the binary and categorical variables in the triple interaction. In our preferred specification (the most disaggregated one), we find that the effect of the treatment on the trust index among the people who had only heard about the commitments hold, and as before, it counteracts the initial negative perception of the trustworthiness of the government among the control individuals who had also heard about the policy. We find no evidence that perceptions of government quality play an important role in the effect of the informational vignettes.

In summary, it is mainly those who can increase their information set, given that they already have a minimum of knowledge about the policy, who respond favorably to the provision of information through the vignettes, especially those who initially had a low perception of the government's ability to meet the needs of the population with quality services.

5 Discussion

Trust is a multidimensional state in which individuals rely on the integrity, ability, and surety of a person or institution. Providing information about the mayor's commitments at the beginning of his mandate and their fulfillment have been shown to induce important changes in

[27] The two-sample Kolmogorov-Smirnov test for equality of distributions reject the null hypothesis of no difference between distributions. Instead, it indicates that perceptions of the quality of the government are shifted left for people who did not know the commitments compared to those who knew about it and compared to those who have heard about it. Similarly, there is a difference between the distribution for those who had heard about the policy and the ones who knew it for sure.

citizens' perceived trustworthiness of the city government as an institution. However, trust is also grounded in the belief that state members can deliver what they promise and commit to it. Citizens rely upon government officials to carry out the investments they promised while in the campaign. Hence, individuals expect public servants and politicians to care about the population and act accordingly.

Following Keefer et al. (2018) and Keefer et al. (2020), we consider two important trust components: whether others can keep their promises and whether they care about people like the respondent to assess how much do people trust those who develop public policy and commit to achieving the city's goals when shown with relevant information about their management. Citizens can update their beliefs about government members' trustworthiness and benevolence when presented with relevant information on what they have done for the city.

We do not find that information on commitments marginally affects politicians' and public servants' perceived trustworthiness. However, we observe that people are 5 pp more likely to express that public servants care about people like the respondent when they receive general information, regardless of whether it is supplemented by accomplishments at the city or commune level (see Table 6). The management model developed in recent years in CABA is based on transparency, both internal and external. Although this study focuses mainly on how information affects citizens' trust in government (external transparency), this result shows us that internal management perceptions are also changing. In particular, we observe that individuals consider public servants to be more empathetic with society.

We have identified that respondents to the experiment make a sharp distinction between "politicians in general" and the "city government." This is reflected in the high initial perception of *government* quality, and the low perception of the politicians ability to keep their promises in the control group (21.5%). Treatments have less impact on the more personalized evaluation of politicians and public employees than on institutional evaluation of the city government. The non-significance of the effect on politicians may be explained in part by relatively low statistical

power.[28,29]

6 Conclusions

Since trust and initiatives to promote transparency are endogenous variables, it is difficult to disentangle their causal relationships. To study this, we developed a theoretical model and conducted an empirical experiment in Buenos Aires City to evaluate how a transparency initiative influenced and was affected by trust in the government and its agents (politicians and public servants). The transparency initiative entailed not only revealing the achievements of the local administration but also the plans that had been initially made. These promises are what the city government calls "commitments".

The model suggests that the effect of transparency initiatives is critically dependent on priors, i.e., baseline trust levels. People at the extremes of no trust and complete trust will not be affected by the initiatives; only those in the middle will be impacted. Furthermore, the model predicts that the treatment will mainly affect those unfamiliar with the transparency policy.

The experimental results corroborate both predictions, showing that the impact of the policy is greatest for intermediate levels of trust, and that there is no effect whatsoever on people who are already familiar with the project (our interpretation is that they are not receiving new information).

However, the experimental results go beyond the models' predictions. When we control for prior knowledge, the data show a stark difference between people who did not know at all about the commitments and those who had heard something about them, so they seemed vaguely familiar. All the effect of the treatment is through this group that had heard of the treatment.

[28] Following McKenzie and Ozier (2019) recommendations, we conduct ex-post Minimal Detectable Effect (MDE) calculations. As we use the estimated standard error to calculate ex-post MDE, it may present some variation from sample to sample. However, this imprecision is lower than the one obtained with ex-post power. Power is set to 80%, the significance level to 5%, and we correct for baseline correlation given that randomization conducted over blocks of age and gender. We conduct pairwise power calculations considering the binary nature of dependent variables, as in Hemming and Marsh (2013), and do not compare two different informational treatments in any case. MDEs are significantly higher than the coefficients we observe in our regressions, which suggest that a greater sample size would have avoided both type I and II errors. An effect lower than 8 pp would not be detected on average.

[29] Our experimental design has three treatment arms, therefore, we also calculate Cohen's δ, which defines the effect size for a one-way analysis of variance, as the square root of the contrast variance to the error or within-group variance. According to Cohen's rule of thumb, we find a small treatment effect. Cohen (2013) indicates that an effect size of 0.2 is small, 0.5, medium, and 0.8 large. This means that if two groups' means do not differ by 0.2 standard deviations or more, the difference is trivial, even if it is statistically significant.

While the group "Heard of" had higher initial trust levels than the group "Do not know", this is not enough to explain the difference, because even individuals with higher initial trust do not react to the treatment. This might have to do with different learning styles; some people only trust what they observe directly, not what others say. The experiment thus suggests that post-electoral commitments may be useful as a complementary tool to solid performance and efficient administrative management. Transparency reforms on their own may be less effective in changing citizens' perceptions than performance, as there is a group of people, those unaware of the commitments, that is not affected by words, only by deeds.

