

Can Intrastate Accountability Reduce Local Capture? Results from a Field Experiment in Mexico*

Ana L. De La O
Yale University
ana.delao[at]yale.edu

Fernando Martel García
Cambridge Social Science
Decision Lab Inc.
fmg444[at]gmail.com

October 23, 2015

PRELIMINARY DRAFT

Abstract

Improving accountability in public service provision is one of the most pressing challenges that young democracies face. This study contributes to an emerging body of literature examining the workings of intrastate accountability. Our evidence is based on a field experiment conducted in collaboration with Mexico's Federal Superior Auditor's Office where we randomly assigned municipalities to be audited by federal auditors, by state auditors, and a control group. We find that federal auditors produce harsher reports and recover substantially more money than state auditors. In turn, local governments assigned to a federal audit decrease the share of expenditures allocated to social infrastructure and increase the share allocated to social assistance the year following the audit. State audits, which are more lenient, have no effect on municipal expenditures. These results are compatible with harsher audits leading municipal administrators to skew public resources towards categories of spending that are more difficult to audit.

*The research protocol for this project, which lists our hypotheses, was registered at EGAP after randomization but before the completion of data collection and before any analysis had been performed. We would like to thank Leonard Wantchekon, Andrew Gelman, Alan Gerber, Don Green, Luke Keele, Craig McIntosh, Jake Bowers, Cyrus Samii, Ken Scheve, and participants of EGAP-Fall 2014 for helpful comments, as well as to the Federal Auditor's Office in Mexico for their collaboration in this project. We acknowledge financial support of Yale's Institute for Social and Policy Studies, the Leitner Program in International and Comparative Political Economy, and NYU. Of course, all errors are ours. Yale IRB Protocol number: 1106008610.

Most democracies in the developing world face the challenge of improving their mechanisms of accountability, which are directly related to the quality of governance and legitimacy of the political system (Mainwaring 2003). For some scholars and practitioners, elections are the vehicle to do so because citizens can control elected officials with their votes (Moreno et al. 2003). Yet adverse selection and moral hazard cripple electoral accountability, like in any other principal-agent relationship (Barro 1973, Ferejohn 1986). Moreover, when elected politicians have exhausted their time in office, elections are but blunt instruments of accountability (Persson 1997, Przeworski et al. 1999, Fearon 1999), and even if elections were perfect instruments to hold elected politicians accountable, what disciplines non-elected public servants?

Intrastate accountability, or the network of state agencies charged with overseeing and/or sanctioning public officials for the discharge of their public duties (Diamond 1999, Mainwaring 2003) is the central yet less studied component of political accountability. O'Donnell (2003) refers to these state agencies as *horizontal accountability*. He explains: “The horizontal accountability characteristic of representative democracy is extremely weak or non-existent in delegative democracies ... [Its] absence is a serious flaw” (p. 61). Leaving the terminology debate aside, the consensus is that intrastate accountability has “theoretical importance in understanding the dynamics and deficiencies of democracy” (Mainwaring 2003, pp. 29). This paper contributes to an emerging strand of literature examining the workings of accountability agencies.

Oversight agencies can affect governance through the information they produce about the competence and behavior of governments. So far, scholarship has for the most part focused on studying whether (and how) voters use this information to punish or reward incumbent governments.¹ Here, we take a step back to empirically study two angles of audit institutions:

¹ Ferraz and Finan (2008, 2011) present evidence that mayors exposed as corrupt by audit reports in Brazil are more likely to lose elections. See also Lopes and Lichand (2008). Such indirect effects may also be moderated by other features of the polity, like term-limits, and may vary over time. Pereira, Melo and Figueiredo (2009) only find effects when audit reports are released on the election year itself. Chong, De La

First, we focus on whether central and decentralized auditors are equally effective at detecting and reporting local capture. Following the decentralization literature, local auditors could be more effective because they have better access to local information. On the other hand, central auditors could be more effective because they are more insulated from local political pressure. Moreover, in most developing countries, federal governments have more resources (human and financial) than local governments, tilting the scale towards the effectiveness of federal auditors.

