Artículo Científico

Motivating bureaucrats with behavioral insights when state capacity is weak:
Evidence from large-scale field experiments in Peru

Autores

Andrew Dustan

Juan Hernández

Stanislao Maldonado

COLOMBIA
CIENTÍFICA
Conocimiento Global para el Desarrollo

ALIANZAEFI
Economía Formal e Inclusiva

Regular article

Motivating bureaucrats with behavioral insights when state capacity is weak: Evidence from large-scale field experiments in Peru☆

Andrew Dustan [a], Juan Manuel Hernandez-Agramonte [b], Stanislao Maldonado [c,*]

[a] *Department of Economics, Vanderbilt University, United States of America*
[b] *Innovations for Poverty Action, United States of America*
[c] *Faculty of Economics, Universidad del Rosario, Colombia*

ARTICLE INFO

ABSTRACT

We study how text messages incorporating behavioral insights can be used as a tool to affect civil servant performance when state capacity is weak. By experimentally varying the content of a messaging campaign targeted to civil servants implementing a school maintenance program in Peru, we test the effectiveness of reminders and treatments making salient either monitoring, social norms, the possibility of public disclosure of noncompliance, or audit risk. All messaging treatments improve compliance by similar magnitudes, increasing the probability of submitting a key expense report by an average of 3.9 percentage points over a base of 74%. The inability of this large-scale experiment to detect differential impacts by treatment arm is consistent with timely reminders being the main driver of increased compliance. We explore generalizability across time and populations in two supplemental experiments, confirming the promise of such campaigns to improve civil servant performance when the state lacks enforcement capacity.

1. Introduction

State capacity depends on a well-functioning bureaucracy in which competent civil servants are incentivized to perform their duties (Bertrand et al., 2020; Best et al., 2017). While improving selection into civil service has received significant attention as a mechanism for improving capacity (Dal Bo et al., 2013; Ashraf et al., 2015; Leaver et al., 2021), much remains to be learned about inducing existing bureaucrats to perform. Evidence is particularly sparse regarding the impact of non-monetary incentives, which are often necessary because of the fiscal costs and political economy challenges inherent to implementing monetary incentives at scale.[1] Non-monetary incentives such as social recognition (Ashraf et al., 2014) and location postings (Khan et al., 2019) have shown the potential to affect bureaucrat behavior in

a flexible, low-cost way, yet few other policies have been rigorously evaluated. In particular, there has been limited use of insights from behavioral economics to design interventions that induce civil servant effort and compliance with public policies.[2]

This paper analyzes the impacts of a randomized, large-scale intervention that leveraged behavioral insights to increase policy compliance. We partnered with Peru's Ministry of Education (MINEDU) to design and execute a text messaging (SMS) campaign, aimed at reducing noncompliance among bureaucrats – typically school principals – in implementing a national school maintenance program. Bureaucrats in this program file a work plan, receive a monetary transfer, execute the plan, and then document their expenditures. Before the SMS campaign, noncompliance was high: 9% filed expense reports late or in an incomplete manner, while 11% did not file at all.[3] Furthermore, 15% of civil servants did not withdraw the designated monetary transfers, representing 10% of the total assigned budget.[4] The SMS campaign randomly assigned bureaucrats to receive a series of messages from MINEDU that targeted one of five specific behavioral mechanisms, discussed in more detail below, or to a pure control group. Treatment status was linked to administrative databases that tracked compliance with program milestones, with a particular focus on withdrawing transfers and filing the culminating expense report.

