ABSTRACT: While external monitoring has been shown to effectively reduce missing government expenditures, its effects on public service delivery have not been credibly documented. This matters because theory is ambiguous about the effects of monitoring on downstream outcomes: since bureaucratic decisions involve multiple dimensions, deterring corruption might distort other margins, such as the quantity and quality of public goods. This paper exploits variation from an anti-corruption program in Brazil, designed by the federal government to enforce guidelines on earmarked transfers to municipalities, to study this question. Combining random audits with a differences-in-differences strategy, we find that the anti-corruption program greatly reduced occurrences of over-invoicing and off-the-record payments, and of procurement manipulation within health transfers. However, evidence from audited amounts suggests that lower corruption came at a high cost: after the program, public spending fell by so much that corruption per dollar spent may have actually increased. Health indicators, such as hospital beds and immunization coverage, became worse as a result. These findings are consistent with those responsible for procurement dramatically reducing purchases after the program, either because they no longer can capture rents, or because they are afraid of being punished for procurement mistakes.

This version: October 8th, 2017

Keywords: Corruption; Audits; Bureaucratic Effectiveness; Public Spending; Mismanagement; Health.

JEL codes: D72, D78, H41, I18, K42, O17

† We would like to thank the invaluable guidance of Sendhil Mullainathan, Nathan Nunn, Edward Glaeser, and Gautam Rao. We are also grateful for comments from George Avelino, Ciro Biderman, Paulo Costa, Rema Hanna, Horacio Larreguy, Joana Naritomi, Rohini Pande, Andrei Shleifer, Jim Snyder, and Laura Trucco. Special thanks to Rafael Barroso, Gustavo Fernandes and Kleber de Souza for their help in understanding the specifics of the municipal spending data. We also thank Flavio Riva and Guilherme Avelar for excellent research assistance. All remaining errors are ours.

* guilherme.lichand@econ.uzh.ch
† marcoslopes@gvmail.br
‡ mcm@econ.puc-rio.br
“Every dollar that a corrupt official or a corrupt business person puts in their pocket is a dollar stolen from a pregnant woman who needs health care. (...) In the developing world, corruption is public enemy number 1”.

– Jim Kim, president of the World Bank

“Even in the case of petty bribery or extortion, it is relevant to ask, what is the alternative?”

– Colin Leys (1965, p. 220)

1 Introduction

There is increasing understanding of what tools can most effectively fight corruption; in particular, external monitoring (from audits to informational campaigns) has been shown to significantly reduce missing government expenditures. Since corruption captures resources from public goods’ provision, as Jim Kim puts it in the opening quote, monitoring is expected to increase resources towards public services. However, rent capture is only one dimension of bureaucratic performance. Exposing corruption may reduce incentives for bureaucrats to exert effort in other dimensions, such as providing the optimal quantity and quality of public goods (Leff, 1964; Leys, 1965; Huntington, 1968; Banerjee, Hanna and Mullainathan, 2012). If that is the case, monitoring could actually hurt public service delivery, as the opening quote by Colin Leys hints at.

To study this question, one must focus on bureaucratic incentives, looking beyond corruption to analyze the effects of monitoring on the quantity and quality of public goods and services, and ultimately on downstream outcomes. This paper exploits variation from an anti-corruption program designed to enforce guidelines on earmarked transfers to local governments in Brazil, estimating its impacts on the incidence of corruption, on spending and implementation quality within health transfers, and on municipalities’ health indicators.

In Brazil, as in most developing countries, many local public goods are provided by subnational governments. Within health, a local bureaucracy employed by each municipal government is responsible

\footnote{1 http://mobile.reuters.com/article/idUSBRE9BI11I20131219?irpc=932}
\footnote{2 E.g.: DiTella and Schargodski (2003) for the effects of a crackdown on corruption in hospital purchases; Olken (2007) for the effects of government audits on road projects; and Reinikka and Svensson (2005) for the effects of a newspaper campaign on government transfers to schools.}
\footnote{3 See Finan, Pande and Olken (2016).}
for a wide array of activities, from contracting hospital reforms to purchasing vaccines to paying public servants’ wages. Managing millions of dollars every year, these bureaucrats have several ways to embezzle public funds: e.g.: by overbilling the government for a hospital reform, by settling off-the-record transactions with vaccine vendors, or by having ghost employees on the public sector payroll. Since most Brazilian municipalities rely almost exclusively on federal redistribution in order to provide local public goods, bureaucrats’ moral hazard problem is particularly acute: citizens have weaker incentives to monitor spending funded by taxes that were not locally raised.4

To address this problem, the federal government introduced a monitoring technology in 2003: the Brazilian anti-corruption program. The program randomly draws municipalities to be audited with respect to their use of federal funds, in a joint venture with the national lottery. Auditors analyze municipalities’ accounts and documentation, and physically inspect public works and service delivery to assess whether earmarked federal transfers are effectively spent according to their guidelines.

To estimate the effects of the program, the key constraint is data: understanding its effects on corruption requires observing how the latter varies over time within municipalities. We take advantage of the fact that auditors follow transfers’ paper trail back, for at least 3 years prior to the time of the audit. Drawing upon a unique dataset for the incidence of corruption across earmarked health transfers between 1997 and 2007, we observe all decisions within transfers that were subsequently audited – even before the audit actually took place.

As in Avis, Ferraz and Finan (2017), we take advantage of the randomness of recent audits that took place at a municipality or its neighbors, as those have been shown to make the perceived probability of being audited by the program in the future more salient. Nevertheless, we go beyond, by also investigating the effects of the announcement of the program, for two reasons. First, we do not have as many draws in our dataset as they do, causing our point estimates of recent audits on overall corruption not very precisely estimated, even if of very similar magnitude to theirs. Second, and most importantly, while recent nearby audits might have induced variation in corruption, this is only part of the story: most of the effects of the program might have occurred when it was announced. If that is the case, then the effects (or lack of thereof) of recent audits on public service delivery may be elusive, since most adjustments may have occurred at the time of announcement, even before any audits took place.

To investigate the overall effect of the program, we resort to a differences-in-differences strategy. The first difference contrasts municipalities before and after local procurement staff learned they could be audited. The Brazilian anti-corruption program is particularly suitable for this strategy: since it was

---

4 This same argument is linked to what the public finance literature calls the “flypaper effect”; see Schwartz (2005).
introduced exactly at the midpoint of the 2001-04 term, we can assess its effects holding decision-makers fixed. The second difference contrasts sets of transfers with different opportunities for embezzlement. The hypothesis is that the program should have had differential effects across transfers with different scope for corruption (while the differential incidence of corruption across these transfers should have remained constant otherwise). Since corruption is fundamentally linked to procurement problems, we compare sets of transfers with different “procurement intensities”. We code this variable as the percentage of transfer’s actions involving procurement-related words (such as “construction”, “expansion” or “acquisition”), following the Health Ministry’s official description of all actions under each transfer.\(^5\)

We find that the anti-corruption program substantially reduced corruption within health transfers, decreasing occurrences of over-invoicing and off-the-record payments, and of procurement irregularities such as participation of ghost firms or tailoring terms of references to specific vendors. Our results suggest that the program decreased corruption by over half its baseline prevalence within the 2001-04 term. Moreover, the effect does not dissipate during the subsequent term, and is unaffected by controlling for proximity to elections. Most of the effect is driven by the program’s announcement: while differential corruption across transfers is also significantly reduced by recent nearby audits, responses to actual audits are much less dramatic than those to the announcement of the program. This pattern validates our empirical strategy: since the bulk of deterrence effects kicked-in even before any audits actually took place, program’s announcement is key to understanding its effects on public service delivery.

Next, we investigate the effects of the anti-corruption program on other dimensions of bureaucratic performance. First, we focus on the quantity margin by looking at public spending. Using audited amounts as a proxy for spending in each set of transfers, the differences-in-differences estimate suggests that public spending plummeted by at least 50% after the program was introduced. In fact, the effect of the program on public spending is so dramatic that corruption per dollar spent may have actually increased after the program. Second, we focus on implementation quality, taking advantage of the richness of information from audit reports. We find that mismanagement rose one-to-one with the fall in corruption, particularly in what comes to problems linked to the quality of health infrastructure and to lack of medication in stock. These results are consistent with those responsible for procurement reducing purchases after the program, either because bureaucrats no longer can capture rents to the same extent, or because they are now afraid of being punished for accidental procurement mistakes.

While few papers document the effects of monitoring on the incidence of corruption (an exception is Avis, Ferraz and Finan, 2017), even fewer analyze its effects on the outcomes that corruption is supposed to detrimentally affect. To investigate if the anti-corruption program improved health indicators, we resort

\(^5\) See Appendix A.
to the Health Ministry’s Monitoring & Evaluation framework, which specifies which outputs and outcomes are used to track the implementation quality of each earmarked federal transfer to municipalities. In order to exploit the same empirical strategy, we include in our analysis outputs and outcomes that are exclusively linked to transfers with high or low procurement intensity. We can contrast outputs and outcomes such as hospital beds and immunization coverage – which are tied to procurement-intensive transfers –, on the one hand, to outcomes such as the population share covered by family doctors and medical consultations per thousand inhabitants – which are tied to transfers that are not procurement-intensive –, on the other. In additional specifications, we contrast outputs that are directly affected by municipalities’ health spending, such as preventable deaths, to those that are not, such as deaths by external causes; and contrast leading health outcomes, such as malnutrition and infant mortality, across municipalities with different corruption prevalence at baseline – hence, differentially exposed to the effects of the program.

Consistent with the effects of the anti-corruption program on other margins of bureaucratic performance, we find that it makes health indicators significantly worse. Relative to the counterfactual expansion of outputs and outcomes linked to transfers with little or no procurement, the program reduced per capita hospital beds, immunization coverage, and the share of households with access to piped water, connected to the sewage network or with septic tanks. Effects are sizable, equivalent to the average effect of losing between half and all support from federal transfers towards municipality’ budget, consistent with the magnitude of the effect of the program on spending. Contrary to the hypothesis that such negative effects might result from temporary adjustments to the new incentives set out by the program, effects become larger and more precisely estimated when we include the subsequent political term in the analysis.

Our findings are at odds with those of Reinikka and Svensson (2005), which document sizable positive effects of a newspaper campaign in Uganda that disclosed official figures about the central government’s transfers to local primary schools. One possible explanation for this difference is that while Reinikka and Svensson (2005) consider a single transfer that is delivered directly to schools, this paper considers the universe of earmarked federal transfers to municipalities. Before they reach citizens, such funds must be used by local bureaucrats to procure goods and services, subject to a complex set of procurement guidelines. This is the typical budget implementation process in developing countries, and the effects we find are consistent with the effects of monitoring bureaucratic performance documented elsewhere (Rasul and Rogger, 2015; Shi, 2008).

Overall, our results provide first-hand evidence that exposing corruption may be detrimental to the outcomes that we ultimately care about. While the Brazilian anti-corruption program represents a major

---

6 For the transfers with no M&E indicators, we resort to the Health Ministry’s description of all actions under each health transfer to define proxies for its outputs and outcomes; see Appendix A.
improvement in monitoring and transparency, the focus of administrative penalties and of public opinion on corruption, instead of on the quality of public services, all seem to have thrown the baby out with the bathwater. Expanding the scope of desirable outcomes beyond formal procedures, differentiating between active and passive waste (Bandiera, Pratt and Valetti, 2009), and supporting local procurement staff in complying with complex guidelines might be important steps towards balancing incentives between procuring, on the one hand, and making proper use of public funds, on the other.

The remainder of this paper is organized as follows. Section 2 briefly introduces the Brazilian anti-corruption program. Section 3 presents a simple model of bureaucratic decisions, illustrating the relevant trade-offs and providing predictions for the effects of an increase in the audit probability on corruption and spending. Section 4 describes the empirical strategy, and presents the results for the effects of the program on corruption. Section 5 analyzes other margins of bureaucratic performance, assessing the impacts of the program on public spending and implementation quality. Section 6 presents the results for the effects of the program on health indicators. Section 7 concludes the paper.

2 The Brazilian anti-corruption program

The Brazilian anti-corruption program (Programa de Sorteios Públicos) is a joint venture between the Office of the Comptroller General (Controladoria Geral da União, CGU henceforth) and the national lottery. It is based on periodic televised random draws that select municipalities to be audited by CGU officials. Auditors analyze municipalities’ accounts and documentation, and physically inspect public works and service delivery, to assess whether earmarked federal transfers are effectively spent according to their guidelines. CGU’s official website points out that auditors also interact with local councils and civil society organizations to empower local citizens in playing an effective role in monitoring the use of tax revenues.

The program was announced in January 2003. Four municipalities were audited as a pilot in February; this was followed by the announcement, in March, of the municipalities selected in the first draw to be audited in April. There was no decree or media announcement of the program prior to 2003. While other overseeing institutions were already in place prior to the introduction of the Brazilian anti-corruption program, its announcement represented not only a substantial increase in the probability of municipalities being audited systematically, but also on the intensity of electoral damages of being exposed as corrupt.

---

7 See Appendix B for more details about the program.
Summary audit reports are systematically broadcasted on the internet, newspapers, television, and radio, and there is evidence that voters punish corrupt mayors exposed by the program (Ferraz and Finan, 2008).\(^8\)

**Figure** – Timeline of the announcement of the program

Beyond electoral punishment, a number of administrative penalties were applied based on CGU audit reports, and even Federal Police operations were triggered by evidence put forward by the anti-corruption program. To date, conviction for involvement in procurement irregularities exposed by the program is one of the very few reasons for which one can lose tenure at the public sector. Furthermore, public servants convicted for corruption have to repay the embezzled amount out of their own pockets or go to prison.