Secondly, we explore how local governments respond to central and decentralized auditors. Oversight may deter public officials' dissonant behavior, especially when, if caught, sanctions are greater than the benefits derived from corruption.² More oversight, however, is not a silver bullet. There is a widespread concern among scholars of young democracies that increased oversight comes at the cost of governmental effectiveness (Mainwaring 2003), policy paralysis, and very low decisiveness (O'Donnell 2003). Moreover, increased oversight may lead to the displacement of local capture, just like increased enforcement can lead criminal activity to be displaced (Yang 2008).³ Such possibilities underscore the importance of better understanding mechanisms of non-electoral accountability.

To explore these issues, we conducted a field experiment in Mexico where, in collaboration with the Mexican Superior Audit, we randomly assigned municipalities to one of three ex-

O, Karlan and Wantchekon (2015) present evidence that exposing mayors' corruption in Mexico decreases turnout, incumbent and challenger support.

²For the economic approach to crime see: Becker (1968). Allingham and Sandmo (1972) theorize the individual decision to misreport income when faced with a positive probability of an audit. Kleven et al. (2010) randomize tax audits and find that "individuals update their beliefs about detection probability based on experiencing an audit.". This is surprising since audit events are independent, and points to behavioral aspects of the audit relation (Fehr and Falk 2002, Gneezy and Rustichini 2000). Baron and Besanko(1984) is a theoretical application of *ex post* audits to regulation. Olken (2007) is a field experiment where project-specific audit probabilities are randomized and found to be effective in reducing corruption. DiTella and Schargrodsy (2003) identify a significant effect of an audit program on procurement in Argentinian hospitals. Zamboni (2011) also find strong effects on procurement of raising the audit probability, but not on the quality of "publicly provided preventive and primary health care services, measured using client satisfaction surveys".

³Bobonis, Fuertes and Schwabe (2009) find that corruption increases after an audit. Olken and Pande (2011) find that audits may only shift corruption around.

perimental conditions: a federal audit, a state audit, or a control group. Audits were part of the national program of audits. We find that federal auditors produce harsher reports and recover substantially more money than state auditors. In turn, local governments assigned to a federal audit decrease the share of expenditures allocated to social infrastructure and increase the share allocated to social assistance the year following the audit. State audits, which are more lenient, have no effect on municipal expenditures. These results are compatible with harsher audits leading municipal administrators to skew public resources towards categories of spending that are more difficult to audit.

We also estimate the effects of federal and state audits on a range of outcomes that allow us to say something about the mechanisms that lead to social spending changes, including municipal authorities' priorities for public spending, knowledge acquisition, perceptions of their own capacity, and compliance with program rules.

Our findings contribute to nascent literature on intrastate accountability in general and audit institutions in particular, which have significant gaps. First, studies commonly do not distinguish between the effects of a national audit program and the effects of actually being audited. In fact, we know of no instances where audit programs are themselves randomized. Much of the empirical work on municipal audits has focused on Brazil where a group of municipal governments are selected randomly every year to be audited (see, for example, Ferraz and Finan 2008 and Bobonis et al. 2009). These studies contrast audited municipalities to non-audited municipalities. Yet all municipalities are familiar with the national program of audits. Our study focuses on municipalities who prior to our study were rarely audited by the federal auditors. Thus, we assume that our experimental municipalities react to federal audits like municipalities that are newly incorporated to a national audit program.

Second, scholarship has predominantly focused on estimating the effect of a threat of an audit on government's performance. Here, we focus instead on the effects of being audited, which could have an effect via the detection and sanctioning of corruption, and could also have an

effect via “soft-touch” strategies, like capacity building, training, and diffusion of knowledge. To disentangle these two mechanisms, we conducted a survey among local public officials to measure potential knowledge acquisition, perceptions of their own capacity, compliance with program rules; and perceptions about audit probabilities.

Finally, studies usually focus on estimating the effects of one type of auditor. We are able to randomize audits by federal and state level institutions across 17 states, to gain some insight on how institutional characteristics and political context may affect effectiveness.