We find that these messages increase compliance on the targeted outcomes. Assignment to any treatment arm causes a 3.9 percentage point (p.p.) increase in the probability of filing the expense report compared to 74% in the control group, a 15% closing in the compliance gap (the distance between the current levels of compliance and full compliance). The messages also increase by 1.5 p.p. the probability of (nearly) complete withdrawal of the transferred amount compared to 89%, closing the compliance gap by 13%. While the absolute size of these impacts is modest, given the low cost and relative simplicity of implementing the intervention at scale, these results indicate that SMS campaigns are one promising tool for increasing compliance. The magnitudes of these effects are similar to other behaviorally-motivated interventions in Peru that did not specifically target civil servants or their particular activities: letters highlighting tax compliance rates decreased the property tax compliance gap by 20% in Del Carpio (2013), while reminder text messages decreased the rate at which savers failed to meet their savings commitments by about 7% (Karlan et al., 2016).

The scale of the intervention and use of administrative data for measurement result in precise impact estimates, allowing us to test for differential impacts from the treatment arms targeting distinct behavioral mechanisms, introduced briefly here. The most basic treatment arm was simply a *reminder* of necessary program activities and well-known deadlines, aimed at helping to overcome limited attention problems by increasing the salience of program requirements. The other treatment arms also increase salience, while attempting to achieve additional impact by targeting additional behavioral barriers and mechanisms. The *salience of monitoring* treatment told civil servants their level of compliance, increasing the salience of administrative oversight that was already known to be occurring. The dynamic *social norm* treatment emphasized that most bureaucrats were already complying, appealing to people's desire to conform to common behavior among peers while also likely serving as a reminder that noncompliance was visible to supervisors. The *soft shaming* treatment reminded bureaucrats

that their names would appear in a public list in the case of noncompliance, while the *salience of auditing* treatment reminded recipients that they would be visited to supervise their activities, increasing the salience of consequences based in the loss of social status and formal discipline, respectively. None of these treatments modify the information civil servants have access to or the structure of the program's monitoring system. At the beginning of the intervention, civil servants know they are being monitored on a regular basis, that they will eventually be visited, and that their names will be published in the case of noncompliance. Responses to these treatments may thus be interpreted as a consequence of the targeted behavioral dimensions.

Interestingly, the pattern of results suggests that the treatment arms were similarly effective in inducing compliance with the targeted outcomes. We cannot reject that all treatments have equal impacts on expense report filing, with point estimates ranging from 3.0 p.p. (reminder treatment) to 4.9 p.p. (salience of monitoring treatment). We are also unable to reject equal impacts on a variety of metrics for transfer withdrawal. For example, equality of impacts cannot be rejected for the outcome of withdrawing 99% of transferred funds, and these impacts range from 1.0 p.p. (social norm and salience of auditing treatments) to 2.1 p.p. (salience of monitoring treatment). There is, however, suggestive evidence that treatments had distinct impacts on the *quality* of expense reports, a dimension which was not targeted by the intervention. The social norms treatment increased the probability of having an expense report approved by the central government by 3.3 p.p., a point estimate greater than any other arm. The overall results are consistent with a conservative interpretation that all treatments acted as reminders, helping civil servants to overcome limited attention problems by making salient the need to comply with program policies, while leaving open the possibility that social norms may have induced differential impacts on implementation quality.

We enrich these primary findings with results from two additional SMS campaign experiments. A follow-up experiment implemented one year later, in 2016, explores the issue of learning and whether this SMS campaign can induce compliance over time. One concern is that civil servants become desensitized to such campaigns, making them ineffective beyond the first iteration. We find no evidence that treatment in 2015 reduces 2016 compliance or responsiveness to the 2016 treatment campaign, suggesting a lack of learning effects and lending additional support to the interpretation that the messages serve as timely reminders. The 2016 data also contain detailed information on civil servants' planned and reported expenditures, allowing us to assess whether the campaign affected the alignment between planned and executed activities.[5]

Finally, we report on a randomized SMS campaign targeting a different national program to learn whether the basic results can be replicated in a population of civil servants with different characteristics. In contrast to the MINEDU-employed bureaucrats characterized by rigid contracts and promotion practices, this experiment targeted field monitors employed by the Ministry of Social Inclusion (MIDIS). These employees have short-term contracts linked to less generous benefits packages than in the education sector.[6] Many field monitors fail to file mandated field reports, affecting planning for future service delivery. We test the role of social norms and the salience of monitoring, two promising interventions identified in the main experiment, in the full population of field monitors. The social norms treatment is ineffective,

---

[3] As a point of reference, Chaudhury et al. (2006) provide evidence on the poor performance of front-line public good providers. Absenteeism of teachers (19%) and health providers (35%) in developing countries are both high and in line with the rate of noncompliance observed in our context.