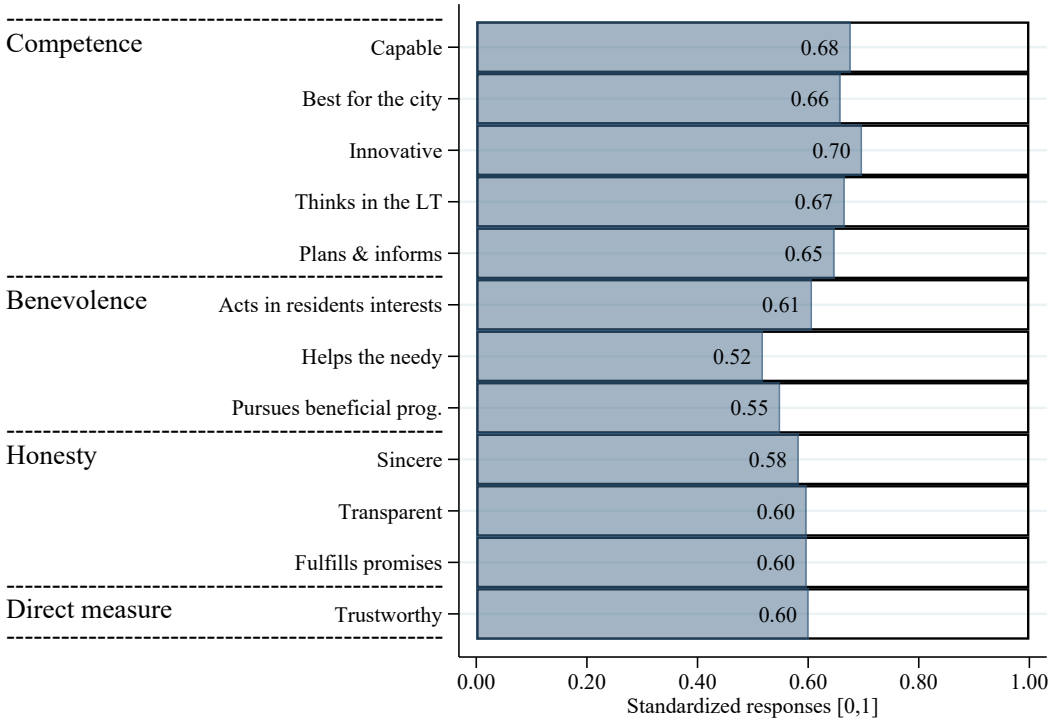
7 Tables and Graphs

Table 2: Summary statistics and randomization balance

	Control	Diff wrt. Control			p-value Wald test equality coefficients				Sample
Variable	(av. & s.d.)	T1	T2	T3	T1=T2=T3	T1=T2	T1=T3	T2=T3	Size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Gender	0.570	0.042	-0.011	0.005	0.197	0.084	0.168	0.538	2,375
	(0.021)	(0.036)	(0.024)	(0.033)					
Age	47.656	-1.829	-2.155**	-2.146*	0.942	0.740	0.780	0.992	2,375
	(1.399)	(1.165)	(0.956)	(1.099)					
College	0.404	0.007	0.015	-0.011	0.615	0.747	0.426	0.366	2,375
	(0.042)	(0.032)	(0.037)	(0.034)					
High school	0.846	-0.022*	0.025	-0.001	0.047	0.041	0.250	0.395	2,375
	(0.016)	(0.011)	(0.021)	(0.017)					
Employed	0.522	0.039	0.017	0.009	0.630	0.449	0.366	0.765	2,375
	(0.021)	(0.030)	(0.026)	(0.031)					
Unemployed	0.152	-0.018	0.010	-0.008	0.307	0.133	0.557	0.413	2,375
	(0.018)	(0.022)	(0.024)	(0.028)					
Socio-economic level (high)	0.244	-0.025	-0.040**	-0.038	0.656	0.463	0.557	0.938	2,375
	(0.025)	(0.020)	(0.017)	(0.027)					
Voluntary Health Insurance	0.430	-0.007	-0.013	0.011	0.768	0.843	0.584	0.483	2,375
	(0.026)	(0.028)	(0.021)	(0.015)					
Internet at home	0.526	0.000	0.000	-0.013	0.908	0.986	0.666	0.734	2,375
	(0.023)	(0.025)	(0.022)	(0.021)					
Credit Card	0.347	0.016	-0.010	0.006	0.697	0.411	0.687	0.520	2,375
	(0.027)	(0.019)	(0.025)	(0.017)					
One or more cars	0.175	0.005	-0.015	0.001	0.625	0.354	0.877	0.483	2,375
	(0.018)	(0.021)	(0.018)	(0.025)					
Perc. Quality of Governm.	7.200	-0.142	-0.055	0.197	0.088	0.616	0.068	0.064	2,331
	(0.103)	(0.154)	(0.098)	(0.126)					
Knowledge of 'Compromisos'	0.219	-0.001	0.047	0.005	0.289	0.123	0.728	0.205	2,375
	(0.016)	(0.015)	(0.030)	(0.020)					
Trust Others	0.662	0.003	-0.026	-0.022	0.589	0.333	0.396	0.867	2,265
	(0.021)	(0.028)	(0.027)	(0.020)					
Collective Action	0.671	0.010	-0.009	0.012	0.603	0.444	0.936	0.344	2,261
	(0.020)	(0.030)	(0.029)	(0.032)					