The paper proceeds as follows. Section 2 motivates the discussion by describing the electoral and institutional mechanisms of municipal accountability in Mexico, with especial emphasis on the use of the Federal Municipal Social Infrastructure funds (FISM in Spanish) by the municipal governments and the role played by auditors in monitoring the use of this fund. Section 3 presents our argument and testable hypotheses. Section 4 describes the research design, paying particular attention to the problem of implementing transparent, replicable, and verifiable randomization protocols in policy experiments. Section 5 presents our results. Section 6 concludes.

Oversight agencies and the local provision of public goods in Mexico

Mexico’s municipalities provide basic public services like drinking water, sanitation, improved road surfaces, and electricity, to 113 million citizens, though access to these services remains uneven across, and within, Mexican municipalities.⁴ Improving access of marginalized populations to basic municipal public services is a key element of Mexico’s National Development

⁴Mexican municipalities are often as large as U.S. counties. A typical municipality comprises a council seat, which is typically the largest and wealthiest locality, and several outlying urban and rural localities.

Plan.⁵ The main instrument available to the Federal Government to achieve this goal is public spending, including earmarked federal grants. For example, the federal Contribution Fund for Social Infrastructure (FISM, in Spanish) provides grants for municipal investments in basic public service infrastructure benefiting local marginalized populations. In FY 2009 it financed one-third of all basic public investment in municipalities, or some 100,000 individual investments (ASF 2011).⁶ However, the reliance on federal transfer schemes as they key instrument for improving access to public services is not without risks.

Municipalities' ability to identify marginalized communities, diagnose their basic public service needs, propose policy solutions, and implement them is weak. Municipal investments are heavily concentrated on paving and roads – even though the federal mandate emphasizes basic services like water and sanitation. Moreover, the use of federal funds for purposes unrelated to the development of marginalized areas, embezzlement, and corruption are a problem (García 2008, Pardini 2010, ASF 2011) – despite claims that decentralization of spending would result in better monitoring and allocation of resources.

The principal mechanism by which the Federal Congress oversees local governments' use of federal resources is the national program of audits, directed by the Superior Federal Auditors (ASF, in Spanish) in coordination with the Superior Audit Entities of States (EFSL, in Spanish).

ASF was created in 2000 as an autonomous agency with the mandate to audit the use of federal resources in hands of federal, state and municipal authorities. ASF has the prerogative

⁵The National Development Plan 2006-2012 states : “It is necessary to bring more actions to [poor] communities in education, health, nutrition, housing, as well as drinking water projects, drainage, electricity, roads, and even telephone, to communicate distant localities to the municipal council seat.” (Gobierno De Los Estados Unidos Mexicanos 2007, pg. 144).

⁶Prior to the decentralization process begun in the early 1980s such public services were grossly underfunded or dependent on state level funding. From 1990 onwards regular mechanisms for federal transfers were put in place, including the Municipal Funds for Solidarity (1990-94), the Solidarity Municipal Fund (1994-95), and the Fund for Municipal Social Development (1996-97). The current framework stems from the Fiscal Coordination Law approved in 1997 and small modifications thereafter. In addition to funding from federal grants, both state and federal governments at times provide additional discretionary grants or intervene directly by, for example, funding a sewage station in a municipality.

to establish compensating liabilities, initiate criminal lawsuits, and to make audit reports publicly available. Nevertheless, ASF does not have the means to sanction criminally or administratively public officers. The Federal Auditor is elected by the Lower House of Congress for an 8-year term. So far, there have been two Federal Auditors, both highly qualified for the position. ASF is perceived as a neutral, autonomous, and professional agency by national and international experts (UNAM 2007). Between 2001 and 2011, ASF began 901 liability procedures of which 709 resulted in the refund of public money, and 192 are still ongoing. Moreover, ASF presented 310 criminal lawsuits against public officials, most of which are still in the courts.