[4] Noncompliance in this setting hinders the government's ability to implement its policy preferences and service delivery. It also affects efficiency, since following up with noncompliers is costly. Lack of compliance is directly related to citizen-level outcomes — school infrastructure has a significant impact on educational progress, learning, and long-run outcomes (Akresh et al., 2018; Barret et al., 2019; Duflo, 2001; Hanushek, 1995).

[5] A subset of schools in the 2016 experiment are linked to a rolling school census that includes information on physical infrastructure, but results regarding impacts on infrastructure quality are imprecise.

[6] The existence of a sizable fraction of temporary workers in the public sector is a characteristic that has persisted in the developed and developing world, despite efforts to minimize patronage and political control (see Grindle (2012)). Meyer-Sahling et al. (2018) estimates that the fraction of temporary workers is about 23% in a group of African, Asian, Eastern European, and Latin American countries.

while the monitoring treatment improves compliance significantly. One explanation for this divergence from the main experiment is that field monitors may not expect to keep their posts far into the future, so they give little weight to their peers' perceptions in their compliance decisions. The salience of monitoring treatment is still effective in this setting because officials do care about the information that upstream bureaucrats have regarding their performance.

This paper relates to an emerging literature about the personnel economics of the state (Finan et al., 2017) and bureaucracies in developing countries (Pepinsky et al., 2017).[7] In particular, we contribute to a growing literature regarding strategies to increase civil servant performance. Existing studies have emphasized the role of monetary incentives in the case of enforcement agents (like tax collectors as in Khan et al. (2016)) and front-line service providers like teachers and health professionals, especially regarding incentives based on outcomes and inputs, with mixed results (Hasnain et al., 2014; Finan et al., 2017; Gilligan et al., 2022).[8] More specifically, we explore the role of non-financial incentives for civil servants. Banerjee et al. (2021) consider the role of transfers for police officers in India, Karachiwalla and Park (2017) study teacher promotion in China, and Khan et al. (2019) study postings for property tax inspectors in Pakistan, with positive results. More in line with the approach followed in this paper, others have found promise in non-financial rewards such as in-kind prizes (Glewwe et al., 2010) and social recognition (Ashraf et al., 2014).

A recent literature explores the role of behavioral biases in civil servants, systematically documenting the existence of present bias (Andreoni et al., 2016), status quo bias (Celhay et al., 2019), confirmation bias, framing effects, inattention bias and optimism bias (Hallsworth et al., 2018; Banuri et al., 2017) among bureaucrats. Rather than documenting the existence of such biases, this paper tests a set of strategies to deal with them. This paper is also related to a literature regarding the use of digital technology to improve policy outcomes. Researchers have used smartphones and other technological devices to monitor bureaucrats' attendance (Callen et al., 2018; Cilliers et al., 2018; Dhaliwal and Hanna, 2017; Duflo et al., 2012), increase accountability (Aker et al., 2017), improve public service delivery (Dal Bo et al., 2018) and minimize corruption (Muralidharan et al., 2016; Lewis-Faupel et al., 2016). This paper contributes an innovative way to use text messaging to induce compliance among civil servants.