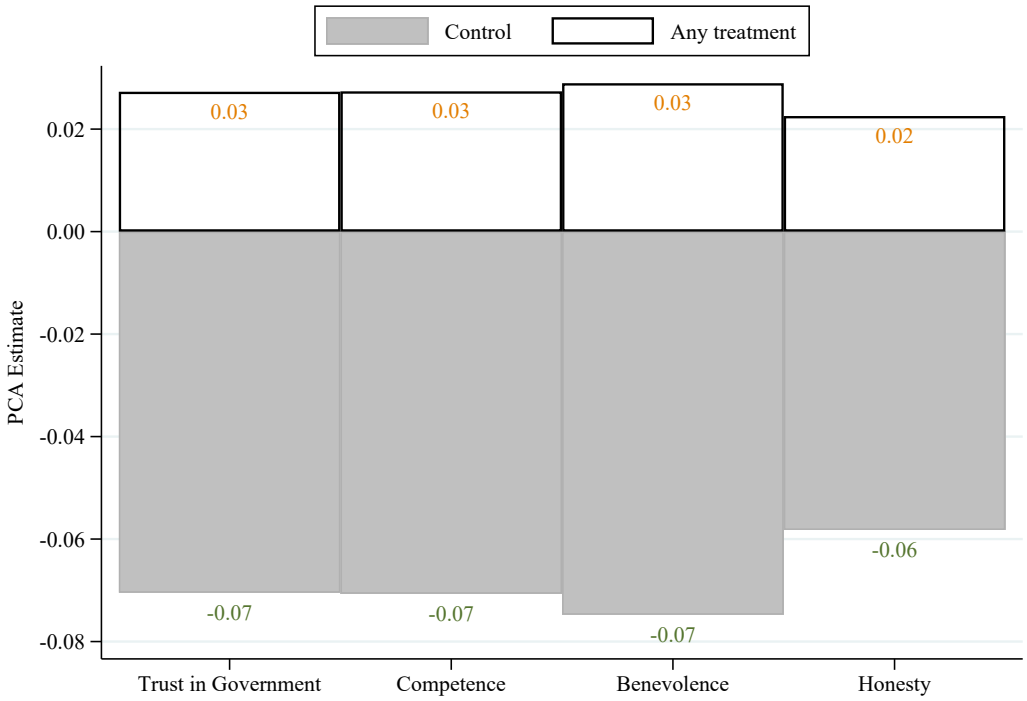
Notes: Column (1) shows the sample average and the standard deviation in parentheses for the control group. Columns (2)-(4) show the regression coefficient and the standard error in parentheses corresponding to an OLS regression - observable is the dependent variable and the treatment variables are the independent ones (T1-T3). Columns (5)-(8) show the p-value of a Wald test of equality of coefficients. Finally, column (9) shows the sample size. *Gender* is a indicator variable for women. *Age* is a continuous variable from 18 to 100 years old. *College* takes the value of one when the individual has a college degree at least, and *High school* is read similarly. *Employed* and *Unemployed* are binary variables for those who have full-time employment (or work from/at home) and those who are looking for a job at the time of the survey, respectively. *Socio-economic level (High)* is a binary variable for those with the highest category in socio-economic level. *Perceived Quality of the Government* is self-explanatory and takes values between 1 and 10, in which the lowest value reflects a very bad score while the greatest an excellent score. *Knowledge* is a binary variable and takes the value one if the participant knows the 'Compromisos' policy and zero otherwise. *Trust Others* is a binary variable that takes the value of one when participants indicate that others are reliable or very reliable. *Collective Action* is a dummy variable that indicates whether participants indicate that they would be able to collect 500 signatures to support a petition for the government among their neighbors. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Components of the index - PCA



Notes: The figure presents categorical variables that assess citizens' perception of the Government, standardized between zero and one. They account for each of the characteristics asked in the question: *Using a scale from 1 to 7, where one is "Completely disagree," and seven is "Completely agree," please show your level of agreement with the following statements about the Government of the city of Buenos Aires.* The interpretation of each bar goes as follows, e.g.: *68%* of the surveyed individuals consider that the CABA Government *is capable.*

Figure 2: Information and trust perception - PCA



Notes: Dependent variables depicted in this figure are constructed using a PCA method, and standardized with mean zero and standard deviation one. They account for indices following Grimmelikhuijsen (2012). Bars are constructed both for the control and treated groups. For illustration purposes, we present the average result for all treatment arms as one. Values in orange correspond to the average of each index for the treated units, and green for their counterpart, the control group

Table 3: Treatment Effect on Trust in Government (by dimension)

VARIABLES	Trust in the Government Global Index			Dimensions of Trust			Direct measure Trustworthiness
				Competence	Benevolence	Honesty	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T1: Commitments	0.051	0.095***	0.095***	0.083**	0.097**	0.099***	0.127***
	(0.053)	(0.028)	(0.027)	(0.028)	(0.037)	(0.030)	(0.028)
T2: Commitments + Fulf. city	0.116**	0.125***	0.124***	0.125***	0.119**	0.113***	0.114***
	(0.046)	(0.029)	(0.030)	(0.027)	(0.042)	(0.032)	(0.033)
T3: Commitments + Fulf. comuna	0.129**	0.104***	0.104***	0.107***	0.119***	0.073**	0.081*
	(0.056)	(0.033)	(0.034)	(0.035)	(0.036)	(0.034)	(0.039)
Constant	-0.070*	-2.233***	-2.206***	-2.187***	-1.922***	-2.242***	-2.191***
	(0.038)	(0.120)	(0.113)	(0.106)	(0.134)	(0.114)	(0.146)
Observations	2,375	2,278	2,278	2,278	2,278	2,278	2,278
R-squared	0.003	0.663	0.665	0.639	0.594	0.625	0.606
Joint significance (p-value)	0.483	0.668	0.679	0.524	0.849	0.573	0.566
T1=T2	0.268	0.422	0.432	0.279	0.584	0.725	0.755
T1=T3	0.309	0.844	0.841	0.638	0.635	0.514	0.330
T2=T3	0.836	0.573	0.582	0.656	0.995	0.307	0.384
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Comuna FE	No	No	Yes	Yes	Yes	Yes	Yes

Notes: All dependent variables are constructed using a PCA method, and standardized with mean zero and standard deviation one. The Competence dimension considers the assessments of the following characteristics, the government: is capable, does what is best for the city, is innovative, thinks in the long run, and plans and informs its plans; the Benevolence dimension considers the following: acts in the interests of its residents, helps those in need and pursues policies and projects beneficial for the families. Finally, the Honesty dimension takes into account: is sincere, is transparent, fulfills its promises. We also study the effect of information on a direct measure of trust in government that asks the respondents to indicate the degree in which she agrees that the city government is trustworthy. Control variables include: age, gender, socio-economic level, labor status, public policy preferences (revealed preferences for public budget allocation between education and infrastructure), being first exposed to information on , pre-treatment beliefs on government quality and the collective action dummy variable. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: Treatment Effect on Trust in Government (by component)