Finally, no other major institutional changes took place over the same period. To date, the anti-corruption program is still the most important corruption-deterrence mechanism in Brazil, and the fact that several political attempts have been made to terminate it, to decrease the number of audited municipalities, or to increase the time span between draws is testimony to its first-order effects on the Brazilian bureaucratic politics’ status quo.

### 3 A simple model

This section discusses a model of bureaucratic decisions. The goal is to illustrate the relevant trade-offs faced by local procurement staff, and to provide predictions for the effects of an increase in the audit probability on different margins of bureaucrats’ action space.

Consider a risk-neutral bureaucrat (\(B\)) deciding whether to use public funds transferred by the federal government to procure or not. Procurement involves an effort cost (\(c\)), and entails a benefit (\(G\)) which can

---

\(^8\) Pande (2011) documents that information experiments lead voters to punish corrupt politicians more broadly.
be thought of as personal utility or managerial rewards for providing public services. If she does not procure, there is no further action. If she does procure, she can decide whether to be corrupt or not. Corruption entails sharing the surplus (\(\pi\)) of the contractual relationship with a vendor, with the share of the surplus allocated to the bureaucrat (\(\alpha\)) determined by an exogenous process, such as Nash bargaining. If she is not corrupt, then the \(N\) identical vendors compete for the surplus (expected surplus for the outbidder is \(\frac{\pi}{N}\)) and the bureaucrat gets no share of it.

The probability of being audited and being found corrupt is \(p\). If exposed as corrupt, the bureaucrat faces a punishment (\(R\)), which can be thought of as an administrative or reputational penalty. If the bureaucrat does not procure, being audited implies no punishment. If she procures and is corrupt, the expected punishment is \(pR\). Even if she is not corrupt, however, it is still possible that she faces a penalty, as procurement mistakes (which happen with probability \(\varepsilon\)) are also coded as corruption by auditors – even when bureaucrats take no share of the surplus.

**Figure** – Bureaucrats’ payoffs from procurement and corruption decisions

\[
\begin{align*}
\text{Procurement} & \quad \text{Corruption} \\
\text{No Procurement} & \quad \text{No Corruption}
\end{align*}
\]

\[
\begin{align*}
(0,0) & \quad (G - c - p\varepsilon R, \frac{\pi}{N}) \\
& \quad (G - c + \alpha\pi - pR, (1 - \alpha)\pi)
\end{align*}
\]

This simple model illustrates the relevant trade-offs. Upon procuring, \(B\) is corrupt if \(\alpha\pi \geq (1 - \varepsilon)pR\), that is, if her share of the contractual surplus is high enough, and if the expected punishment \textit{in excess of} that coming from the possibility of procurement mistakes is not too high. An increase in the probability of being exposed as corrupt, \(p\), decreases the likelihood of corruption. What is more, if there is a range of
transfers with different embezzlement opportunities ($\alpha \pi$), an increase in $p$ will deter corruption particularly for transfers with the lowest upside from corruption.

Next, $B$ procures if $c + pR \leq G + \alpha \pi$ when $B$ is corrupt, or if $c + p\varepsilon R \leq G$ otherwise, that is, if the expected costs of procurement – which include punishment from being exposed as corrupt (having embezzled resources or not) – are not higher than the benefits of public service provision. An increase in the probability of being exposed as corrupt, $p$, decreases the likelihood of procurement. It does so particularly for transfers with the lowest upside from corruption.

The model illustrates how an increase in the audit probability is expected to decrease both corruption and spending. Spending is expected to decrease particularly for transfers with the lowest embezzlement opportunities, in which bureaucrats would not have chosen to capture resources had they used those funds to procure public goods and services. For this reason, if embezzlement opportunities are proportional to transfer amounts, corruption per dollar spent is expected to increase with an increase in $p$.

While the model certainly abstracts from many relevant features, its predictions are robust to the introduction of more complex elements, such as punishments for lack of procurement (as long as lower than those for corruption), vendors’ interactions with bureaucrats (as long as punishment lays primarily on bureaucrats), or making the procurement decision a continuum (allowing for partial spending) rather than binary.

Since both corruption and spending are expected to decrease with a higher audit probability, the prediction of the model for the quality of public service delivery is ambiguous. Ultimately, the effect of an anti-corruption program on outcomes depends on the combination of, one the one hand, more resources directed to public service delivery within transfers for which bureaucrats are no longer corrupt and still find it worthwhile to undertake procurement, and, on the other, less resources directed to public service delivery within transfers for which bureaucrats no longer find it worthwhile to undertake procurement.

4 Has the anti-corruption program reduced corruption?

This section first introduces our unique dataset in subsection 4.1. Next, subsection 4.2 discusses our empirical strategy for identifying the causal effect of the program on corruption prevalence, and for decomposing it into that of its announcement and that of actual audits. Subsection 4.3 presents the results, followed robustness checks in subsection 4.4.
4.1 Data

We draw upon a unique dataset assembled by Fundação Getulio Vargas' Center of Politics and Economics of the Public Sector (CEPESP-FGV/SP), based on audit reports for constitutionally mandated health transfers. Although auditors investigate transfers linked to several ministries, including Education, Transports, and Environment, it is only for audit reports within Health transfers (up to July 2007) that the year of each instance investigated by auditors was coded. Each audit report was independently coded by two research assistants, who did not know what the data would be used for; disagreements were solved by assigning the report to a third research assistant.

In our dataset, transfers are dated according to the auditor's record of the year the action investigated accrues to. For example, consider the audit report for Quissamã, selected in the 8th draw (March 2004). While the audit took place in 2004, auditors have looked into transfers’ life cycle back to 2001. As an example, auditors note that resources transferred in December 2001, which should have been kept in a separate bank account, were rather kept in municipality’s health secretariat general account. We categorize this irregularity under documentation or accounting problems, coded in our dataset as evidence of mismanagement.

Our sample has 11,419 investigations. Of those, 10,538 are linked to health transfers for which we have coded procurement intensity, distributed across 9 health transfers and 731 municipalities, between 1997 to 2007. Procurement intensity is coded as follows. First, drawing upon Health Ministry’s list of actions under each transfer (Funcional Programática), we coded each action as procurement-related or not, based on whether its official description involves words such as “inputs”, “modernization”, or “acquisition”. Second, for each transfer, we compute the percentage of actions coded as procurement-related. In our regressions, we use an indicator variable for high procurement intensity, equal to 1 if the transfer’s procurement intensity is 50% or higher, and 0 otherwise. In our sample, 44.8% of investigations are linked to procurement-intensive transfers.

Due to the retrospective nature of audits, we have a significant number of observations (1,324 investigations) for the 2001-2002 period, although, as expected, there are more observations for the 2003-2004 period (5,748 investigations). Our dataset includes 33 types of irregularities (besides compliance, when auditors find no irregularity), ranging from documentation problems to off-the-record invoice. One

---

9 More specifically, we have data for draws 2 to 24 (draw 1 was a pilot).
10 See Table A1 in Appendix A for the complete classification list, and Appendix C for more examples of pre- and post-program investigations described in audit reports.
11 Those are the 9 most prevalent health transfers in CEPESP’s dataset, representing 92.3% of the total number of coded investigations, see Table A2 in Appendix A.
12 See Table A3 in Appendix A for the complete list.
13 Results are robust to the choice of the cutoff, see the Supplementary Appendix.
third of these irregularities are coded as evidence of corruption, and the remainder 21 as evidence of mismanagement. We follow Ferraz and Finan (2008, 2011) in defining corruption as procurement problems, over-invoicing and off-the-record invoicing.\(^\text{14}\) Over 2001 and 2002, the first half of the political term, 30.8\% of the investigations in our dataset are coded as evidence of corruption. The complete classification list is included in Appendix A.

We also have socio-demographic data at the municipality level from the Brazilian Institute of Geography and Statistics’ 2000 census; mayor characteristics and administrative attributes from the 2000 Municipal Information Dataset (Base de Informações Municipais, BIM); political variables, such as turnout and political alignment with the state governor, for each political term from the Superior Electoral Tribunal (Tribunal Superior Eleitoral, TSE); and public finance data from the National Treasury and from the Information System for Public Budgets in Health (Sistema de Informações sobre Orçamentos Públicos em Saúde, SIOPS).

4.2 Empirical strategy

As in Avis, Ferraz and Finan (2017), we take advantage of the randomness of recent audits that took place at a municipality or its neighbors, as those have been shown to make the perceived probability of being audited by the program in the future more salient. Nevertheless, we go beyond, by also investigating the effects of the announcement of the program, for two reasons. First, we do not have as many draws in our dataset as they do, causing our point estimates of recent audits on overall corruption not very precisely estimated, even if of very similar magnitude to theirs.\(^\text{15}\) Second, and most importantly, while recent nearby audits might have induced variation in corruption, this is only part of the story: most of the effects of the program might have occurred when it was announced. If that is the case, then the effects (or lack of thereof) of recent audits on public service delivery may be elusive, since most adjustments may have occurred at the time of announcement, even before any audits took place.

To investigate the overall effect of the program, we resort to a differences-in-differences strategy. The first dimension of comparison contrasts municipalities before and after local procurement staff learned they

\(^{14}\) Whenever auditors point out evidence of corruption, we restrict attention to the incumbent’s extensive margin decision of whether or not to be corrupt, rather than to the intensive margin decision of how much to embezzle. The reason is that the exact embezzled amount is rarely included in audit reports. We have not tried to replicate Ferraz and Finan (2008)’s effort to proxy for this amount from the description of “missing items” in the event of corruption.

\(^{15}\) We show in the Supplementary Appendix that, with municipality fixed-effects and year fixed-effects, nearby audits have a negative but statistically insignificant effect on overall corruption prevalence in our dataset. In Avis, Ferraz and Finan (2017), access to administrative CGU data allows for a sample that extends over a longer time span, in particular with many more cases of municipalities audited more than once, improving precision of the estimates.
could be audited. The second dimension of comparison contrasts sets of transfers with high and low procurement intensities.

The idea is that each group is subject to different treatment intensities: procurement-intensive transfers, for which 37% of investigations are coded as corruption in 2001-2002, have a higher scope for reducing corruption after the program is introduced than other transfers, for which only 17% of investigations are coded as corruption over the same period. The differences-in-differences estimator documents whether procurement-intensive transfers experienced a systematically different variation in corruption prevalence following the introduction of the program.

The differences-in-differences estimate captures the causal effect of the program under the identification assumption of identical potential reported outcomes. Under additive measurement error from misreporting by auditors, sufficient conditions for this assumption to hold are (1) parallel trends for actual potential outcomes across the two sets of transfers; and (2) parallel trends for misreporting by auditors across the two sets of transfers. The first hypothesis is the typical identification assumption for the differences-in-differences estimator. The second hypothesis means that auditors can differentially misreport corruption over time – as long as equally so for the two sets of transfers –, or differentially misreport corruption across the two sets of transfers – as long as equally so over time. If auditors misreport corruption differentially across transfers and over time, then misreporting would confound the effects of the program captured by the differences-in-differences strategy.

The empirical strategy for identifying the overall effect of the program is described by the following equation:

\[
\text{Corruption}_{mit} = \alpha + \theta_m + \theta_t + \beta \text{Post}_t \times \text{ProcIntensive}_{mit} + \gamma \text{ProcIntensive}_{mit} + \epsilon_{mit} \tag{1}
\]

In equation (1), \(\text{Corruption}_{mit}\) equals 1 if investigation \(i\) in municipality \(m\) at year \(t\) was coded as evidence of corruption, and 0 otherwise; \(\text{Post}_t = 1\) from 2003 on, and 0 otherwise; \(\text{ProcIntensive}_{mit}\) equals 1 for procurement-intensive transfers, and 0 otherwise; \(\theta_m\) are municipal fixed-effects; \(\theta_t\) are year fixed-effects; and \(\epsilon_{mit}\) is a zero-mean error term. Under the assumption of identical potential reported outcomes between high and low procurement-intensity transfers, \(\beta\) identifies the causal effect of the anti-corruption program on corruption. All estimates are from OLS regressions, clustering standard errors at the municipality level to deal with potential serial correlation of residuals (Bertrand, Duflo and Mullainathan, 2004). The prediction we are interested in is \(\beta < 0\).