Finally, this paper relates to a recent but growing literature about experimentation at scale (Muralidharan and Niehaus, 2017; Davis et al., 2017). Large-scale experiments are important for many reasons, including the search for external validity.[9] Among the strategies used to generalize the results from small-scale experiments to a relevant population are theoretical approaches (Banerjee et al., 2017b) and statistical designs based on reweighting (Andrews and Oster, 2017) or bounds (Kowalski, 2018). This paper's design addresses several concerns regarding external validity. By implementing the interventions with a population of civil servants, concerns about scalability and generalizability of results from small samples are not present. The subsequent follow-up intervention addresses the concern that its impacts

are not externally valid due to time-specific aggregate shocks (Rosenzweig and Udry, 2020). Furthermore, by implementing a similar SMS campaign with a different population of civil servants, we address the issue of how generalizable the results are for alternative settings and populations.

The rest of the paper is organized as follows. Section 2 describes the institutional setting. Section 3 introduces the intervention and research design. Section 4 describes the results of the main experiment. Section 5 presents the results of the subsequent two experiments. Section 6 concludes.

2. Institutional setting

The field experiment was run in collaboration with the National Educational Infrastructure Program (PRONIED) at the Ministry of Education (MINEDU) in Peru. The goal of this program is to expand, improve, replace, rehabilitate, and construct public educational infrastructure, including both buildings and furniture. One of its critical functions is the School Infrastructure and Furniture Maintenance Program, which entails the allocation of monetary transfers to a civil servant in charge of regular maintenance activities in each school. The progression of the intervention is summarized in Fig. 1.

The program has a strong participatory component. At the beginning of each school year, a Maintenance Committee is created with the participation of teachers, students and parents. This committee is the unit at the school level in charge of implementing all activities related to the maintenance of basic infrastructure and furnishings. In addition, an Oversight Committee is also formed along the same lines. A coordinator of the Maintenance Committee is selected, which is usually the school principal but occasionally a teacher. This coordinator is required to submit a signed legal document called a commitment act to the Educational Local Management Unit (UGEL, similar to a school district in the US case), declaring that they fully understand the program's policies and procedures and that they promise to execute and declare all expenditures according to stated deadlines.

Once this institutional structure is in place, a technical form is produced with the maintenance investments that are considered relevant for the Maintenance Committee. The technical form establishes not only the expenditure items but also the estimated budget. This form is submitted to the UGEL for approval. An UGEL is similar to a school district in the US case. Once the technical form is approved, maintenance civil servants are able to start carrying out maintenance activities and withdraw maintenance funds from accounts at the National Bank assigned to them exclusively for investments in infrastructure. Typical maintenance activities include refurbishing classrooms, such as repairing floors and walls or replacing windows; repairing and improving bathrooms; upgrading electrical installations; and repairing or replacing classroom furniture.

Once the execution phase is completed, the coordinator is required to write an expense report accompanied by invoices for all expenditures, submitted with a report prepared by the Oversight Committee, which evaluates whether they believe that funds were used for their intended purpose. Unused funds are returned. The final step is the approval of the expense report, an activity performed by an infrastructure specialist at the corresponding UGEL.

The maintenance and reporting activities carried out by the coordinator are in addition to the usual workload for which they are responsible. For school principals, these duties include all aspects of school administration, including day-to-day management of teachers and school activities and compliance with MINEDU norms. That is, maintenance tasks are undertaken alongside many other obligations, creating the potential for inattention to deadlines for implementation and reporting.

It is important to emphasize that the maintenance funds are assigned to a single maintenance civil servant in each school. While

---

[7] Recent reviews in the political economy literature include Dal Bo and Finan (2016) and Azulai et al. (2014).

[8] Barlevy and Neal (2012) present optimal financial incentives for educators, highlighting context-specific factors such as the possibility that standardized tests used to assess performance may change over time.

[9] Researchers have explored several dimensions of external validity, in particular the scalability of interventions, the existence of market equilibrium effects and externalities, site selection and piloting bias, the effect of treatments on different populations, the effect of treatment in the same population under different circumstances, and the effect of different, but related, technologies (Allcott, 2015; Banerjee et al., 2017a; Al-Ubaydli et al., 2017; Banerjee et al., 2017b). See Peters et al. (2018) for an overview of these issues in the development literature.