As part of the yearly National Program of Audits, ASF selects a few municipalities per state to be audited. Because of budget constraints, a share of these municipalities is audited by ASF's auditors and the rest is audited by EFSL's auditors. Outside of the National Program of Audits, ESFLs conduct their own audits on regular basis. EFSLs vary in terms of their capacity, their legal mandate, and their autonomy. Most EFSLs are underfunded, do not have the prerogative to initiate criminal lawsuits, and are required to submit their audit reports to local congress for approval before making them public (IMCO 2013). State Superior Auditors are selected by local congress, and in numerous cases there are no clear guidelines for their appointment and removal. High rotation among State Superior Auditors is a problem and so is the lack of civil service (IMCO 2013). Moreover, the lack of transparency of ESFLs themselves is a source of concern (IMCO 2013). In sum, despite ASF's efforts, municipal accountability seems to remain lacking (Comisión de Vigilancia, 2012).

Hypotheses

In Besley (2006) model of political agency, which combines moral hazard with adverse selection, voters can punish dissonant behavior by not reelecting representatives who have failed

them. Thus, electoral accountability acts as a disciplining instrument. Yet, with term limits, electoral accountability unravels when politicians' time in office has exhausted because voters can't punish representatives that are not in the ballot. This does not mean that incumbents cannot be disciplined. We discuss two extensions to the basic model of political agency that could result in accountable governments.

One possibility is that accountability is achieved through institutional means. SAIs audit and punish dissonant behavior. Thus, through criminal accountability audits can deter public officials from engaging in dissonant behavior. We would expect that audit institutions endowed with more resources to detect and to punish dissonant behavior, as well as with more autonomy, produce harsher reports and have a more effective deterrence effect. In the Mexican context, we would expect that federal auditors are more effective compared to state auditors in detecting dissonant behavior and increasing compliance with the regulations of federal transfer schemes to municipalities. One substantively important manifestation of an increase in compliance in Mexico would be an increase in spending allocated to social infrastructure projects and actions that benefit low-income areas in the municipality.

SAIs may also improve accountability through "soft-touch" strategies, including transfers of knowledge and assistance to build local capacity. This is particularly relevant if the lack of accountability and under provision of public goods is caused by public officers' low administrative capacity, unproductive administrative habits or simple lack of information. If audits increase knowledge and capacity, then we should see an increase in compliance with the regulations of federal resources.

Another possibility is that career concerns create incentives for accountability (via political parties or the private market). The effect of SAIs would work here through indirect channels. SAIs provide information on public officials' competence and behavior in office, but leave the punishment to other principals including political parties or organized groups in the private sector. The indirect mechanism is at work if public officials have career concerns, which

link their performance in office with their career prospects, and if the information revealed by SAIs influences those career concerns. Information on municipalities' competence and behavior may increase accountability, but could have ambiguous effects on citizens' welfare.

Audits could help public officers to better assess their own "type". Public officers concerned with their career prospects may have incentives to reallocate resources after the revelation of their type. For example, the mayor of a municipality deemed low capacity may reallocate resources towards categories of spending that require less capacity to be disbursed in compliance with regulations.

On the other hand, the mayor of a municipality deemed as corrupt may have no further incentive to pretend to be honest during his remaining time in office. If such mayor considers his or her principal is the political party, and the political party cares about its reputation, then the mayor may have incentives to reallocate resources to counteract the report by, for example, increasing spending in private goods or public goods that are easy to roll out. If such mayor considers his principal is an interest group, he or she may have incentives to reallocate resources towards categories of spending that benefit such group.

To explore whether audits increase accountability and the provision of public goods, we analyze our experimental data by first conducting a few tests of the effects of random assignment to state and federal audits on awareness of treatment status and beliefs about the probability of being audited in 2 and 3 years. Then, we test whether federal auditors produce harsher reports (including the number of remarks, their severity, and the amounts of mandated reimbursements to federal treasury of misspent grant money) compared to state auditors. Next, we explore impacts of randomized assignment to audits on outcomes such as knowledge about program requirements, capacity building, and career concerns. Finally, we estimate the effects of audits on public official's spending priorities, and actual spending patterns. Table 1 includes a list of expected outcomes.