---

\(^{16}\) See Supplementary Appendix.
We can combine the differences-in-differences strategy with recent nearby audits to separately estimate the effect of the announcement of the program and that of the actual audits. On the one hand, the differences-in-differences estimate captures the total effect of the anti-corruption program. On the other hand, once we control for the differential effect of recent nearby audits across transfers with different procurement intensities, the change in the differences-in-differences coefficient across specifications allows pinning down the effect of the program’s announcement separately from that of the audits themselves. The following equation describes how we modify equation (1) in order to estimate the two effects separately:

\[
\text{Corruption}_{mit} = \alpha + \theta_m + \theta_t + \beta_1 \text{Post}_t \times \text{ProcIntensive}_{mit} + \gamma \text{ProcIntensive}_{mit} + \\
\beta_2 \text{NearbyAudits}_{m,t-1} \times \text{ProcIntensive}_{mit} + \delta \text{NearbyAudits}_{m,t-1} + \epsilon_{mit} \tag{2}
\]

\( \text{NearbyAudits}_{m,t-1} = 1 \) at year \( t \) if there was an audit within \( R \) km of municipality \( m \)’s centroid in the previous year (including, possibly, \( m \) itself facing an audit), and 0 otherwise; where \( R \) is the maximum distance for which nearby audits are still estimated to decrease subsequent corruption. To compute this indicator variable, we estimate a non-parametric relationship between the minimum distance from municipality’s centroid to audits in the previous year and the within-municipality subsequent variation in corruption prevalence, and compute \( R \) as the maximum distance for which the upper bound of the confidence interval of this estimate is below 0.\(^{17}\)

Once again, all estimates are from OLS regressions, clustering standard errors at the municipality level. Because \( \text{NearbyAudits}_{m,t-1} \) is a generated independent variable, we bootstrap standard errors using a two-step procedure described in Supplementary Appendix A. The predictions we are interested in are \( \beta_1 \) \( < 0 \) and \( \beta_2 \) \( < 0 \). The difference between \( \hat{\beta} \) estimated from equation (1) and \( \hat{\beta}_1 \) estimated from equation (2) is that \( \hat{\beta}_1 \) parses out the effect of the actual audits, and hence stands for the effect of the announcement of the program. We compute \( \hat{\beta}_1 / \hat{\beta} \), the estimated share of the overall effect of the program coming from its announcement.

Finally, even though the identification assumption for the differences-in-differences estimator cannot be tested, if the differential effect of recent nearby audits across transfers of different procurement

\(^{17}\) In the local polynomial smoothing, confidence intervals weight observations by the number of investigations used to generate municipality’s share of investigations coded as corruption. To compute those shares, we use all 11,419 investigations between 1997 and 2007 in our dataset, regardless of whether we have coded procurement intensity for the transfer to which they are linked.
intensities ($\hat{\beta}_2$) agrees with the differences-in-differences estimate ($\hat{\beta}_1$), then we can confidently rule out alternative explanations for the latter.

4.3 Results

We start by considering the yearly average of the share of investigations coded as corruption for each set of transfers, from 1997 on, displayed in Figure 1. Restricting attention to the 2001-2004 term, not only does corruption for all transfers fall systematically after the program, but also, from 2003 onwards, the gap in corruption prevalence between high and low procurement-intensity transfers basically ceases to exist.

[Figure 1]

Furthermore, there seems to be an upward trend in corruption right before the program’s announcement, consistent with the idea that the latter was not anticipated by local procurement staff. Last, even though the identification assumption for the differences-in-differences strategy cannot be tested, within the term, the two sets of transfers had basically parallel trends for the share of investigations coded as corruption before the program was introduced.

Before the program, procurement-intensive transfers had a 20 p.p. higher share of investigations coded as corruption relative to other transfers (37% vs. 17%); after the program, however, this difference was reduced to zero. Corruption among low procurement-intensity transfers fell by 5 percentage points (from 17% to 12% of investigations) after the program, while that for transfers with high procurement intensity fell by 25 percentage points (from 37% to 12%). All in all, the differences-in-differences estimate based on unconditional averages suggests that the program decreased corruption by 20 percentage points, an effect over two-fold the baseline prevalence of corruption among low procurement-intensity transfers.

Since differences in unconditional averages could be confounded by the effects of other variables, in particular by compositional changes of municipal characteristics linked to investigations in each year, we take on to the regression analysis, exploiting within-municipality differences over time across the two sets of transfers. Results are presented in Table 1. Columns (1) and (2) restrict attention to the 2001-2004 term, over which we can estimate the effects of the program while holding decision-makers fixed. Column (1) controls for a post-program indicator variable, while column (2) includes year fixed-effects. In both cases, the program is estimated to have reduced corruption by at least 17.8 p.p., and the effect is statistically significant at the 1% level.
Column (3) includes the previous and subsequent political terms in the analysis, extending it to the universe of investigations for which we have coded procurement intensity. Doing so provides evidence about the extent to which the overall effect of the program previously identified is short-lived or whether it consolidates over the following term. Results are basically unchanged relative to the estimates for the 2001-04 sample. Beyond a larger sample size, including other political terms can help us rule out alternative explanations for the effects previously documented; in particular, proximity to elections, since the program kicks off in the second-half of the term. For this reason, column (4) also includes a second-half of the term indicator, equal to 1 for 1999, 2000, 2003, 2004, and 2007, and 0 otherwise, allowing it to differentially affect procurement-intensive transfers. While proximity to elections indeed significantly reduces corruption (and particularly so for procurement-intensive transfers), the effect of the program on corruption remains at -17.6 p.p. and significant at the 1% level, remarkably close to the original estimate.

Finally, columns (5) and (6) decompose the overall effect of the program into that of the announcement and that of actual audits, the former for the 2001-04 sample, and the latter for 1997-2007. Results are as follows. First, recent nearby audits decrease differential corruption prevalence across transfers of different procurement intensities, by at least 4.5 p.p. in both cases, although only statistically significant at the 10% level when we include the full sample. This is reassuring, since this source of variation is random conditional on municipality fixed-effects and year fixed-effects (see Supplementary Appendix), and since the direction of the effects agrees with the differences-in-differences estimate. Second, about 88.5% of the overall effect of the program comes from its announcement, confirming the caution about the potential limited effects of recent nearby audits. This pattern validates our empirical strategy: since the bulk of deterrence effects kicked-in even before any audits actually took place, program’s announcement is key to understanding its effects on public service delivery.

We present two additional analyses to provide further evidence that the effects are driven by the program. In Figure 2, we allow the differences-in-differences coefficient to vary by year, in order to flexibly estimate the timing of effects. Effects becomes negative only for the post-program, and coefficients are not statistically different from zero up until that year.
Next, we pursue a variation of the previous analysis, allowing the differences-in-differences coefficient to vary by draw. Because the local procurement staff may take some time to adjust after learning about the program (and since tampering with procurement paper trail is also evidence of corruption), one would expect to see the effects kick-in only after several draws. Consistent with that hypothesis, Figure 3 shows basically no effects until the 7th draw, followed by a clear downward slope for the coefficients, particularly after the 9th draw. The analysis suggests corruption to be quite sticky: it takes almost 2 years for the differences-in-differences coefficient to differ systematically from zero (lottery 14 was drawn on November 17, 2004).

[Figure 3]

Last, the Supplementary Appendix breaks down the analysis by mayor’s political term. Results point out that the program reduced corruption amongst second-term mayors by about two-fold its effect on those serving their first political term. This pattern is consistent with both higher corruption prevalence amongst lame duck politicians (Ferraz and Finan, 2010) and with most of the effect of audits operating through legal discipline, rather than electoral discipline (Avis, Ferraz and Finan, 2017).

4.4 Robustness checks

This subsection addresses alternative explanations for the effects previously documented, and assesses the sensitivity of results to the definition of some of the variables. First, it might be that the political pressure around the anti-corruption program led auditors to investigate recent evidence to a lesser extent, particularly with respect to transfers with higher scope for embezzlement. While CGU highlights that auditors have no discretion over investigations (they have to follow service orders which pre-specify all investigations; see Appendix B), selective auditing could still result from the way that service orders are issued. To test this hypothesis, the Supplementary Appendix shows that the number of investigations conducted by auditors does not differentially affect corruption across different procurement-intensity categories after the onset of the program.

The Supplementary Appendix also documents that results are robust to using the number of neighbor municipalities audited within 50km and 100km in the previous year as an alternative definition of audit threats, to using different cutoffs for the procurement-intensive indicator variable, and to replacing the latter by a continuous measure of transfers’ procurement-intensity.
5 Other dimensions of bureaucratic performance

Having shown that the program significantly reduced corruption prevalence, we move on to investigate its effects on other dimensions of bureaucrats’ decisions. Subsection 5.1 analyzes the quantity margin, by focusing on how public spending evolved for transfers of different procurement-intensities, before and after the introduction of the program. In subsection 5.2, we analyze the quality of implementation, as auditors record a range of implementation problems in audit reports, allowing us to explore the richness of our dataset to test whether implementation quality falls alongside the documented fall in public spending. Subsection 5.3 presents robustness checks. In subsection 5.4, we discuss how to interpret our results in face of minimum spending mandates in Brazil, and tackle the issue of what happened to idle resources – a question with direct implications for welfare analysis.

5.1 Quantity

Budget execution of federal transfers to municipalities in Brazil is notably low. For the federal government’s Growth Acceleration Program (PAC) in health (a voluntary transfer), budget execution has been around 10%, and municipalities’ inability to streamline procurement has been credited as the most important reason for such low execution rates. In this subsection, we test the hypothesis that the Brazilian anti-corruption program reduced public spending among earmarked federal transfers to municipalities within health.

The biggest challenge to answering this question is spending data. Municipal health spending is recorded by the Brazilian Dataset on Municipal Budgets (FINBRA) only for the total budget, not separately for constitutionally mandated transfers. Since these transfers do not legally admit budget leftovers (such that amounts only partially executed cannot be returned to the federal government), there are no official records of how much municipalities actually spent out of those transfers. As an example, to analyze potential displacement effects of audits on spending, Avis, Ferraz and Finan (2017) have to rely on data from the Brazilian Institute of Applied Economic Research (IPEA), which reflect planned budget figures, rather than budget execution.

Having said that, the richness of our dataset allows us to proxy for public spending using audited amounts. Once again, because auditors leave CGU’s headquarters with pre-set service orders, they do not have discretion over which transfers (or how much) to audit. Any systematic differences in audited amounts across transfers over time must accrue to differential spending patterns. More specifically, the idea is that,

---

19 This has led some authors to wrongly assume that budget execution for these transfers is always 100% (Zamboni and Litschig, 2013).
beyond compliance with documentation and account keeping – with respect to which all transfers are always audited on –, only spending moves a transfer up in its life-cycle and generates further objects of investigations by the anti-corruption program.

We again combine the differences-in-differences estimator with recent nearby audits as our empirical strategy. The dependent variable is the natural logarithm of the audited amount for each investigation, as in the following equation:

\[
\ln(Amount_{mit}) = \alpha + \theta_m + \theta_t + \beta_1 Post_t \times ProIntensive_{mit} + \gamma ProIntensive_{mit} + \\
\beta_2 NearbyAudits_{m,t-1} \times ProIntensive_{mit} + \delta NearbyAudits_{m,t-1} + \epsilon_{mit} \tag{3}
\]

In equation (3), \(\ln(Amount_{mit})\) is the audited amount (in logarithms) for investigation \(i\) at municipality \(m\) at year \(t\), \(Post_t = 1\) from 2003 on, and 0 otherwise; \(ProIntensive_{mit}\) equals 1 for procurement-intensive transfers, and 0 otherwise; \(NearbyAudits_{m,t-1} = 1\) at year \(t\) if there was an audit within \(R\) km of municipality \(m\)’s centroid in the previous year (including, possibly, \(m\) itself facing an audit), and 0 otherwise, where \(R\) is the maximum distance for which nearby audits are still estimated to decrease subsequent corruption; \(\theta_m\) are municipal fixed-effects; \(\theta_t\) are year fixed-effects; and \(\epsilon_{mit}\) is a zero-mean error term. Under the assumption of identical potential reported outcomes between high and low procurement-intensity transfers, \(\beta_1\) and \(\beta_2\) identify the causal effects of the anti-corruption program on public spending. All estimates are from OLS regressions, clustering standard errors at the municipality level. The predictions we are interested in are \(\beta_1, \beta_2 < 0\).

Figure 4 shows that, within the 2001-2004 term, audited amounts for the two sets of transfers followed similar trends over the pre-program period. While amounts for both sets of transfers fall substantially after the program (at least until 2005), amounts for procurement-intensive transfers fall significantly more, and the gap between the low and high procurement-intensity transfers is increasing over time.

[Figure 4]

Table 2 considers the empirical regularity hinted at by the previous figure within a regression framework. Columns (1) and (2) restrict attention to the 2001-2004 term, over which we can estimate the effects of the program while holding decision-makers fixed. Column (1) controls for a post-program
indicator variable, while column (2) includes year fixed-effects. In both cases, the program decreases spending by about 50%, a huge effect statistically significant at the 1% level.

[Table 2]

Column (3) includes the previous and subsequent political terms in the analysis, extending it to the universe of investigations for which we have coded procurement intensity. Effects are substantially larger (about 70%) and still significant at the 1% level. Column (4) includes a second-half of the term indicator, allowing it to differentially affect procurement-intensive transfers. While proximity to elections also significantly deteriorates spending (what is expected, since municipalities cannot procure for most of the final year of the term), the effect of the program on corruption remains significant at the 1% level and remarkably close to column (3)’s estimate.

Finally, columns (5) and (6) include the indicator of recent nearby audits and its interaction with procurement-intensity, considering the 2001-2004 term and our whole sample, respectively. We find that audits have a negative and significant effect on public spending of about 37% in both cases. Once more, since the effects of recent nearby audits are in the same direction and order of magnitude of those of the differences-in-differences estimator, we can be confident that results are not driven by alternative explanations. Last, between 50% and 75% of the overall effect of the program on public spending comes from its announcement, consistent with the effects of the program on corruption.

To highlight how striking the results of Table 2 are, we can decompose variation in corruption over time as follows:

\[ C_t = \frac{Y_t}{T_t} = \left( \frac{Y_t}{X_t} \right) \left( \frac{X_t}{T_t} \right) \] (4)

In equation (4), the first equality defines corruption at time \( t \), \( C_t \), as the ratio between the number of corruption events at time \( t \), \( Y_t \), and the number of investigations at time \( t \), \( T_t \). The second equality just rewrites this ratio as the product between corruption events per dollar spent at time \( t \), \( \frac{Y_t}{X_t} \), multiplied by the average spending per investigation at time \( t \), \( \frac{X_t}{T_t} \). Taking logs and differentiating with respect to time yields:
\[
\frac{d \ln(C_t)}{d t} = \frac{d \ln(Y_t)}{d t} + \frac{d \ln(X_t)}{d t} \tag{5}
\]

The first term on the right-hand side of equation (5) is the percentage change in corruption if average spending per investigation was held constant; the second term is the percentage change in average spending.