	Competence					Benevolence			Honesty		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
T1: Commitments	0.069*	0.053*	0.031	0.120***	0.102**	0.047	0.119***	0.106**	0.104***	0.063*	0.116***
	(0.035)	(0.027)	(0.035)	(0.038)	(0.044)	(0.038)	(0.037)	(0.038)	(0.031)	(0.035)	(0.024)
T2: Commitments + Fulf. city	0.097***	0.095**	0.052	0.142***	0.177***	0.079**	0.148***	0.106**	0.084*	0.072**	0.169***
	(0.030)	(0.033)	(0.038)	(0.031)	(0.038)	(0.035)	(0.046)	(0.047)	(0.041)	(0.030)	(0.030)
T3: Commitments + Fulf. comuna	0.107**	0.073*	0.053	0.134***	0.116***	0.056	0.123***	0.154***	0.063	0.033	0.113***
	(0.038)	(0.040)	(0.055)	(0.038)	(0.038)	(0.040)	(0.040)	(0.039)	(0.039)	(0.034)	(0.031)
Constant	-2.103***	-1.929***	-2.060***	-2.057***	-1.658***	-1.838***	-1.787***	-1.738***	-2.001***	-2.133***	-2.285***
	(0.106)	(0.133)	(0.094)	(0.164)	(0.117)	(0.118)	(0.151)	(0.134)	(0.119)	(0.134)	(0.111)
Observations	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278
R-squared	0.556	0.596	0.509	0.501	0.441	0.560	0.493	0.499	0.566	0.565	0.581
Joint significance (p-value)	0.738	0.462	0.892	0.800	0.0401	0.659	0.707	0.443	0.667	0.537	0.415
T1=T2	0.543	0.228	0.641	0.549	0.0458	0.420	0.556	0.983	0.649	0.831	0.225
T1=T3	0.537	0.673	0.749	0.803	0.808	0.870	0.932	0.314	0.379	0.444	0.924
T2=T3	0.874	0.595	0.988	0.824	0.144	0.613	0.456	0.238	0.598	0.292	0.240

Notes: All regressions include controls and commune fixed effects. All dependent variables are standardized with mean 0 and standard deviation 1. Each column presents the result for perceived performance of the CABA government. The first column displays the global effect on the Index of Trust in the Government. Following Grimmelikhuijsen (2012), next five columns reflect Government Competence: (1) is capable, (2) does what is best for the city, (3) is innovative, (4) thinks in the long-term, and (5) plans and informs; following three columns show Benevolence: (6) acts in the interests of the residents, (7) helps those in need, (8) pursues policies and projects that are beneficial for families; next three, Honesty: (9) is sincere, (10) is transparent, (11) fulfills its promises. Control variables include: age, gender, socio-economic level, labor status, public policy preferences (revealed preferences for public budget allocation between education and infrastructure), first exposure to 'Compromisos' and pre-treatment beliefs on government quality. Robust standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5: Average Marginal Treatment Effect on Trust in Government - Generalized Ordered Logit

Panel A: Competence

	1	2	3	4	5	6	7
The CABA Government...	**Strongly disagree**	**Disagree**	**Somewhat disagree**	**Neutral**	**Somewhat agree**	**Agree**	**Strongly agree**
Competence 1: is capable							
T1: Commitments	-0.009**	-0.004*	-0.005*	-0.007**	-0.007**	-0.003*	0.035**
	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.017)
T2: Commitments + Fulf. city	-0.011**	-0.005**	-0.006**	-0.008**	-0.009**	-0.004**	0.042**
	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.017)
T3: Commitments + Fulf. comuna	-0.013***	-0.006***	-0.008***	-0.010***	-0.011***	-0.005**	0.054***
	(0.004)	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)	(0.017)
Competence 2: does what is best for the city							
T1: Commitments	-0.008*	-0.003*	-0.005*	-0.005*	-0.004*	0.000	0.025*
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.014)
T2: Commitments + Fulf. city	-0.012***	-0.005**	-0.008**	-0.008***	-0.007**	0.000	0.040***
	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.015)
T3: Commitments + Fulf. comuna	-0.010**	-0.004**	-0.006**	-0.007**	-0.006**	0.000	0.032**
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.014)
Competence 3: is innovative							
T1: Commitments	-0.007	-0.002	-0.003	-0.005	-0.004	-0.001	0.021
	(0.005)	(0.001)	(0.003)	(0.004)	(0.003)	(0.001)	(0.017)
T2: Commitments + Fulf. city	-0.010*	-0.003*	-0.005*	-0.007*	-0.006*	-0.001	0.033*
	(0.006)	(0.002)	(0.003)	(0.004)	(0.003)	(0.001)	(0.018)
T3: Commitments + Fulf. comuna	-0.009	-0.002	-0.005	-0.006	-0.005	-0.001	0.028
	(0.005)	(0.001)	(0.003)	(0.004)	(0.003)	(0.001)	(0.017)
Competence 4: thinks in the long term							
T1: Commitments	-0.018***	-0.003**	-0.009***	-0.012***	-0.010***	-0.005**	0.056***
	(0.006)	(0.001)	(0.003)	(0.004)	(0.003)	(0.002)	(0.018)
T2: Commitments + Fulf. city	-0.022***	-0.004***	-0.011***	-0.014***	-0.012***	-0.006***	0.068***
	(0.006)	(0.001)	(0.003)	(0.004)	(0.004)	(0.002)	(0.018)
T3: Commitments + Fulf. comuna	-0.016	0.013	-0.041***	-0.013	-0.033**	0.027	0.063***
	(0.012)	(0.009)	(0.010)	(0.014)	(0.016)	(0.019)	(0.022)
Competence 5: plans and informs its plans							
T1: Commitments	0.006	-0.013	-0.044***	-0.001	0.002	0.021	0.028
	(0.010)	(0.009)	(0.012)	(0.015)	(0.018)	(0.018)	(0.021)
T2: Commitments + Fulf. city	-0.021***	-0.008***	-0.016***	-0.015***	-0.012***	0.003*	0.069***
	(0.005)	(0.002)	(0.004)	(0.004)	(0.003)	(0.002)	(0.017)
T3: Commitments + Fulf. comuna	-0.015***	-0.006***	-0.012***	-0.011***	-0.008**	0.002	0.049***
	(0.006)	(0.002)	(0.004)	(0.004)	(0.003)	(0.002)	(0.018)