Table 1: *Expected outcomes*[†]

Treatment status and probability of audits
Municipal administrators in treated municipalities are aware of their treatment status
Municipal administrators in treated municipalities audited in year 1 believe the probability of being audited in year 2 is lower than in year 3
Municipal administrators in treated municipalities have higher long-run beliefs about the probability of being audited
Differences between state and federal audits
Federal auditors (ASF) yield more remarks, more severe opinions, and more refunds to the federal treasury than state auditors (EFSL)
Soft-touch strategies and compliance
Municipal administrators in treated municipalities have more knowledge of FISM grant rules and regulations
Municipal administrators in treated municipalities are more aware of lack of capacity and more likely to manifest plans for improving capacity
Municipal administrators in treated municipalities are more likely to comply with FISM reporting and data accessibility rules
Municipal administrators in treated municipalities manifest preferences for municipal investments more in accordance with FISM priorities
Career prospects
Municipal administrators in treated municipalities have different expectations about future political appointments
Municipal administrators in treated municipalities have different expectations about career prospects
Municipal administrators in treated municipalities perceive the ASF as a more important principal
Spending
Municipal administrators reallocate expenditures to comply with federal regulations
Governor's reaction: audited municipalities receive more funds to compensate federal refunds mandated by the auditors (section in progress).

Study Design

This project started in April 30, 2009, when we presented a proposal to ASF senior staff to conduct a RCT to evaluate the impact of federal and state audits coordinated in the National Program of Audits. During the first two years of the project, we were in continuing conversation with the ASF regarding the RCT, and collaborated with them in a series of data collection efforts in order to establish a trustworthy relationship between the parts. In the third year, ASF approved the RCT. The random assignment of audits took place in March 26, 2011. Auditors were in the field from April 2011 to approximately November 2011. In compliance with our collaboration agreement with ASF, data collection for the project began in April, 2011 with the collection of baseline characteristics based on official sources, and is still undergoing. The survey was on the field between April 27, 2012 and June 7, 2012.

We conducted a block randomized field experiment in a convenience sample of eighty-five municipalities in Mexico. The process to select the convenience sample and our randomization strategy is explained in detail below. We blocked our sample by state. Each of 17 blocks contains 5 municipalities. Using random assignment, one municipality per block is assigned to be audited by the ASF, another by the EFSL, and the remaining three municipalities serve as controls.

No power analysis was done for this field experiment. The reason is our implementing partner (the ASF) gave us a strict limit on the number of audits they would allow us to randomize. Given this restriction, we sought to power the study by using an unbalanced block design, which improves covariate balance and efficiency.⁷

⁷We plan to conduct an ex-post power analysis to determine minimum detectable effects.

Selection of experimental group

As part of the National Program of Audits, the ASF selects every year a group of municipalities (as well as state and federal entities that rely on funds from the federal budget) to be audited. Our study focuses on the set of audits that examine how municipalities spend the FISM federal grant, which is intended to fund municipal public services and basic infrastructure. ASF coordinates all of the FISM audits, but a subgroup of these audits is directly conducted by ASF (direct), and another subgroup of audits is delegated by the ASF to the EFSL (indirect). Results of direct and indirect audits are collated into the ASF's annual report. The EFSL may carry out additional audits as part of the state program of audits, but these are not part of the National Program of Audits and are not included in the final ASF annual report.

The ASF gave us the opportunity to add to the national program of audits a set of randomly assigned direct and indirect audits to municipalities' use of FISM. We conformed our experimental group in such a way that minimized disruption to the National Program of Audits. In order to do so, we selected our convenience sample using the following criteria:

Stage 1 From the universe of 2,440 municipalities located in 31 states select:

1. States with more than 20 municipalities;
2. Municipalities with FISM transfers in 2010 of 10 million pesos or more;
3. Municipalities not audited in the previous two years (2009, 2010);
4. Municipalities not amongst the 43 pre-selected by the ASF for the 2011 National Program of Audits.