With \(\frac{d \ln(C_t)}{d t} \equiv -0.18\), according to the results in column (2) of Table 1, and with \(\frac{d \ln(X_t)}{d t} \equiv -0.49\), according to the results of column (2) of Table 2, average corruption per dollar spent must have increased by 31 percentage points; otherwise the share of investigations coded as corruption would have been reduced even further.

Our findings for the effects of the program on corruption and public spending are consistent with the simple model introduced in Section 3. The higher probability of being audited (and exposed as corrupt for being involved in procurement problems) discourages both corruption and spending, particularly among transfers involving lower embezzlement opportunities. If those opportunities are proportional to transfer amounts, the model predicts that a higher audit probability increases corruption per dollar spent.

### 5.2 Quality

This subsection provides additional evidence about the mechanism linking lower corruption to worse health indicators, by looking at the quality of policy implementation. We start by decomposing investigations not coded as “evidence of corruption” into mismanagement and compliance.\(^{20}\) We follow the same empirical strategy as equation (3), only replacing the dependent variable on the left-hand-side by indicator variables of mismanagement (equal to 1 if the investigation is coded as evidence of mismanagement, and 0 otherwise) or compliance (equal to 1 if the investigation is neither coded as evidence of corruption nor of mismanagement, and 0 otherwise).

Table 3 presents the results. Columns (1) to (3) restrict attention to the 2001-04 term, while columns (4) to (6) include all investigations between 1997 and 2007 in the analysis. Looking first at the effect of the program’s announcement, we find that the decrease in corruption (columns 1 and 4) translates basically one-to-one into an increase in mismanagement (columns 2 and 5). Compliance (columns 3 and 6) is not significantly affected by the announcement of the anti-corruption program.

\[\text{[Table 3]}\]

\(^{20}\) See Appendix A for the complete list of irregularities and how they are coded in our dataset.
For recent nearby audits, it seems that results are mixed: in both samples, about half of the decrease in corruption seems to translate into higher mismanagement, and the remainder, to higher compliance. Having said that, the marginal effect of actual audits on corruption is small enough that their effect on compliance is not statistically significant at the 10% level.

To shed further light on what ensued after the Brazilian anti-corruption program was introduced, still resorting to the empirical strategy described by equation (3), we look separately at each category of irregularities within mismanagement in Table 4, including all investigations between 1997 and 2007. In column (1), resource diversion represents diverting resources meant to be used for the purposes of one transfer towards other transfers (within health or not); notice that this category does not include over-invoicing or off-the-record invoice, instances of resource diversion that are coded as evidence of corruption. In column (2), health council problems range from precarious facilities to below-required frequency of meetings. In column (3), performance problems stand for complaints from final users about frustrated consultations or admissions, about lack of availability of medication, or about low-quality health services. In column (4), infrastructure and stock problems encompass precarious facilities in health units (including mobile units such as ambulances), medication not properly kept in stock, or ins and outs not properly accounted for. In column (5), human resources problems represent problems with late wages, with missing personnel, or with the composition of health teams relative to transfer’s guidelines. Last, in column (6), documentation or accounting problems range from mix-ups in bank accounts supposed to be kept separate to invoices with illegible information.

First, the results show that mismanagement increased particularly for infrastructure and medication stock problems. This pattern is consistent with plummeting public spending. Second, we also find a significant increase in documentation and accounting problems. This is suggestive that local procurement staff might have tried to tamper with evidence for past procurement problems, an additional piece of evidence for the extent to which bureaucrats diverted energy from public service delivery after the program was announced.

Third, we find that resource diversion significantly decreased after the program. This finding potentially raises a different concern: reshuffling resources across different programs does not necessarily capture resources from public goods’ provision, but, possibly, reallocates those resources to local priorities (which local bureaucrats might be better equipped to identify). Last, evidence for performance problems is mixed:
while, to some extent, the announcement of the program seems to discipline the quality of service delivery to end users, recent nearby audits increase these problems (although this is not statistically significant).

Overall, the evidence suggests that the Brazilian anti-corruption program significantly decreased corruption, but distorted other margins of bureaucrats’ decisions. In response to the introduction of the program, local procurement staff reduced spending within earmarked federal transfers by at least 50%. As a result, infrastructure and medication suffered and bureaucrats felt more constrained about channeling resources towards local needs. These negative effects do not seem to result from temporary adjustments: not only did they not fade in the subsequent political term, but also became stronger and more precisely estimated.

### 5.3 Robustness checks

One channel potentially missing from the previous analysis is voluntary transfers. Brollo (2010) argues that lower federal support towards municipalities’ budget might be central to the electoral punishment documented by Ferraz and Finan (2008) once a municipality is exposed as corrupt. In our setting, this might imply that municipalities’ total health budgets become lower on average after the program, a possibility that could also deteriorate the quality of public services. Having said that, this channel would only matter for our results if voluntary transfers have differential trends before and after the program across transfers of different procurement-intensities.

While there is no annual data on municipal-level voluntary funds released to each transfer that we could link to procurement-intensity, we have annual data on municipal-level voluntary funds towards two sets of health transfers: Basic Attention, which can use those funds for a variety of purposes, including procurement, and Medium- and High-Complexity, whose guidelines explicitly preclude municipalities from using its funds to procure goods or services.\(^{21}\) For the purposes of our analysis, we code these sets of transfers as high and low procurement-intensity, respectively. The Supplementary Appendix shows that voluntary transfers do not differentially affect corruption or spending across different procurement-intensity categories after the onset of the anti-corruption program.

Last, to provide additional evidence about the mechanism, the Supplementary Appendix investigates whether public spending was particularly reduced within sets of transfers for which bureaucrats used to exert higher discretion at baseline. We first define a discretion-intensive indicator, equal to 1 for transfers

---

\(^{21}\) Basic Attention encompasses Basic Attention in Health and other health transfers, such as Basic Pharmacy. Pages 55-56 of DENASUS’ ‘Manual de auditoria na gestão dos recursos financeiros do SUS’ (2004) explicitly preclude the utilization of funds for Medium- and High-Complexity transfers for procurement. Although data for voluntary transfers towards Strategic Actions transfers is also available, we do not include it because such voluntary funds are not systematic, accruing to only about 10% of Brazilian municipalities each year.
with above-median share of resource diversion problems unrelated to corruption before the program, and 0 otherwise. We then generate a triple interaction, multiplying the discretion-intensive indicator with the procurement-intensive indicator and the post-program indicator. The hypothesis we test is whether corruption is further reduced by the program within discretion-intensive transfers, beyond its differential effect working through procurement-intensity. Consistent with the mechanism and with the results for resource diversion in Table 4, we find that corruption is further reduced among discretion-intensive transfers.

5.4 Where did the money go?

We have shown that the anti-corruption program had a very large negative effect on health spending within earmarked federal transfers to municipalities. Where did that money go? Answering this question matters for two reasons. First, this result seems to be inconsistent with mandated minimum budget shares: Brazilian municipalities have to spend at least 15% of their tax revenues on health, and most of them struggle to stay above that threshold. If spending on constitutional transfers dropped by as much as 50%, how did municipalities comply with that mandate?

Second, if municipalities really did not spend those funds, their destination has implications for welfare. For instance, if idle funds were re-budgeted by the federal government and reallocated to municipalities able to spend them, then the analysis of the effects of the program on downstream outcomes would have to take such reallocation into account.

In what comes to the first issue, minimum spending mandates are verified based on planned budget, not on budget execution. Part of the reason is that earmarked transfers have no leftovers, leading to the implicit assumption that their execution rate is 100% – clearly not the case based on our findings. It is also worth mentioning that lower spending within earmarked federal transfers may have been partially matched by higher spending within voluntary transfers and tax revenues, since those cannot be audited by CGU.22 However, even if that were the case, effects on implementation quality and outcomes (see Section 6) suggest that, even if (partial) reallocation has taken place, it introduced distortions that ultimately hurt public service delivery.

In what comes to the second issue, precisely because earmarked federal transfers admit no leftovers, idle funds just sit in municipalities’ accounts. Even though it is not possible to track leftovers for those transfers, as previously discussed, such pattern is consistent with the increase in leftovers and low overall budget execution rates since 2003.23

---

22 Fernandes and Lichand (2017) investigate this hypothesis with data from the São Paulo State Court of Accounts.
23 See https://civitarese.wordpress.com/2016/07.
6 Has the anti-corruption program improved public service delivery?

This section starts by discussing the data we use for health indicators in subsection 6.1. Next, subsection 6.2 presents the empirical strategy for identifying the causal effects of the program, including how we implement the differences-in-differences strategy and how we deal with multiple testing. Results are presented in subsection 6.3, followed by robustness checks in subsection 6.4. Subsection 6.5 discusses competing interpretations for our findings. Last, subsection 6.6 discusses how those findings relate to previous results from the corruption literature.

6.1 Data

For the annual municipal-level data for health outputs and outcomes, we rely on the National Health Database (Base de Informações de Saúde, DATASUS), restricting attention to the municipalities and years for which we have data on corruption. We assemble a dataset with the following indicators: number of hospital beds per thousand inhabitants; immunization shots per thousand inhabitants; share of households with proper sewage disposal (either connected to the general network or with septic tanks, which are widespread in rural areas); share of households connected to piped water; population share assisted by the Family Health program (Programa Saúde da Família); medical consultations per thousand inhabitants; preventable deaths; deaths by external causes; share of under-1-year-olds diagnosed with malnutrition by family doctors; share of under-2-year-olds with diarrhea episodes; and under-1-year-old infant mortality per thousand.

6.2 Empirical strategy

We start by combining the differences-in-differences strategy with recent nearby audits to estimate the effects of the program on health outputs and outcomes. In order to exploit this empirical strategy, however, we need outputs and outcomes that accrue exclusively to transfers of different procurement intensities. We

24 The immunization indicator includes 28 vaccine-preventable diseases, listed under the Epidemiological and Environmental Surveillance in Health program (VIGISUS).
25 Preventable deaths are all deaths excluding those from external causes (such as accidents or homicides). Both preventable and external deaths are coded by municipality of in-patient care, rather than by that of residence, since we are interested in the effects of the program mediated by health services.
26 The Supplementary Appendix presents correlations between corruption and health indicators at baseline. More corrupt municipalities even have better health outcomes in some dimensions. While this contrasts with the findings in Ferraz, Finan and Moreira (2012), which document a negative cross-sectional correlation between corruption and educational outcomes, in our sample more corrupt municipalities at baseline have more educated mayors and better access to public goods, such as internet coverage, which might confound the cross-sectional relationship between corruption and health outputs and outcomes.
deal with this issue by resorting to the Health Ministry’s Monitoring & Evaluation framework, which specifies which outputs and outcomes are used to track the implementation quality of each constitutionally mandated federal transfer to municipalities. According to this framework, the outputs and outcomes we include in our analysis are linked either to transfers with high procurement-intensity only, or to transfers with low procurement-intensity only. For transfers with no M&E indicators, we resort to the Health Ministry’s description of all actions under each health transfer to define proxies for its outputs and outcomes.27

The outputs and outcomes that we can include are the following. For high procurement-intensity transfers: hospital beds per thousand inhabitants, immunization shots per thousand inhabitants, household share with proper sewage disposal, and household share with access to piped water – all of which are linked to transfers with procurement-intensity 50% or higher. For low procurement-intensity transfers: population share covered by the Family Health program, and medical consultations per thousand inhabitants – both of which are linked to transfers with procurement-intensity below 50%.

The relationship between each health output/outcome and the anti-corruption program is described by the following equation:

\[
y_{mj}^{k} = \alpha + \theta_{m} + \theta_{t} + \beta_{1}^{k} \text{Post} \times \text{ProcIntensive}_{mit} + \gamma^{k} \text{ProcIntensive}_{mit} + \\
\beta^{k}_{2} \text{NearbyAudits}_{mt-1} \times \text{ProcIntensive}_{mit} + \delta^{k} \text{NearbyAudits}_{mt-1} + \epsilon_{mit} \tag{6}
\]

In equation (6), as before, \(\text{Post}_{t} = 1\) from 2003 on, and 0 otherwise; \(\text{ProcIntensive}_{mj}\) equals 1 for procurement-intensive transfers, and 0 otherwise; \(\text{NearbyAudits}_{mt-1} = 1\) at year \(t\) if there was an audit within \(R\) km of municipality \(m\)’s centroid in the previous year (including, possibly, \(m\) itself facing an audit), and 0 otherwise, where \(R\) is the maximum distance for which nearby audits are still estimated to decrease subsequent corruption; \(\theta_{m}\) are municipal fixed-effects; \(\theta_{t}\) are year fixed-effects; and \(\epsilon_{mt}\) is a zero-mean error term. \(Y_{mj}^{k}\) is health output/outcome \(k\) linked to set of transfers \(j\) in municipality \(m\) at year \(t\), where \(j\) represents either high or low procurement-intensity transfers.