Panel B: Benevolence

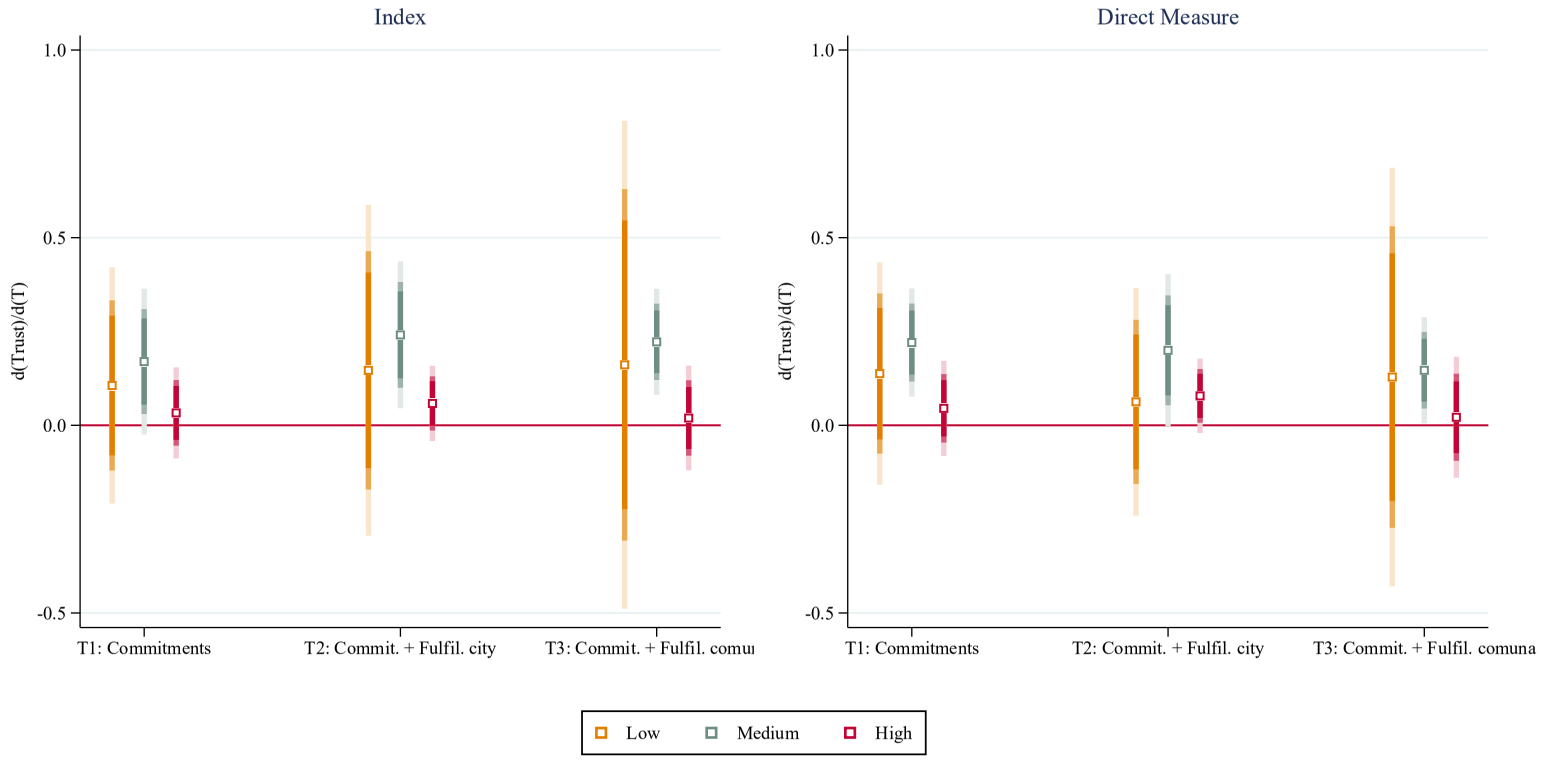
	1	2	3	4	5	6	7
The CABA Government...	**Strongly disagree**	**Disagree**	**Somewhat disagree**	**Neutral**	**Somewhat agree**	**Agree**	**Strongly agree**
Benevolence 1: acts in the interests of neighbors							
T1: Promises	-0.004	-0.001	-0.002	-0.002	-0.001	0.001	0.009
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.013)
T2: Commitments + Fulf. city	-0.010*	-0.004*	-0.005*	-0.006*	-0.003*	0.004*	0.025*
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.014)
T3: Commitments + Fulf. comuna	-0.008	-0.003	-0.004	-0.005	-0.003	0.003	0.020
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.013)
Benevolence 2: does everything in its power to help those in need							
T1: Promises	-0.022***	-0.007***	-0.008***	-0.007***	0.000	0.010***	0.033***
	(0.008)	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.012)
T2: Commitments + Fulf. city	-0.028***	-0.009***	-0.010***	-0.009***	0.000	0.013***	0.043***
	(0.008)	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.013)
T3: Commitments + Fulf. comuna	-0.025***	-0.008***	-0.009***	-0.008***	0.000	0.012***	0.038***
	(0.008)	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.012)
Benevolence 3: pursues policies and projects that my family cares about							
T1: Promises	-0.022***	-0.009***	-0.006***	-0.010***	-0.002	0.009***	0.039***
	(0.007)	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.013)
T2: Commitments + Fulf. city	-0.017**	-0.007**	-0.005**	-0.008**	-0.001	0.007**	0.031**
	(0.007)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.013)
T3: Commitments + Fulf. comuna	-0.030***	-0.013***	-0.009***	-0.014***	-0.003	0.013***	0.055***
	(0.007)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.013)