Stage 2 From this selection of [767] municipalities located in [21] states select:

1. States with 5 or more municipalities;

2. For each state, rank municipalities in decreasing order of FISM transfers and choose by state the five municipalities with ranks 6 to 10.

The first stage of the selection process of our convenience sample guarantees that our experimental sample includes municipalities that are of relevance to the ASF in terms of the amount of transfers received through the FISM transfer scheme.

The second stage of the selection process ensures we have 5 municipalities per state in the experimental group; that our experimental group includes municipalities that are unlikely to have been audited since 1998, when the current audits to FISM expenditures began; and that, within states, municipalities in our sample are similar in terms of the amount of transfers received through the FISM scheme.

The final selection includes 5 municipalities in each of 17 states for a total experimental group sample of 85 municipalities.

Assignment to treatment and control arms

We randomly assigned municipalities in the experimental group to one of the three experimental arms: audit by ASF, audit by EFSL, and no audit.

For the random assignment to be acceptable to the implementing agency, it had to meet a number of constraints. These constraints responded to ASF need to explain, justify, and replicate the randomization mechanism to Congress members. First, the randomization process had to be objective, replicable, and verifiable. Second, the randomization process had to be compatible with the existing operational and technological infrastructure of the implementing agency (effectively limiting applicable software solutions to Microsoft Excel).

The block randomization process was as follows:

1. By state, we provided each municipality with a pair of single-digit “tickets”:
 - (a) Block municipalities by state
 - (b) In Excel list municipalities in increasing order based on their individual identifier provided by the Mexican National Institute of Statistics and Geography (INEGI, in Spanish).
 - (c) Assign each municipality two single-digit “tickets”, and do this sequentially for all municipalities (e.g. 0-1, 2-3, 4-5, 6-7, 8-9 ...).

2. We generated a random vector of “winning digits”:
 - (a) To generate the random “winning digits”, we used the winning numbers of the seven largest prizes of the Mexican National Lottery of the first Tuesday of March 2011.
 - (b) Each winning number has 5-digits.
 - (c) We ordered the 5-digit winning numbers in decreasing order of prize.
 - (d) Our first five “winning digits” come from the number associated with the highest prize (e.g. for the date we used, the number was 23862 and the price 5 million pesos), the next ten “winning digits” digits come from the second and third prizes.
 - (e) The fourth largest prize (of 80,000 pesos) was won by four numbers. To order these tied lottery numbers randomly, we (1) ordered the numbers in increasing order; (2) grab the number associated with the largest prize in the lottery of 22 February (e.g. number 36625), delete one repeated digit (e.g. becomes 3625); (3) assign one of these digits to each of the four tied lottery numbers; (4) use this assigned digit to sort the four tied lottery numbers in increasing order (e.g. 2,3,5,6).
 - (f) Concatenating the 15 “winning digits” from three lottery numbers associated with the three top prizes, and the random ordering of the four lottery numbers tied for

fourth prize, gives us a random vector of 35 “winning digits”, enough to randomly assign 17 municipalities to ASF audit, and 17 municipalities to EFSL audit.

3. We then assigned municipalities to treatment arms based on the random vector of “winning digits”:
 - (a) Start reading from the top of the vector of “winning digits”. The first winning digit is a 2, so assign the municipality in the first state holding the single-digit “ticket” 2 to an ASF audit. Then, use the second “winning digit” from the vector to assign a municipality in the second state to ASF audit, and so on for all 17 states.
 - (b) Repeat the procedure – starting from the 18th element of the vector of winning digits – to allocate one municipality by each of the seventeen states to an audit by the EFSL.
 - (c) Municipalities not allocated to EFSL or ASF serve as control.

A worked example of the randomization procedure is provided in Table 2.