Since there are multiple outputs and outcomes within each set of transfers, there is no obvious way of pairing indicators across high and low procurement-intensity transfers in the analysis, and estimating separate regressions for each pair would substantially inflate the probability of false positives above stated significance levels. To deal with this issue, we convert all outputs and outcomes to z-scores, and define

---

27 See Table A4 in Appendix A.
summary measures as the average of z-scores within each set of transfers. Following Kling, Liebman and Katz (2007), effect sizes of the overall effect of the program are obtained by replacing $Y_{mjt}^k$ in equation (6) by the summary measure of high and low procurement-intensity transfers in each municipality and year.\(^{28}\)

$$Z_{mjt} = \alpha + \theta_m + \theta_t + \beta_1 Post_t \times ProcIntensive_{mit} + \gammaProcIntensive_{mit} + \beta_2 NearbyAudits_{m,t-1} \times ProcIntensive_{mit} + \delta NearbyAudits_{m,t-1} + \epsilon_{mit}$$ (7)

In equation (7), $Z_{mjt}$ is the summary measure for health outputs and outcomes linked to set of transfers $j$ in municipality $m$ at year $t$. Under the assumption of identical potential outcomes between high and low procurement-intensity transfers, $\beta$ identifies the causal effect of the anti-corruption program on corruption. Because our corruption measure is computed as the share of investigations coded as corruption in each municipality-year, all estimates are from Generalized Least Squares (GLS) regressions using the number of investigations in municipality $m$ and year $t$ as weights, clustering standard errors at the municipality level.\(^{29}\) The prediction we are interested in is $\beta_1, \beta_2 \leq 0$. $\beta_1, \beta_2 > 0$ implies that exposing corruption mainly reallocates resources towards public goods’ provision, whereas $\beta_1, \beta_2 < 0$ implies that exposing corruption mainly reduces incentives for bureaucratic performance.

The previous strategy has an important caveat: it does not allow for including health outcomes such as malnutrition and mortality counts, which are affected by all health transfers. To circumvent this limitation, we consider two additional strategies. First, we compare two mortality indicators: preventable deaths, and deaths by external causes. Because only the former should be sensitive to municipal health services, we assign it as an outcome of procurement-intensive transfers, and rely on death by external causes as the counterfactual outcome for the comparison category. As long as the two mortality indicators would have followed parallel trends on the absence of the program, the differences-in-differences estimate captures the causal effects of the program on mortality rates.

Second, we compare immunization coverage, incidence of malnutrition amongst under-1-year-old children, incidence of diarrhea amongst under-2-year-old children, and infant mortality across municipalities that should have been differentially exposed to the effects of the program. The idea is similar to that of procurement intensity: where corruption prevalence was higher before the program, there was more scope for reducing corruption afterwards. The Supplementary Appendix shows that, in fact,

\(^{28}\) While Kling, Liebman and Katz (2007) show that the procedure generates effect sizes with accurate confidence intervals for binary treatments, it is straightforward to see that the algebra generalizes to the differences-in-differences estimator.

\(^{29}\) The robustness of our results to estimating (7) using OLS is discussed in subsection 6.4.
corruption was reduced significantly more after the onset of the program among municipalities with above-
median baseline corruption prevalence. As long as, for each outcome, below- and above-median baseline
corruption prevalence municipalities would have followed parallel trends on the absence of the program,
the differences-in-differences estimate captures the causal effects of the program on health outcomes.

6.3 Results

We start by considering the yearly average of the health indicators’ summary measure for each set of
transfers, from 1997 onwards, displayed in Panel A of Figure 5. If the summary measure for low
procurement-intensity transfers was about 0.4 standard deviation lower than that of procurement-intensive
transfers over the baseline period, after 2003 the former systematically increases while the latter plateaus.
By 2007, in fact, the baseline difference between summary measures of different sets of transfers was
basically flipped.

[Figure 5 – Panel A]

Moving on to the regression framework, we exploit within-municipality differences over time across
the two sets of transfers. Effect sizes are presented in Table 5. Columns (1) and (2) restrict attention to the
2001-2004 term, over which we can estimate the effects of the program while holding decision-makers
fixed. Column (1) controls for a post-program indicator variable, while column (2) includes year fixed-
effects. In both cases, the program has a negative point estimate on the summary measure of health outputs
and outcomes, of the order of 0.1 standard deviation, although not statistically significant at the 10% level.

[Table 5]

Column (3) includes the previous and subsequent political terms in the analysis, encompassing the
universe of investigations for which we have coded procurement intensity. Effect sizes increase
substantially (to almost 0.3 standard deviation) relative to the estimates for the 2001-04 sample, and become
statistically significant at the 1% level. The magnitude of the effects in columns (1) to (3) is equivalent to
losing between two-thirds and all the support of federal transfers towards the municipality’s health budget
(in a cross-sectional comparison at the 2001-02 period, within low procurement-intensity transfers). This
benchmark is consistent with the magnitude of the effects on spending documented in the previous section.
Column (4) includes a second-half of the term indicator, allowing it to differentially affect procurement-intensive transfers. While proximity to elections also significantly deteriorates health outputs and outcomes relative to low procurement-intensity transfers (an effect size of similar magnitude to that of the program), the effect of the program on corruption remains significant at the 1% level and remarkably close to column (3)'s estimate.

Finally, columns (5) and (6) include the indicator of recent nearby audits and its interaction with procurement-intensity. We find that recent nearby audits have a negative and significant (at the 10% level) effect on health outputs and outcomes when we consider the full sample, between 1997 and 2007.

Taken together, these are striking results. They suggest that health indicators became systematically worse not because of a short-term disruptive effect of the anti-corruption program, followed by improvements once corruption has been brought down; instead, negative effects become larger and more precise once we include the subsequent term in the analysis.

Because the direction of the effects of recent nearby audits agrees with the differences-in-differences estimate, we can credibly rule out that results are driven by alternative explanations; in particular, by differential pre-existing trends in health indicators linked to different sets of transfers (as Panel A of Figure 5 might suggest). Last, about 75% of the overall effect of the program on health outputs and outcomes comes from its announcement, consistent with the effects of the program on corruption documented in the previous sections.

Next, we consider the comparison between preventable deaths and deaths by external causes. Panel B of Figure 5 shows that both indicators followed similar pre-trends within the 2001-2004 and that, after the program, preventable deaths remain basically still while external deaths fall, most pronouncedly after 2005.

[Figure 5 – Panel B]

Table 6 confirms that pattern, estimating a positive effect of the program on death rates. The coefficient, statistically significant at the 10% level for the overall effect of the program within the term, is not precisely estimated when we control for the second-half of the term indicator, but remains constant steady at around 7% of the baseline average mortality rate. What is more, nearby audits have a positive and statistically significant (at the 5% level) effect of similar magnitude on preventable death rates, relative to deaths from external causes. Since those are randomly assigned conditional on municipality fixed-effects and year fixed-effects, results corroborate our previous findings that the anti-corruption program deteriorated health outcomes.
Last, we consider the alternative counterfactual for the effects of the program on health indicators, comparing municipalities above- and below-median baseline corruption prevalence. Panel C of Figure 5 displays how infant mortality evolved for the two subsets of municipalities, before and after the onset of the program. It shows that, within the 2001-2004 term, while higher infant mortality among above-median municipalities seemed to be converging to that of the other group before the program, it drifts away from 2003 onwards, and that difference remains in the years that follow.

Table 7 presents the results for the immunization, incidence of malnutrition and diarrhea, and infant mortality. Two features of the table are striking. First, although the effect of the program on infant mortality is not precisely estimated, it is very large in magnitude, of about 12% its average at the baseline period. Second, the fact that no indicators are estimated to improve as a result of the program – in particular, point estimates of the effects of the program on malnutrition and diarrhea are very close to zero – despite decreasing corruption by over 38 percentage points across groups, is evidence against the hypotheses that monitoring primarily redirected resources towards public service delivery.

Altogether, results are consistent with our findings for public spending and implementation quality. With 50% to 66% lower budget execution within earmarked transfers within health, more precarious facilities (possibly fewer hospital beds) and with missing medication (possibly fewer vaccines), it is not surprising that health indicators linked to procurement-intensive transfers deteriorate after the program.

6.4 Robustness checks
The Supplementary Appendix presents results without weighting observations by the number of investigations used to compute the share of investigations coded as corruption in each municipality and year. While our preferred specification is the one which weights municipalities by the same criteria we use for computing municipal-level measures of corruption, it is not unreasonable to give all municipalities equal weight since reduced-form estimates do not rely on corruption directly.
Results are similar and, if anything, stronger for the differences-in-differences estimate, which become about 50% larger and statistically significant at the 5% or 1% level in all columns. However, the effect of recent nearby audits is not as robust. While those still have a negative effect on health indicators for the 1997-2007 sample, this effect size is no longer significant.

6.5 ‘Greasing the wheels’ or procurement risk?

We have shown that exposing corruption has made outcomes worse. This effect could be explained by two very different mechanisms, as the model in Section 3 makes it clear. On the one hand, for transfers in which bureaucrats would be corrupt, net expected benefits are lower under a higher audit probability, making procurement undesirable for some transfers that would lead to procurement otherwise. In this case, corruption “greases the wheels” of public service delivery (Banerjee, Hanna and Mullainathan, 2013): bureaucrats are only willing to pay the effort costs of procurement when they can benefit to a greater extent from embezzling resources.

On the other hand, for transfers in which bureaucrats would not be corrupt, net expected benefits are also lower under a higher audit probability, due to procurement risk arising from inaccurate punishment. As a result, procurement becomes undesirable for some transfers that would lead to procurement otherwise. Even though such mechanism is not commonly discussed in the Political Economy literature, it is the leading cause raised by public managers for explaining low budget execution rates in focus groups.

Which mechanism is the most important? As the model shows, variation in $p$, the probability of being exposed as corrupt, is not enough to disentangle the two mechanisms. While it is likely that both play a role, documenting their relative contribution to why audits backfire would be an important direction for future work.

6.6 Relation to the literature

While corruption is conjectured to have high social costs, from static resource misallocation to inefficient investment in factors of production for which returns are seized (Rose-Ackerman, 1997), there is limited evidence of its effects (Banerjee, Hanna and Mullainathan, 2013; Pande, 2007). On the one hand, corruption has been documented to pose a major obstacle to the decentralization of public service provision in developing countries, with embezzlement levels sometimes higher than the amount that actually reaches targeted individuals (Olken, 2006), and sometimes even reversing the progressivity of public expenditures.

Ferraz, Finan and Moreira (2012), Méon and Weil (2010), and Méon and Sekkat (2005) present cross-sectional evidence on the relationship between corruption and efficiency.
(Reinikka and Svensson, 2004). On the other hand, whether it actually induces inefficient outcomes or rather represents transfers to bureaucrats, which ‘grease the wheels’ of public service delivery, remains largely an open question.31

There are at least two reasons why exposing corruption might be detrimental to efficiency within a bureaucratic politics’ setting. First, under limited liability, the optimal principal-agent contract must leave rents to the agent (Bolton and Dewatripont, 2004). If such rents are reduced, then incentives must become less powerful (so that the participation constraint is still satisfied). Second, because it is hard to separate corruption from discretion (Bandiera, Pratt and Valla, 2009; Huntington, 1968; Leys, 1965; Leff, 1964), discouraging the former often discourages the latter. This paper provides first-hand evidence for these two mechanisms. After the introduction of the anti-corruption program, we observe both plummeting public spending and less reshuffling of resources across different transfers.

Analyzing the causal effects of corruption is challenging. Since it is hard to obtain experimental variation in corruption, researchers have resorted to two alternative approaches. First, mechanism experiments (Ludwig, Kling and Mullainathan, 2012) that identify the effects of corruption in narrow settings in which it can be experimentally varied (e.g., Zamboni and Litschig, 2013; Bertrand et al., 2007). The advantage of this approach is control: the variation is randomly assigned and tightly linked to the mechanism of interest. Its disadvantage is external validity: it is unclear to what extent results in these narrow settings would carry over to the effects of corruption in society at large. Second, natural experiments that exploit policy changes (e.g., Reinikka and Svensson, 2005; DiTella and Schargrodsky, 2003). The advantage of this approach is external validity: this is exactly the variation that one is interested in when thinking about the effects of corruption. Its disadvantages are manifold. In particular, anti-corruption policies usually affect everyone, such that counterfactual analysis must rely on strong assumptions. Furthermore, because corruption is observed only after the program kicks-off, it is often inferred from other variables (such as prices) that include variation unrelated to the mechanism of interest.

This paper exploits variation that is externally valid and that allows for counterfactual analysis under weaker assumptions than previous papers. By leveraging on a national anti-corruption program, we take advantage of a shock that affects corruption in society at large. Since the program is based on random audits, we can estimate its causal effects on corruption prevalence and on public service delivery. That allows us to go beyond the findings of Olken (2007), Bobonis, Fuertes and Schwabe (2015) and Avis, Ferraz and Finan (2017), which document the effects of monitoring on corruption but do not analyze its effects on the quality of public goods.

31 For a theoretical perspective, see Shleifer and Vishny (1993) and Banerjee, Hanna and Mullainathan (2012).
Ferraz and Finan (2008) exploit variation induced by the Brazilian anti-corruption program to study whether voters punish corrupt politicians. By comparing municipalities equally corrupt – according to audit reports – that had reports publicized by the media right before or right after elections (only because of the randomness of the program’s lotteries), they can estimate the effect of information on reelection rates. Ferraz and Finan (2011) rely on the same program to answer a different question: can electoral institutions reduce corruption? Through a regression discontinuity design, they document that second-term mayors are corrupt to a greater extent than first-term mayors. Ferraz, Finan and Moreira (2012) analyze the cross-sectional correlation between the incidence of corruption measured by the program’s audit reports and educational outcomes, finding a negative association between missing government funds and students’ achievement. None of these papers estimate the effect of the program on corruption.