Panel C: Honesty

	1	2	3	4	5	6	7
The CABA Government...	**Strongly disagree**	**Disagree**	**Somewhat disagree**	**Neutral**	**Somewhat agree**	**Agree**	**Strongly agree**
Honesty 1: is sincere							
T1: Commitments	-0.019***	-0.007***	-0.007***	-0.007***	-0.004**	0.003**	0.039***
	(0.007)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.014)
T2: Commitments + Fulf. city	-0.013*	-0.005*	-0.005*	-0.005*	-0.003*	0.002*	0.028*
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.014)
T3: Commitments + Fulf. comuna	-0.012*	-0.004*	-0.005*	-0.005*	-0.002*	0.002*	0.026*
	(0.007)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.014)
Honesty 2: is transparent							
T1: Commitments	-0.013**	-0.004**	-0.005**	-0.005**	-0.003*	0.002*	0.028**
	(0.007)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.014)
T2: Commitments + Fulf. city	-0.015**	-0.005**	-0.006**	-0.005**	-0.003**	0.002*	0.032**
	(0.007)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)	(0.015)
T3: Commitments + Fulf. comuna	-0.008	-0.003	-0.003	-0.003	-0.002	0.001	0.018
	(0.007)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.015)
Honesty 3: fulfills its promises							
T1: Commitments	-0.019***	-0.006***	-0.010***	-0.010***	-0.007***	0.006***	0.046***
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.013)
T2: Commitments + Fulf. city	-0.027***	-0.008***	-0.014***	-0.014***	-0.010***	0.009***	0.065***
	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.014)
T3: Commitments + Fulf. comuna	-0.017***	-0.005***	-0.009***	-0.009***	-0.006***	0.006***	0.042***
	(0.005)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.013)
Observations	2,278	2,278	2,278	2,278	2,278	2,278	2,278

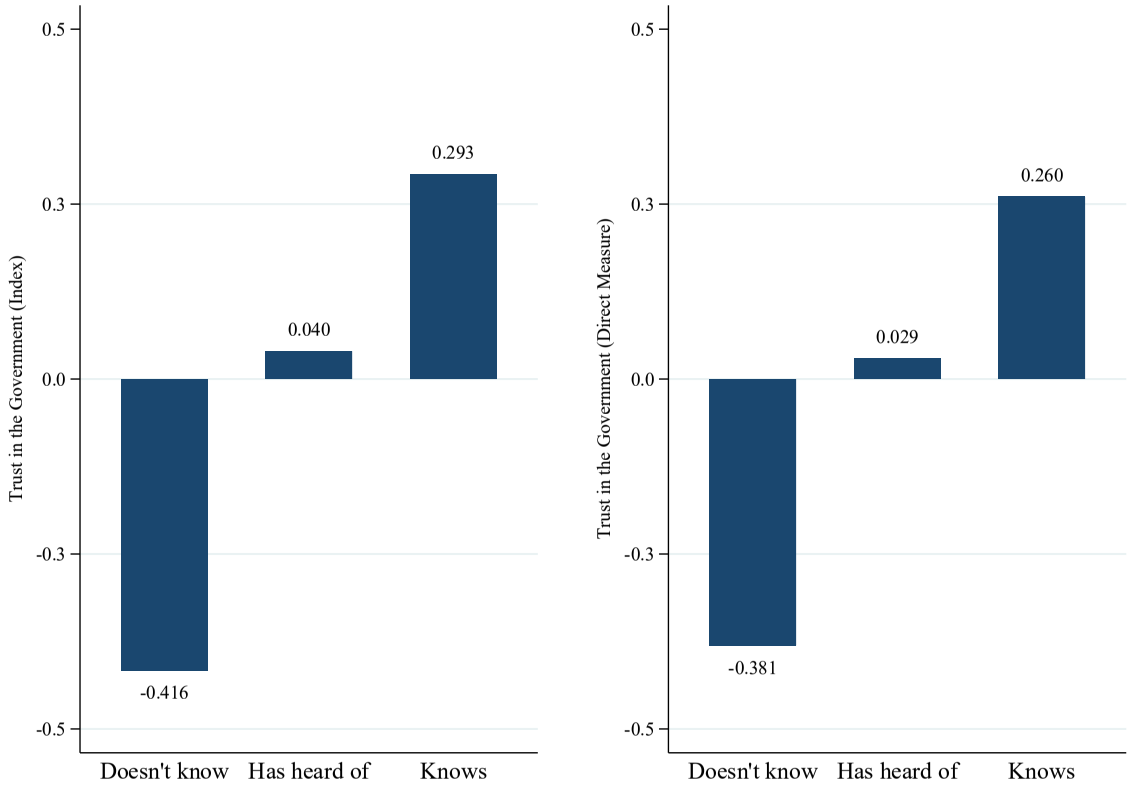
Notes: Robust standard errors are shown in parenthesis. Control variables include: age, gender, socio-economic level, labor status, pre-treatment beliefs on government quality and pre-intervention preferences for public education and infrastructure. *p<0.10, **p<0.05, ***p<0.01.

Figure 3: Treatment Effect on Trust in the Government by perceived quality of the government