This method of randomization has advantages as well as disadvantages. The advantages are that the randomization is transparent, replicable, and verifiable. In addition, the only software requirements are a web browser (to access the lottery numbers) and Microsoft Excel. These features were key for the ASF to accept the procedure. The disadvantages are two. First, it is somewhat complicated, mainly because we were limited to using Excel.⁸ Second, lottery numbers span the range 00000 to 59999.⁹ Accordingly, the first digit of every winning lottery number can only take the values 0 through 5, compared to all other digits that can take values from 0 to 9. Thus, the fourth and fifth municipalities in the first

⁸Originally we had planned to use the winning number associated with the highest prize as a random seed. Combining this random seed with a specific version of the statistical software R (which fixes the pseudo random number generator) would have generated a replicable randomization. There exist R front ends for Excel, but our partners were not allowed to install this front ends in their systems.

⁹This happens in the “Sorteo Mayor” lotteries.

Table 2: *Example of Random Allocation for Two States*

Panel A: Creating the Sequence of Random “Winning Digits”

Lottery 3/1/2011		Lottery 2/22/2011	
Number	Prize (millions)	Number	Prize (millions)
23862	5	36625	5
19186	0.4		
54595	0.2	Sort order (ascending)	
02437	0.08	3	
09502	0.08	6	
42585	0.08	2	
45776	0.08	5	

Panel B: Randomization of Municipalities[†]

ID	State	Municipality	FISM transfer (millions)	“Ticket” Digits	ASF	EFSL
07022	Chiapas	Comitán	de 69	0-1		
		Domínguez				
07028	Chiapas	Chenalhó	61	2-3	1	
07076	Chiapas	Ocozocoautla de Es-	68	4-5		1
		pinosa				
07092	Chiapas	San Cristóbal de Las	65	6-7		
		Casas				
07111	Chiapas	Tecpatán	59	8-9		
08012	Chihuahua	Carichí	12	0-1		
08021	Chihuahua	Delicias	15	2-3	1	
08030	Chihuahua	Guazapares	11	4-5		
08032	Chihuahua	Hidalgo del Parral	13	6-7		
08066	Chihuahua	Uruachi	13	8-9		1

[†]The sequence of random numbers from Panel A is: 23862, 19186, 54595, 42585, 02437, 45776, 09502. We use the first 17 winning digits to allocate one municipality by state to an audit by ASF. For example, the first winning digit in the random sequence is a 2. Because Chenalhó was allocated that “ticket” (see Panel B, Digits column), it is selected to be audited by ASF. The second winning digit in the random sequence is a 3, and so Delicias is selected, and so on for the remaining 15 states. To allocate EFSL we begin at the top again, starting with the 18th digit in the random sequence, a 5. Accordingly, Ocozocoautla de Espinosa is allocated to EFSL, and so on. Had the 18th digit been a 2 or a 3, we would have skipped that digit, moved to the next digit different from 2 or 3, and used that digit to allocate the first municipality to EFSL. One municipality cannot be assigned to both ASF and EFSL.

state of our study have in practice zero chance of being audited by the ASF because they hold “tickets” (6,7) and (8,9) respectively. After the first assignment, this happens every fifth assignment, when a new lottery number is added to the sequence of “winning digits”. In other words, the randomization procedure generates known non-uniform probabilities of treatment in a subset of the blocks.¹⁰ To take this into account in our analysis, we will use inverse probability weighting.

Interventions

Other than the random assignment, the intervention consists of audits carried out as usual by federal and state level auditors. See Merino and Aramburo (2009) for details of the standard auditing process. Figure 1 provides a schematic layout of municipal audit activities for the FISM program.

We did not directly monitor implementation of intervention since this would have compromised the regular operation of the National Audits Program.

Instruments and measurements

Our outcome data come from audit reports, other official sources, direct observations by the investigators, and from a proprietary survey of municipal administrators.

The survey was developed by the investigators and implemented by the Mexican survey firm *Data Opinion Publica y Mercados*. The survey firm was blind to treatment status.

The survey was pilot tested on four municipalities similar to the ones in the experimental group, and the results were used to clarify the meaning of questions and adapt the length

¹⁰Because there are 17 states in our study, out of 34 allocations to treatment, only 4 ASF and 3 EFSL are affected by the non-homogenous randomization.