Avis, Ferraz and Finan (2017) estimate the effect of the program on corruption by relying on nearby audits, and find similar results to ours. Besides separating the effects of the program’s announcement from that of the actual audits, our contribution is to estimate the effect of the program on other margins of bureaucratic decisions, and on downstream outcomes that we ultimately care about.

Our findings are at odds with those of Reinikka and Svensson (2005), which document sizable positive effects of a newspaper campaign in Uganda that disclosed official figures about the central government's transfers to local primary schools. One possible explanation for this difference is that while Reinikka and Svensson (2005) consider a single transfer that is delivered directly to schools, this paper considers the universe of earmarked federal transfers to municipalities. Before benefitting citizens, such funds must be used by local bureaucrats to procure goods and services, subject to a complex set of procurement guidelines. This is the typical budget implementation process in developing countries, and the effects we find are consistent with the effects of bureaucratic rigidity (Bertrand et al., 2016), who also find negative effects of distorted incentives on downstream outcomes – quite substantial in terms of GDP growth – and that of monitoring bureaucratic performance documented elsewhere (Rasul and Rogger, 2015; Shin, 2008).

7 Discussion and concluding remarks

In this paper, we have documented that the Brazilian anti-corruption program significantly decreased corruption, but made health indicators significantly worse. In response to the introduction of the program, local procurement staff seem to have reduced spending within constitutionally mandated transfers by at least 50%. As a result, infrastructure and medication suffered, bureaucrats felt more constrained about channeling resources towards local needs, and the quality of health services significantly deteriorated.
These negative effects do not seem to result from temporary adjustments. Not only do effects persist in the subsequent political term; they also become stronger and more precisely estimated. Moreover, it is not the case that unspent resources are reallocated to other local priorities (that would be coded as resource diversion) or to other municipalities (constitutionally mandated transfers cannot be returned to the federal government). While there are no official data for budget leftovers with respect to these transfers, evidence for other federal health programs which rely on decentralized budget execution is consistent with extremely low execution rates. All in all, our findings suggest that extreme views such as that held by Jim Kim, president of the World Bank, who calls corruption “developing world’s public enemy number 1”, should be taken with a grain of salt.

Our results provide first-hand evidence that exposing corruption may hurt public service delivery. Does this mean corruption is welfare-improving in this setting? On the one hand, it may be that local bureaucracies have more accurate information about the local demand for public goods or about the quality of local suppliers, along the lines of Huntington (1968). In this case, being able to deviate from transfers’ guidelines by diverting resources or targeting procurement processes to specific vendors – all coded by auditors as evidence of corruption – would improve welfare. On the other hand, it may be that bureaucrats are motivated by other objectives, such as extracting personal rents from vendors, or minimizing implementation effort. In that case, corruption would not necessarily improve welfare. Better public services would come at the cost of redistribution towards bureaucrats, which can generate a broad set of distortions (from vendors’ incentives to outbid other bribe offers to individuals’ perceptions about trustworthiness in society at large) that might out-weight its benefits. Our results for implementation quality provide preliminary evidence that both mechanisms might be at play in this setting.

While the Brazilian anti-corruption program represents a major improvement in monitoring and transparency, the focus of administrative penalties and of public opinion on corruption, instead of on the quality of public services, all seem to have thrown the baby out with the bathwater. These findings suggest that policies that expand the scope of desirable outcomes beyond formal procedures, that differentiate between active and passive waste, and that support local procurement staff in complying with complex guidelines might be important steps towards balancing incentives between procuring and making proper use of public funds.

Some recent advances have moved the country in that direction. In particular, electronic procurement has streamlined acquisition of homogenous goods and services nationwide, and some municipalities have devised centralized procurement agencies to concentrate their best human resources across different agencies. Progress in other critical dimensions, however, does not show the same promise. Recurrent corruption scandals in national politics have inflamed public opinion, with the logic of “crime and
punishment” prevailing over that of implementation quality. Moreover, the Brazilian procurement framework – organized around Law 8.666 – has not undergone substantive improvements since 1993, despite being systematically criticized by its complex and restrictive guidelines, in particular in what comes to the complex procedures for procuring works and consulting services. Finally, there has not been significant progress in expanding the scope of irregularities for which administrative penalties can be applied since 2007; Draft Law 931 still waits to be voted upon Congress.

Despite being the highest per capita GDP country in Latin America, Brazil fares poorly in health indicators. In 2010, according to the World Health Organization, Brazil only placed 10th out of 20 Latin American countries with respect to infant mortality, and 13th with respect to life expectancy at birth. Given that the decentralization of federal spending, from 1996 on, increased municipal health budgets seven-fold (from 0.1% of GDP in 1995 to 0.72% in 2010), we believe this paper sheds light on some critical constraints to bringing the country’s health outcomes more in line with its development stage.

Our findings for the effects of the program on corruption and public spending are consistent with a model in which local procurement staff decides whether or not to procure, and then, conditionally upon procuring, decides whether or not to be corrupt. Such a model has at least two normative implications. First, audits should focus on investigations involving larger amounts (with the largest embezzlement opportunities), since those involve the bulk of embezzlement, and since, within those transfers, audits discourage procurement by the least. Second, supporting local procurement staff, building capacity to decrease the probability of accidental procurement mistakes, may both decrease corruption (passive waste) and increase public spending (and hence downstream outcomes).

Interestingly, in line with the first normative implication, in 2014 and 2015 the anti-corruption program has restricted audits to larger municipalities (for which transfers involve larger amounts). The reason, however, was of a different nature: given the fragile fiscal situation of the federal government, CGU’s budget only allowed investigating municipalities closer to the state capital, where auditors are based. CGU has indicated it would like to return to its all-encompassing strategy once the fiscal situation is alleviated.

The second normative implication seems to have been taken on board by CGU. In 2006, it launched the Public Management Strengthening (“Fortalecimento da Gestão Pública”, FGP) program, which randomly draws municipalities for capacity building, particularly with respect to procurement guidelines – provided through in-site training over the course of a week, combined with complementary online materials. In fact, Lopes (2011) provides preliminary evidence that FGP may decrease corruption within health transfers. Although effect sizes are small and not systematically estimated with precision (perhaps not surprisingly, given the short duration of the training program), these results suggest that in-site capacity-building programs could be a promising way forward.
All in all, the optimal design of capacity-building interventions to disseminate best practices among local procurement staff, and the extent to which those interventions can improve public service delivery, are promising avenues for future research.
REFERENCES


BROLLO, F. (2010) “Who is punishing corrupt politicians: voters or the central government? Evidence from the Brazilian anti-corruption program,” Bocconi University, mimeo.


Appendix A – Classification lists

Table A1 – List of irregularities

<table>
<thead>
<tr>
<th>Panel A: Corruption</th>
<th>Category</th>
<th>Irregularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procurement</td>
<td>Irregular receipts</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Evidence for ghost firms</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Contracts not signed or falsified signatures</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Favored vendor</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Lack of publicity</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Documents set with different dates</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Other procurement problems</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>Irregular class</td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td>No realization</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Overinvoicing</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Off-the-record payments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Mismanagement</th>
<th>Category</th>
<th>Irregularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource diversion</td>
<td>Unconfirmed payments</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Diversion of resources for other goals</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Diversion of resources for other goals within Health</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Diversion of resources for other goals within Program</td>
<td></td>
</tr>
<tr>
<td>Resource diversion</td>
<td>Under-application of resources</td>
<td></td>
</tr>
<tr>
<td>Health council</td>
<td>Irregular Composition</td>
<td></td>
</tr>
<tr>
<td>Health council</td>
<td>Irregular Operation</td>
<td></td>
</tr>
<tr>
<td>Health council</td>
<td>Poor infrastructure and work conditions</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Unaccomplished goals</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Unfinished projects</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Poorly evaluated services to health system users</td>
<td></td>
</tr>
<tr>
<td>Supplies and facilities</td>
<td>Precarious facilities</td>
<td></td>
</tr>
<tr>
<td>Supplies and facilities</td>
<td>Signs and logos not properly set</td>
<td></td>
</tr>
<tr>
<td>Supplies and facilities</td>
<td>Lack of medical supplies</td>
<td></td>
</tr>
<tr>
<td>Supplies and facilities</td>
<td>Stock control of medication</td>
<td></td>
</tr>
<tr>
<td>Supplies and facilities</td>
<td>Maintenance of medication</td>
<td></td>
</tr>
<tr>
<td>Human Resources</td>
<td>Professionals that don't fulfill work time requirements</td>
<td></td>
</tr>
<tr>
<td>Human Resources</td>
<td>Staff training</td>
<td></td>
</tr>
<tr>
<td>Human Resources</td>
<td>Staff composition</td>
<td></td>
</tr>
<tr>
<td>Human Resources</td>
<td>Public servants’ payments</td>
<td></td>
</tr>
<tr>
<td>Documentation/Accounting</td>
<td>Incomplete documentation or inadequate account keeping</td>
<td></td>
</tr>
</tbody>
</table>
### Table A2 – Procurement-intensity by program

<table>
<thead>
<tr>
<th>Health Ministry Code</th>
<th>Health Program</th>
<th>% of action coded as procurement-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>0119</td>
<td>Sanitation (Saneamento básico)</td>
<td>100.00</td>
</tr>
<tr>
<td>0004</td>
<td>Quality and efficiency of the Unified Health System (Qualidade e eficiência do SUS)</td>
<td>54.17</td>
</tr>
<tr>
<td>0005</td>
<td>Pharmaceutical Assistance (Assistência Farmacêutica)</td>
<td>50.00</td>
</tr>
<tr>
<td>0013</td>
<td>Epidemiological and environmental surveillance in Health (Vigilância epidemiológica e ambiental em Saúde)</td>
<td>50.00</td>
</tr>
<tr>
<td>0002</td>
<td>Prevention and control of vector-transmitted diseases (Prevenção e controle de doenças transmitidas por vetores)</td>
<td>38.46</td>
</tr>
<tr>
<td>1214</td>
<td>Basic Attention in Health (Atenção básica em Saúde)</td>
<td>4.55</td>
</tr>
<tr>
<td>0001</td>
<td>Family Health (Saúde da família)</td>
<td>0.00</td>
</tr>
<tr>
<td>0023</td>
<td>Admission, emergency and hospital services (Atendimento ambulatorial, emergencial e hospitalar)</td>
<td>0.00</td>
</tr>
<tr>
<td>1335</td>
<td>Conditional Cash Transfer (Transferência de renda com condičionalidades)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### Table A3 – List of procurement-related words

- “Insumos” (Inputs)
- “Implantação” (Adoption of)
- “Modernização” (Modernization of)
- “Adequação” (Tailoring of)
- “Aparelhamento” (Related to equipment/infrastructure)
- “Ampliação” (Enlargement)
- “Manutenção” (Maintenance)
- “Construção” (Construction)
- “Produção” (Production)
- “Aquisição” (Acquisition)
- “Estruturação” (Structuring)

### Notes to Table A3:

1. The procurement-related words listed are used to classify each action under each health transfer as procurement-related or not, based on the description of each action in Health Ministry’s *Funcional Programática*;
2. The classification of all actions under the transfers we analyze in this paper is presented in the Supplementary Appendix.
Table A4 – Health Ministry’s M&E indicators by transfer

<table>
<thead>
<tr>
<th>Health Ministry Code</th>
<th>Health Program</th>
<th>Outputs/Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0119</td>
<td>Sanitation</td>
<td>(i) Household share connected to the water network; (ii) Household share connected to the sewage network.</td>
</tr>
<tr>
<td>0004</td>
<td>Quality and efficiency of the Unified Health System</td>
<td>(i) Hospital beds per thousand inhabitants [imputed].</td>
</tr>
<tr>
<td>0005</td>
<td>Pharmaceutical Assistance</td>
<td>(i) Distribution of specific medication [unavailable].</td>
</tr>
<tr>
<td>0013</td>
<td>Epidemiological and environmental surveillance in Health</td>
<td>(i) Immunization shots per thousand inhabitants [imputed]. (ii) Incidence of HIV, tuberculosis, and leprosy [unavailable]; (iii) Population share covered by dental health teams [unavailable]; (iv) Population share covered by family health teams.</td>
</tr>
<tr>
<td>0002</td>
<td>Prevention and control of vector-transmitted diseases</td>
<td>(ii) Population share covered by family health teams. (i) Medical consultations per thousand inhabitants; (ii) Hospital admissions per thousand inhabitants; (iii) Transplantations per thousand inhabitants [unavailable].</td>
</tr>
<tr>
<td>1214</td>
<td>Basic Attention in Health</td>
<td></td>
</tr>
<tr>
<td>0001</td>
<td>Family Health</td>
<td></td>
</tr>
<tr>
<td>0023</td>
<td>Admission, emergency and hospital services</td>
<td></td>
</tr>
<tr>
<td>1335</td>
<td>Conditional Cash Transfer</td>
<td></td>
</tr>
</tbody>
</table>

Notes to Table A4:

2. In the case of the two programs for which the Health Ministry does not list M&E indicators, we impute outputs/outcomes based on the description of all actions under each transfer in Health Ministry’s Funcional Programática. *Quality and efficiency of the Unified Health System* includes action 0004.1823 (“aparelhamento de unidades de saúde do SUS”, *equipment/infrastructure for public health units*), which we proxy with hospital beds per thousand inhabitants;
3. *Epidemiological and environmental surveillance in Health* includes action 0013.3994 (“modernização do Sistema Nacional de Vigilância em Saúde – VIGISUS”[^1][32], *modernization of the national health surveillance system*), which we proxy with immunization;
4. *Unavailable* indicators have no annual municipal-level data available in DATASUS.