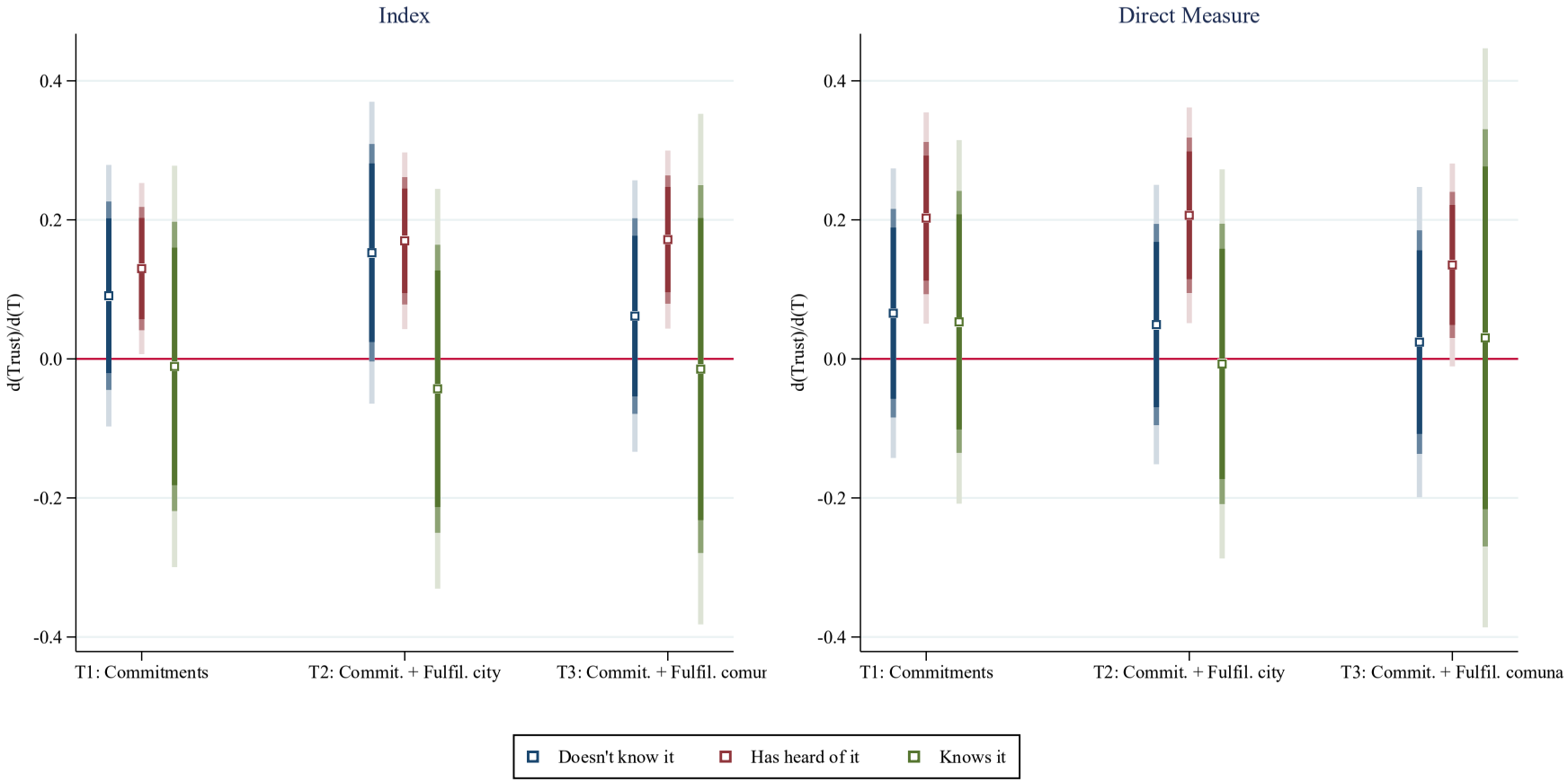
Notes: Perceived quality of the government is a categorical variable that takes values 1 to 10, with one meaning very low quality and 10 very high quality. We reclassified the participants assessments in three categories, low, med-high and high, given the sample composition. People who scored the government quality below 4 are part of the "Low" category, people who scored 4 to 7 in the "Mid" and people who scored 8 or more in the "High" category. The color intensity of confidence intervals represents the confidence level, from darker to lighter, 90%, 95%, and 99%.

Figure 4: Trust in the Government in the control group by previous knowledge of the policy



Notes: Dependent variables depicted in this figure are constructed using a PCA method, and standardized with mean zero and standard deviation one. The bars depict the level of trust in the government individuals from the control group have grouped by level of initial knowledge of the commitments policy.

Figure 5: Treatment Effect on Trust in the Government by level of knowledge of 'Compromisos'



Notes: Dependent variables depicted in this figure are constructed using a PCA method, and a standardization with mean zero and standard deviation one. The estimate points correspond to the effect of each treatment in the level of knowledge indicated in the legend. For example, people who had heard of the commitments but did not know them for sure before the survey increase their trust in the government by a greater extent than people who did not know them at all or those who knew the policy, regardless of the trust measure used. The color intensity of confidence intervals represents the confidence level, from darker to lighter, 90%, 95%, and 99%.

Table 6: Treatment Effect on Trust in Institutional Agents

Dimension of Trust	**Keep their promises**		**Care about people like you**	
Agent	Politicians in General	Public Servants from the City	Politicians in General	Public Servants from the City
	(1)	(2)	(3)	(4)
T1: Commitments	0.020	-0.003	0.036*	0.045***
	(0.021)	(0.025)	(0.021)	(0.017)
T2: Commitments + Fulf. city	0.023	0.027	0.004	0.039
	(0.021)	(0.028)	(0.019)	(0.027)
T3: Commitments + Fulf. comuna	-0.001	0.007	-0.007	0.049**
	(0.023)	(0.035)	(0.022)	(0.023)
Constant	0.228***	0.531***	0.266***	0.513***
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	2207	2221	2192	2208
Joint significance	0.601	0.524	0.030	0.957
T1=T2	0.830	0.280	0.024	0.855
T1=T3	0.397	0.784	0.041	0.808
T2=T3	0.318	0.493	0.614	0.779

Notes: All regressions include controls and commune fixed effects. All dependent variables are binary and take the value of one when the individual scored 'Very common' or 'Somewhat common' to the question *Do you think it is very common, somewhat common, unusual, or not at all common that* ***the agent*** KEEP THEIR PROMISES/CARE ABOUT THE INTERESTS OF PEOPLE LIKE YOU? Control variables include: age, gender, socio-economic level, labor status, public policy preferences (revealed preferences for public budget allocation between education and infrastructure), first exposure to , pre-treatment beliefs on government quality and the collective action dummy variable. Robust standard errors are shown in parentheses ***p<0.01, **p<0.05, *p<0.1

Figure 6: Treatment Effect on Trust in Government members - Heterogeneous effects

Notes: The color intensity of confidence intervals represents the confidence level, from darker to lighter, 90%, 95%, and 99%.

References

Alessandro, M., Cardinale Lagomarsino, B., Scartascini, C., Streb, J., and Torrealday, J. (2021). Transparency and trust in government. evidence from a survey experiment. *World Development*, 138:105223.