Appendix B – More details about the Brazilian anti-corruption program

Created in February 2001, CGU is in charge of oversight and fraud detection in every issue related to federal public funds, and it is also responsible for developing mechanisms to prevent corruption. The Brazilian anti-corruption program is a federal government’s initiative to inhibit corruption across all levels of the public administration.33 Ferraz and Finan (2008) perform a thorough analysis of the actual randomness of the program.

The anti-corruption program began with a sample of five municipalities in the first draw. The second draw included 25 municipalities. From the third to the twelfth draws, 50 municipalities were audited, and from the thirteenth on, 60 municipalities are now drawn. The distribution of draws over time and the number of municipalities drawn in each lottery are presented on the CGU website. Although states have also been audited under the program from 2004 on, the focus of the program is on municipalities. The program currently audits municipalities up to 500,000 inhabitants. Maximum population thresholds have increased over time: from 100,000 to 250,000 inhabitants at the third draw; to 300,000 inhabitants between the fourth to the eighth; and, finally, to 500,000 inhabitants for the ninth draw and after. Some draws also had a minimum population threshold of 10,000 inhabitants. The sampling procedure was designed so that the drawn samples are geographically representative, and selection probabilities currently approximate 1% for each of the 5,526 municipalities (which represent over 99% of Brazilian municipalities or about 70% of the country's population) which currently lie within the maximum population eligibility thresholds.

The program investigates earmarked transfers linked to national health and education policies (constitutionally mandated transfers), direct transfers to citizens, and politically negotiated (voluntary) transfers. Once a municipality has been randomly drawn, a service order is issued by CGU, indicating the set of transfers to be audited in each municipality. According to CGU officials interviewed by the authors, service orders encompass every transfer currently at a “critical stage in its life-cycle”: (i) disbursement, (ii) procurement, or (iii) approval of previous records by municipal councils. Once service orders have been issued, a team of auditors visits the municipality to investigate irregularities.

For both education and health, all federal transfers are subject to auditing in every municipality drawn, discarding selection concerns with respect to the issuance of service orders. Auditors are then entitled to inspect the complete paper trail of the audited transfers, from the National Treasury's account to its current stage under the municipality's discretion, whether in previous years or under a previous political ruling.

33 For a full description, see http://www.cgu.gov.br/AreaAuditoriaFiscalizacao/ExecucaoProgramasGoverno/Sorteios/index.asp.
Once the audit is complete, CGU officials describe all irregularities detected for each transfer (if any) in official reports.\textsuperscript{34}

Irregularities documented by auditors are followed-up by the public entities responsible for implementing sanctions, including the Prosecutor's Office, the Brazilian Court of Audits ("Tribunal de Contas da União", TCU), the Federal Police, and municipal legislative houses.

\textsuperscript{34} Mayors can challenge the conclusions of such reports; when the CGU takes such claims into consideration, a CGU team comes back to the municipality to reassess prior analysis until a final report can be issued.
Appendix C – Examples from audit reports

Examples of pre-program findings (1997-2002)

COMPLIANCE

When evaluating the implementation of a public program for treatment of Hansen’s disease (leprosy) in 2001-02 in Alvaraes, State of Amazonas, auditors found no evidence of irregularity. The local health unit, maintained by the Unified Health System (SUS), operates in compliance with legislation, keeping patients’ records and providing vaccination, counseling, and staff training. We code this finding as a case of compliance (in the biennium 2001/2002). The municipality was drawn by lottery 2.

MISMANAGEMENT

When evaluating the municipal inventory of medical supplies in 2002, auditors found no formal controls to be in place concerning stock inflows or outflows. As a consequence, there was no way to determine either the quantity of stocked medication or the quantity delivered to each local health unit. We code this irregularity as mismanagement related to medication stock control. This irregularity occurred in 2002, in Amajari, State of Roraima, municipality drawn by lottery 2.

CORRUPTION

When evaluating the procurement process to purchase two mobile health units (modified vehicles that operate as small health units), in 2002 by the municipal government of Dourados, auditors found evidence of fraud. The only public outbidder, Santa Maria Comércio e Representações Ltda., does not legally exist according to the local branch of the Federal Revenue Secretariat in Cuiabá, State of Mato Grosso do Sul. Despite this fact, the municipal government has concluded the public bid and paid the company the agreed-upon amount. Unfortunately, auditors could not find further evidence of what may have encouraged the fraud. We code this irregularity as an evidence of ghost firm, a clear indication of corruption. This irregularity occurred in 2002, in Dourados, Mato Grosso do Sul, drawn by lottery 4.

Examples of post-program findings (2003-2007)

COMPLIANCE

When evaluating the process of purchasing, stocking, and supplying medical supplies to local health units, auditors found no evidence of irregularity related to the municipal government in Teresina de Goiás, Goiás. The only such case was due to a delay caused by the state government, which failed to transfer
resources in due time. We code this finding as a case of compliance (by the municipal government, in 2007). The municipality was drawn in the 23rd round.

MISMANAGEMENT

When evaluating the medication accessibility to citizens of Londrina in 2006, auditors found out that several prescribed drugs in local health units were out of stock. Evidence was provided by interviews with patients and local employees (nurses), who reported lack of, for example, Clorana, Dipirona, Enalapril, Captopril, Cloritisona, Nifedipina, and Istamin. We code this irregularity as lack of medical supplies, an evidence of mismanagement. This irregularity occurred in 2006, in Londrina, Parana, drawn by lottery 21.

CORRUPTION

When analyzing a procurement process to purchase medical supplies in 2004, auditors found that the municipal government of Poloni had paid higher prices for medication than the one agreed upon the public-bid contract. For example, according to receipt number 115655 (Procurement number 2004/01696), the correct price of 150 mg of the medication Ranitidine was R$ 0.18 per tablet, but the municipality paid R$ 0.28 per tablet. No further documentation was presented by the municipal government, and the outbidder Empresa Soquímica Laboratórios Ltda., embezzled the resources. We code this irregularity as overinvoicing, an evidence of corruption. This irregularity occurred in 2004, in Poloni, São Paulo, drawn by lottery 17.
Appendix D – Figures

Figure 1 – Share of investigations coded as corruption by year and procurement-intensity

Notes on Figure 1:

1. Weighted average of the share of investigations coded as corruption, by year and by the high procurement-intensity indicator, equal to 1 if procurement-intensity is 50% or higher, and 0 otherwise;
2. Number of investigations at each municipality and year used as weights;
3. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A.
4. Solid lines indicate term limits.
Notes on Figure 2:

1. Regression coefficients of interactions of year indicators with the high procurement-intensity indicator, equal to 1 if procurement-intensity is 50% or higher, and 0 otherwise, on the share of investigations coded as corruption;
2. Ordinary Least Squares (OLS) regression including municipality fixed-effects and year fixed-effects;
3. Dashed lines represent 95% confidence intervals, computed from robust standard errors clustered at the municipal level;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A.
Figure 3 – Differences-in-differences coefficient by draw

Notes on Figure 3:

1. Regression coefficients of interactions of draw indicators with the high procurement-intensity indicator, equal to 1 if procurement-intensity is 50% or higher, and 0 otherwise, on the share of investigations coded as corruption;
2. Ordinary Least Squares (OLS) regression including municipality fixed-effects and year fixed-effects;
3. Dashed lines represent 95% confidence intervals, computed from robust standard errors clustered at the municipal level;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A.
Figure 4 – Average public spending (proxied by audited amounts) by year and procurement-intensity

Notes on Figure 4:

1. Weighted average of audited amount per investigation, by year and by the high procurement-intensity indicator, equal to 1 if procurement-intensity is 50% or higher, and 0 otherwise;
2. Number of investigations at each municipality and year used as weights;
3. There is no data on budget execution for constitutionally mandated health transfers. For this reason, we resort to audited amounts. The idea is that, beyond compliance with documentation and account keeping, with respect to which all transfers are always audited on, only spending moves a transfer up in its life-cycle and generates further objects of investigations by the anti-corruption program;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A.
5. Solid lines indicate term limits.
Figure 5 – Average health outputs’ and outcomes’ by year and procurement-intensity

Panel A

[Graph showing health outputs and outcomes by year and procurement-intensity.]

Panel B

[Graph showing death rates by year and type of cause.]

Pre-program

Post-program
Notes on Figure 5:

1. The y-axis stand for: (i) in Panel A, the weighted average of summary measures for health outputs and outcomes, by year and by the high procurement-intensity indicator (equal to 1 if procurement-intensity is 50% or higher, and 0 otherwise); (ii) in Panel B, the weighted average of mortality rates by cause of death – external causes matched to low-procurement intensity indicator and general causes (all but external) matched to the high-procurement intensity indicator; and (iii) in Panel C, weighted average of infant mortality rates by baseline corruption prevalence (equal to 1 if above the 1997-2002 median corruption prevalence, and 0 otherwise);
2. Number of investigations at each municipality and year used as weights;
3. In Panel A, the summary measure for high procurement-intensity transfers includes hospital beds per thousand inhabitants, immunization, the household share connected to the general sewage network, and the household share with connected to piped water. The summary measure for low procurement-intensity transfers includes medical consultations and the population share covered by the Family Health program;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A.
5. Solid lines indicate term limits.
### Appendix E – Tables

#### Table 1 - Differences-in-differences estimates of the anti-corruption program on corruption

<table>
<thead>
<tr>
<th></th>
<th>Share of audits coded as corruption</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[2001-04]</td>
<td>[1997-2007]</td>
<td>[2001-04]</td>
<td>[1997-07]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Post-program x</td>
<td>-0.179***</td>
<td>-0.178***</td>
<td>-0.180***</td>
<td>-0.176***</td>
<td>-0.163***</td>
<td>-0.159***</td>
</tr>
<tr>
<td></td>
<td>[0.034]</td>
<td>[0.033]</td>
<td>[0.032]</td>
<td>[0.032]</td>
<td>[0.033]</td>
<td>[0.034]</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-program</td>
<td>-0.094***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.023]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>0.215***</td>
<td>0.214***</td>
<td>0.186***</td>
<td>0.187***</td>
<td>0.215***</td>
<td>0.190***</td>
</tr>
<tr>
<td></td>
<td>[0.034]</td>
<td>[0.034]</td>
<td>[0.032]</td>
<td>[0.032]</td>
<td>[0.031]</td>
<td>[0.031]</td>
</tr>
<tr>
<td>Second-half of term x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.083**</td>
<td>-0.083**</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.031]</td>
<td>[0.034]</td>
</tr>
<tr>
<td>Second-half of term</td>
<td>-0.223***</td>
<td></td>
<td></td>
<td></td>
<td>-0.198*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.052]</td>
<td></td>
<td></td>
<td></td>
<td>[0.115]</td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td></td>
<td></td>
<td></td>
<td>-0.049</td>
<td>-0.046**</td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td></td>
<td></td>
<td></td>
<td>[0.030]</td>
<td>[0.022]</td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>0.067</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.111]</td>
<td>[0.034]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipality fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>7072</td>
<td>7072</td>
<td>10538</td>
<td>10538</td>
<td>7072</td>
<td>10538</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>623</td>
<td>623</td>
<td>728</td>
<td>728</td>
<td>623</td>
<td>728</td>
</tr>
</tbody>
</table>

**Notes on Table 1:**

1. All columns are regressions with dependent variable equal to 1 if the investigation is coded as corruption, and 0 otherwise. See Appendix A for the definition of corruption;
2. Robust standard errors in brackets, clustered at the municipal level, in columns (1)-(4);
3. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A;
5. Coefficients and standard errors in columns (5) and (6) were bootstrapped using the procedure described in the Supplementary Appendix;
6. *** p<0.01, ** p<0.05, * p<0.1.
**Table 2** - Differences-in-differences estimates of the anti-corruption program on public spending (proxied by audited amounts)

<table>
<thead>
<tr>
<th></th>
<th>log(audited amount)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[2001-04]</td>
<td>[1997-2007]</td>
<td>[2001-04]</td>
<td>[1997-07]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Post-program x</td>
<td>-0.496***</td>
<td>-0.490***</td>
<td>-0.711***</td>
<td>-0.677***</td>
<td>-0.372***</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.134]</td>
<td>[0.133]</td>
<td>[0.126]</td>
<td>[0.125]</td>
<td>[0.129]</td>
</tr>
<tr>
<td>Post-program</td>
<td>-0.374***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0965]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>-0.371***</td>
<td>-0.377***</td>
<td>-0.439***</td>
<td>-0.434***</td>
<td>-0.374***</td>
</tr>
<tr>
<td></td>
<td>[0.133]</td>
<td>[0.132]</td>
<td>[0.123]</td>
<td>[0.123]</td>
<td>[0.124]</td>
</tr>
<tr>
<td>Second half of term x</td>
<td>-0.855***</td>
<td>-0.711***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.283]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second half of term</td>
<td>-0.683***</td>
<td>-0.685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.248]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>-0.379**</td>
<td>-0.372**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.197]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>0.257</td>
<td>0.197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipality fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>7072</td>
<td>7072</td>
<td>10538</td>
<td>10538</td>
<td>7072</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>623</td>
<td>623</td>
<td>728</td>
<td>728</td>
<td>623</td>
</tr>
</tbody>
</table>