Anderson, J. E. (2017). Trust in government and willingness to pay taxes in transition countries. *Comparative Economic Studies*, 59:1–22.

Ansolabehere, S. and Snyder, J. M. J. (2000). Valence politics and equilibrium in spatial election models. *Public Choice*, 103:327–336.

Ardanaz, M., Otálvaro-Ramírez, S., and Scartascini, C. (2022). Does citizen participation in budget allocation pay? a survey experiment on political trust and participatory governance. Technical report, Inter-American Development Bank.

Aruguete, N., Calvo, E., and Ventura, T. (2021). News by popular demand: Ideology, reputation, and issue attention in social media news sharing. *International Journal of Press/Politics*.

Ashworth, S. (2012). Electoral accountability: recent theoretical and empirical work. *Annual Review of Political Science*, 15:183–201.

Baekgaard, M. (2015). Performance information and citizen service attitudes: Do cost information and service use affect the relationship? *International Public Management Journal*, 18(2):228–245.

Barrows, S., Henderson, M., Peterson, P. E., and West, M. R. (2016). Relative performance information and perceptions of public service quality: Evidence from american school districts. *Journal of Public Administration Research and Theory*, 26(3):571–583.

Blanco, L. and Ruiz, I. (2013). The impact of crime and insecurity on trust in democracy and institutions. *American Economic Review*, 103(3):284–88.

Bolsen, T., Druckman, J., and Cook, F. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behavior*, 36(2):235–262.

Bouwman, R. and Grimmelikhuijsen, S. (2016). Experimental public administration from 1992 to 2014: A systematic literature review and ways forward. *International Journal of Public Sector Management*, 29(2):110–131.

Butler, D. M., Hughes, A. C., Volden, C., and Wiseman, A. E. (2021). Do constituents know (or care) about the lawmaking effectiveness of their representatives? *Political Science Research and Methods*, pages 1–10.

Cai, M., Liu, P., and Wang, H. (2020). Political trust, risk preferences, and policy support: A study of land-dispossessed villagers in china. *World Development*, 125:104687.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

Evans, D. K., Holtemeyer, B., and Kosec, K. (2019). Cash transfers increase trust in local government. *World Development*, 114:138 – 155.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2):703–745.

Grimmelikhuijsen, S. (2012). Linking transparency, knowledge and citizen trust in government: an experiment. *International Review of Administrative Sciences*, 78(1):50–73.

Hemming, K. and Marsh, J. (2013). A menu-driven facility for sample-size calculations in cluster randomized controlled trials. *The Stata Journal*, 13(1):114–135.

Hertz, U., Bell, V., and Raihani, N. (2021). Trusting and learning from others: immediate and long-term effects of learning from observation and advice. *Proceedings Royal Society B*, 288.

Izquierdo, A., Pessino, C., and Vuletin, G. (2018). *Better Spending for Better Lives*. Inter-American Development Bank.

Jones, D., Molitor, D., and Reif, J. (2019). What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study*. *The Quarterly Journal of Economics*, 134(4):1747–1791.

Keefer, P. and Scartascini, C. (2022). *Trust: The Key to Social Cohesion and Growth in Latin America and the Caribbean*. Inter-American Development Bank.

Keefer, P., Scartascini, C., and Vlaicu, R. (2018). Shortchanging the Future: The Short-Term Bias of Politics. In Izquierdo, A. and Pessino, C., editors, *Better Spending for Better Lives*, chapter 10. Inter-American Development Bank, Washington, DC, USA, 1 edition.

Keefer, P., Scartascini, C., and Vlaicu, R. (2020). Voter Preferences, Electoral Promises and the Composition of Public Spending.

Keefer, P., Scartascini, C., and Vlaicu, R. (2022). Demand-side determinants of public spending allocations: Voter trust, risk and time preferences. *Journal of Public Economics*, 206:104579.

Kendall, C., Nannicini, T., and Trebbi, F. (2015). How do voters respond to information? evidence from a randomized campaign. *American Economic Review*, 105(1):322–53.

Khemani, S., Dal Bo, E., Ferraz, C., Finan, F., Stephenson, C., Odugbemi, A. M., Thapa, D., and Abrahams, S. D. (2016). Making politics work for development: Harnessing transparency and citizen engagement. Technical report. World Bank Policy Research Report.

Lupu, N., Ramírez Bustamante, M. V., and Zechmeister, E. (2020). Social media disruption: Messaging mistrust in latin america. *Journal of Democracy*, 31(3):160–171.

Marvel, J. D. (2016). Unconscious bias in citizens' evaluations of public sector performance. *Journal of Public Administration Research and Theory*, 26(1):143–158.

McKenzie, D. and Ozier, O. (2019). Why ex-post power using estimated effect sizes is bad, but an ex-post mde is not. Development Impact - World Bank Blogs.

Miranda, J. J., Datta, S., and Zoratto, L. (2020). Saving Water with a Nudge (or Two): Evidence from Costa Rica on the Effectiveness and Limits of Low-Cost Behavioral Interventions on Water Use. *The World Bank Economic Review*, 34(2):444—463.

Piotrowski, S., Grimmelikhuijsen, S., and Deat, F. (2019). Numbers of narratives? how government message strategies affect citizens' attitudes. *Public Performance and Management Review*, 42(5):1005–1028.

Porumbescu, G. A., Grimes, M., Grimmelikhuijsen, S., et al. (2021). Capturing the social relevance of government transparency and accountability using a behavioral lens. *Journal of Behavioral Public Administration*, 4(1).

Rogoff, K. and Sibert, A. (1988). Elections and macroeconomic policy cycles. *The review of economic studies*, 55(1):1–16.

Sances, M. W. (2021). Presidential approval and the inherited economy. *American Journal of Political Science*.

Slothuus, R. and De Vreese, C. (2010). Political parties, motivated reasoning and issue framing effects. *Journal of Politics*, 72(3):630–645.

Yamamura, E. (2014). Trust in government and its effect on preferences for income redistribution and perceived tax burden. *Economics of Governance*, 15:71–100.