**Notes on Table 2:**

1. All columns are regressions with investigation’s audited amount (in log) as dependent variable;
2. Columns (1) to (6) are Ordinary Least Squares (OLS) regressions;
3. Robust standard errors in brackets, clustered at the municipal level, in columns (1)-(4);
4. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
5. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A;
6. Coefficients and standard errors in columns (5) and (6) were bootstrapped using the procedure described in the Supplementary Appendix;
7. *** p<0.01, ** p<0.05, * p<0.1.
Table 3 – Differences-in-differences estimates of the anti-corruption program on implementation quality

<table>
<thead>
<tr>
<th></th>
<th>[2001-04]</th>
<th></th>
<th></th>
<th>[1997-2007]</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Post-program x</td>
<td>-0.163***</td>
<td>0.164***</td>
<td>-0.001</td>
<td>-0.163***</td>
<td>0.155***</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>[0.033]</td>
<td>[0.042]</td>
<td>[0.028]</td>
<td>[0.034]</td>
<td>[0.041]</td>
<td>[0.026]</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>0.215***</td>
<td>-0.224***</td>
<td>0.009</td>
<td>0.190***</td>
<td>-0.200***</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>[0.031]</td>
<td>[0.040]</td>
<td>[0.026]</td>
<td>[0.031]</td>
<td>[0.038]</td>
<td>[0.025]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td>-0.049</td>
<td>0.018</td>
<td>0.031</td>
<td>-0.044**</td>
<td>0.027</td>
<td>0.017</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.030]</td>
<td>[0.038]</td>
<td>[0.034]</td>
<td>[0.022]</td>
<td>[0.024]</td>
<td>[0.016]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>0.067</td>
<td>-0.034</td>
<td>-0.033</td>
<td>0.011</td>
<td>-0.009</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>[0.111]</td>
<td>[0.088]</td>
<td>[0.050]</td>
<td>[0.035]</td>
<td>[0.040]</td>
<td>[0.016]</td>
</tr>
<tr>
<td>Municipality fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>7.072</td>
<td>7.072</td>
<td>7.072</td>
<td>10.538</td>
<td>10.538</td>
<td>10.538</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>623</td>
<td>623</td>
<td>623</td>
<td>728</td>
<td>728</td>
<td>728</td>
</tr>
</tbody>
</table>

Notes on Table 3:

1. Columns (1) and (4) are regressions with dependent variable equal to 1 if the investigation is coded as corruption, and 0 otherwise; columns (2) and (5) are regressions with dependent variable equal to 1 if the investigation is coded as mismanagement, and 0 otherwise; and columns (3) and (6) are regressions with dependent variable equal to 1 if the investigation is coded as compliance (neither evidence of corruption, nor of mismanagement), and 0 otherwise. See Appendix A for the definition of corruption and mismanagement;
2. Columns (1) to (6) are Ordinary Least Squares (OLS) regressions;
3. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A;
5. Coefficients and standard errors were bootstrapped using the procedure described in the Supplementary Appendix;
6. *** p<0.01, ** p<0.05, * p<0.1.
## Table 4 - Differences-in-differences estimates of the anti-corruption program within mismanagement (1997-2007)

<table>
<thead>
<tr>
<th></th>
<th>(1) Resource diversion</th>
<th>(2) Health council problems</th>
<th>(3) Performance problems</th>
<th>(4) Infrastr. and stock problems</th>
<th>(5) Human resources problems</th>
<th>(6) Documentation problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-program x</td>
<td>-0.066***</td>
<td>0.032**</td>
<td>-0.068**</td>
<td>0.210***</td>
<td>0.032**</td>
<td>0.066**</td>
</tr>
<tr>
<td></td>
<td>[0.024]</td>
<td>[0.016]</td>
<td>[0.030]</td>
<td>[0.026]</td>
<td>[0.015]</td>
<td>[0.031]</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>-0.006</td>
<td>-0.078***</td>
<td>0.086***</td>
<td>0.010</td>
<td>-0.035**</td>
<td>-0.169***</td>
</tr>
<tr>
<td></td>
<td>[0.022]</td>
<td>[0.015]</td>
<td>[0.027]</td>
<td>[0.023]</td>
<td>[0.014]</td>
<td>[0.030]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td>-0.004</td>
<td>0.007</td>
<td>0.042</td>
<td>0.024</td>
<td>0.010</td>
<td>-0.017</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.029]</td>
<td>[0.012]</td>
<td>[0.051]</td>
<td>[0.036]</td>
<td>[0.032]</td>
<td>[0.028]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>-0.014</td>
<td>-0.002</td>
<td>-0.015</td>
<td>0.018</td>
<td>0.005</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>[0.055]</td>
<td>[0.023]</td>
<td>[0.068]</td>
<td>[0.034]</td>
<td>[0.019]</td>
<td>[0.030]</td>
</tr>
<tr>
<td>Municipality fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>10538</td>
<td>10538</td>
<td>10538</td>
<td>10538</td>
<td>10538</td>
<td>10538</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>728</td>
<td>728</td>
<td>728</td>
<td>728</td>
<td>728</td>
<td>728</td>
</tr>
</tbody>
</table>

**Notes on Table 4:**

1. Columns (1) to (6) are regressions with dependent variable equal to 1 if the investigation is coded as the mismanagement category which labels the column, and 0 otherwise. See Appendix A for the definition of all mismanagement categories;
2. Columns (1) to (6) are Ordinary Least Squares (OLS) regressions;
3. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
4. Coefficients and standard errors were bootstrapped using the procedure described in the Supplementary Appendix;
5. *** p<0.01, ** p<0.05, * p<0.1.
Table 5 - Differences-in-differences estimates of the anti-corruption program on health outputs’ and outcomes’ summary measure

<table>
<thead>
<tr>
<th>Health outputs/outcomes summary measure</th>
<th>[2001-04]</th>
<th></th>
<th>[1997-2007]</th>
<th>[2001-04]</th>
<th>[1997-07]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Post-program x</td>
<td>-0.118</td>
<td>-0.125</td>
<td>-0.288***</td>
<td>-0.279***</td>
<td>-0.132*</td>
</tr>
<tr>
<td></td>
<td>[0.0830]</td>
<td>[0.0825]</td>
<td>[0.0775]</td>
<td>[0.0777]</td>
<td>[0.078]</td>
</tr>
<tr>
<td>Post-program</td>
<td>0.127**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0524]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>0.232***</td>
<td>0.242***</td>
<td>0.234***</td>
<td>0.236***</td>
<td>0.246***</td>
</tr>
<tr>
<td></td>
<td>[0.0876]</td>
<td>[0.0873]</td>
<td>[0.0812]</td>
<td>[0.0814]</td>
<td>[0.074]</td>
</tr>
<tr>
<td>Second half of term x</td>
<td>-0.194*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.115]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>1.024***</td>
<td>1.319**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.219]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td>-0.064</td>
<td></td>
<td>-0.155*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.207]</td>
<td></td>
<td>[0.086]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>0.014</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.208]</td>
<td>[0.069]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipality fixed-effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,874</td>
<td>1,874</td>
<td>2,703</td>
<td>2,703</td>
<td>1,874</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>611</td>
<td>611</td>
<td>725</td>
<td>725</td>
<td>611</td>
</tr>
</tbody>
</table>

Notes on Table 5:
1. All columns are regressions with the summary measure explained below as the dependent variable. Columns (1) to (6) are Generalized Least Squares (GLS) regressions, with the number of investigations in each municipality and year used as weights;
2. The summary measure for high procurement-intensity transfers includes hospital beds per thousand inhabitants, immunization, the household share connected to the general sewage network, and the household share with connected to piped water. The summary measure for low procurement-intensity transfers includes medical consultations and the population share covered by the Family Health program;
3. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A;
5. Corr(health summary measure, share of federal transfers in mun. health budget) = 0.186, over the 2001-02 period for low procurement-intensity transfers.
6. Robust standard errors in brackets, clustered at the municipal level, in columns (1)-(4);
7. Coefficients and standard errors in columns (5) and (6) were bootstrapped using the procedure described in the Supplementary Appendix, Section, A;
8. *** p<0.01, ** p<0.05, * p<0.1
### Table 6 – Differences-in-differences estimates of the anti-corruption program on mortality rates, by cause

<table>
<thead>
<tr>
<th></th>
<th>Mortality rates</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[2001-04]</td>
<td>[1997-2007]</td>
<td>[2001-04]</td>
<td>[1997-2007]</td>
<td></td>
<td>[1997-2007]</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Post-program x</td>
<td>0.263*</td>
<td>0.261*</td>
<td>0.0847</td>
<td>0.107</td>
<td>0.225</td>
<td>-0.036</td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.148]</td>
<td>[0.149]</td>
<td>[0.136]</td>
<td>[0.135]</td>
<td>[0.126]</td>
<td>[0.143]</td>
</tr>
<tr>
<td>Post-program</td>
<td>-0.0918</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.150]</td>
<td>[0.151]</td>
<td>[0.141]</td>
<td>[0.141]</td>
<td>[0.151]</td>
<td>[0.142]</td>
</tr>
<tr>
<td>Second half of term x</td>
<td>-0.436</td>
<td>-0.430</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.358]</td>
<td>[0.348]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second half of term</td>
<td>0.237</td>
<td>0.310</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.257]</td>
<td>[0.269]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td>0.0841</td>
<td>0.270**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement-intensive</td>
<td>[0.343]</td>
<td>[0.131]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>-0.124</td>
<td>-0.184*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.196]</td>
<td>[0.101]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Municipality fixed-effects</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year fixed-effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Observations: 1,864 1,864 2,695 2,695 1,864 2,695
Number of clusters: 613 613 723 723 613 723

Notes on Table 6:

1. All columns are regressions with dependent variable defined as mortality rates by cause of death – external causes matched to low-procurement intensity indicator and general causes (all but external) matched to the high-procurement intensity indicator; (ii) general mortality rates (excluding external);
2. Columns (1) to (6) are Generalized Least Squares (GLS) regressions, with the number of investigations in each municipality and year used as weights;
3. Procurement-intensive equals 1 if procurement-intensity is 50% or higher, and 0 otherwise;
4. Procurement-intensity is coded as the share of actions under each transfer that include procurement-related words; see Appendix A;
5. Robust standard errors in brackets, clustered at the municipal level, in columns (1)-(4);
6. Coefficients and standard errors in columns (5) and (6) were bootstrapped using the procedure described in the Supplementary Appendix, Section A;
7. *** p<0.01, ** p<0.05, * p<0.1.
### Table 7 - Differences-in-differences estimates of the anti-corruption program on health outcomes, alternative counterfactual

<table>
<thead>
<tr>
<th></th>
<th>(1) Imunization (shots / 1,000 inhabitants)</th>
<th>(2) % under 1-year-old children with malnutrition</th>
<th>(3) % under 2-year-old children with diarrhea</th>
<th>(4) Infant mortality per 1,000 live births</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-program x</td>
<td>0.187</td>
<td>-0.003</td>
<td>0.003</td>
<td>2.689</td>
</tr>
<tr>
<td>Above-median</td>
<td>[1.267]</td>
<td>[0.004]</td>
<td>[0.004]</td>
<td>[2.134]</td>
</tr>
<tr>
<td>Second-half of term x</td>
<td>0.251</td>
<td>-0.004</td>
<td>-0.013</td>
<td>0.576</td>
</tr>
<tr>
<td>Above-median</td>
<td>[3.478]</td>
<td>[0.015]</td>
<td>[0.017]</td>
<td>[3.980]</td>
</tr>
<tr>
<td>Second-half of term</td>
<td>17.564</td>
<td>-0.031</td>
<td>0.004</td>
<td>4.634</td>
</tr>
<tr>
<td></td>
<td>[10.024]</td>
<td>[0.020]</td>
<td>[0.021]</td>
<td>[7.871]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1 x</td>
<td>1.206</td>
<td>0.004</td>
<td>-0.001</td>
<td>-0.204</td>
</tr>
<tr>
<td>Above-median</td>
<td>[2.631]</td>
<td>[0.006]</td>
<td>[0.005]</td>
<td>[3.598]</td>
</tr>
<tr>
<td>Audits (own or nearby) in t-1</td>
<td>-2.273</td>
<td>0.004</td>
<td>-0.002</td>
<td>-1.933</td>
</tr>
<tr>
<td></td>
<td>[2.171]</td>
<td>[0.006]</td>
<td>[0.008]</td>
<td>[4.398]</td>
</tr>
</tbody>
</table>

**Observations**: 1192 1141 1141 1193  
**Number of clusters**: 359 344 344 359  
**Municipality fixed-effects**: Yes Yes Yes Yes  
**Year fixed-effects**: Yes Yes Yes Yes  
**Mean (pre-2003)**: 77.95 0.045 0.580 22.08

**Notes on Table 7:**
1. Columns (1) to (4) are Ordinary Least Squares (OLS) regressions;  
2. Coefficients and standard errors in columns were bootstrapped using the procedure described in the Supplementary Appendix, Section A;  
3. Above-median baseline corruption equals 1 if the prevalence of corruption in the municipality is greater or equal to the median sample prevalence in the 1997-2002 period, and 0 otherwise;  
4. Robust standard errors in brackets, clustered at the municipal level, in columns (1)-(4);  
5. Coefficients and standard errors in columns (5) and (6) were bootstrapped using the procedure described in the Supplementary Appendix, Section A;  
6. *** p<0.01, ** p<0.05, * p<0.1.