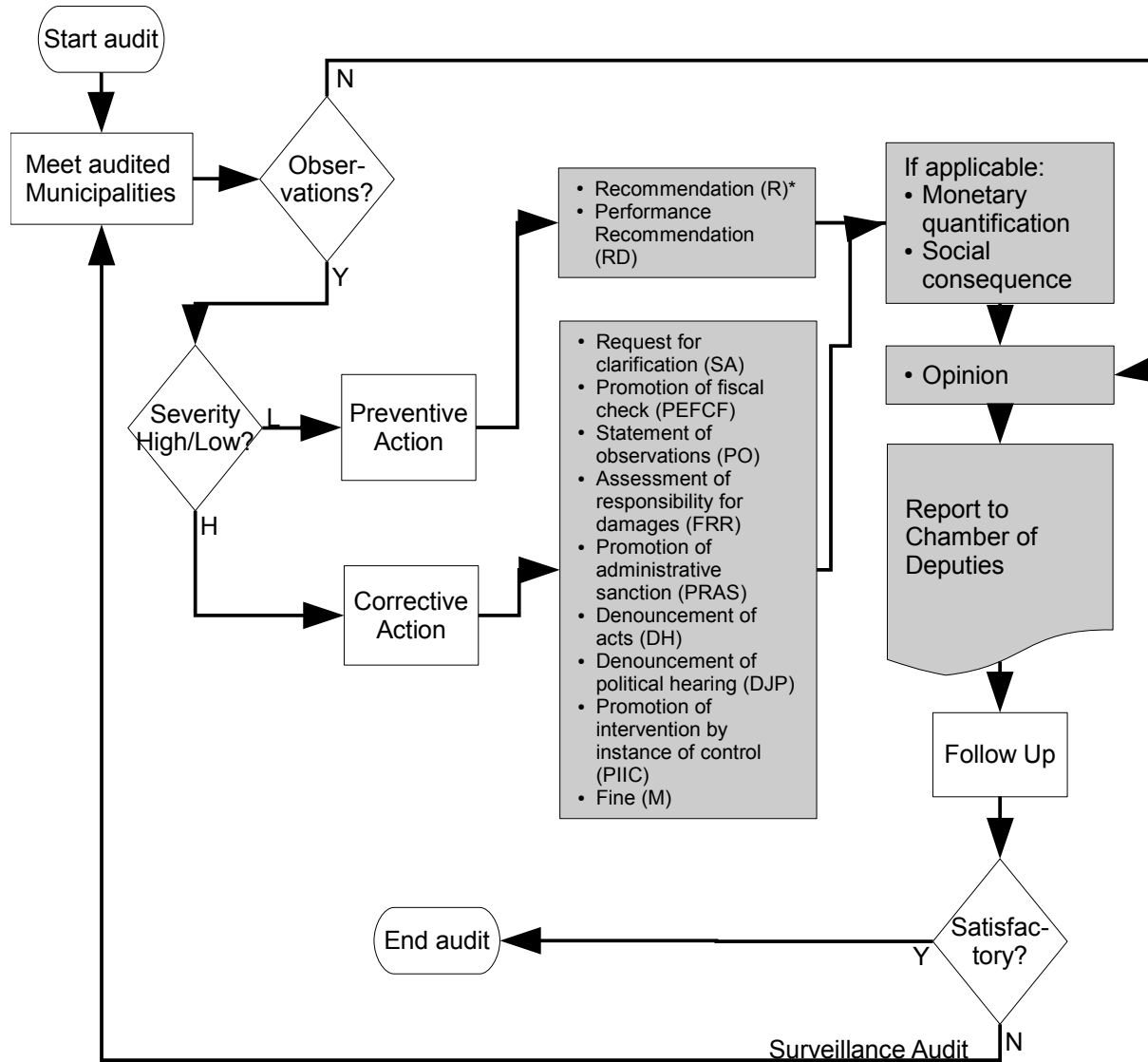


Figure 1: Flow chart depicting the Superior Federal Auditor's (ASF) audit process of municipal expenditures under the federal Contribution Fund for Social Infrastructure (FISM) grant program. Highlighted in grey are ASF judgements, opinions, and outputs. Explanations of the preventive and corrective actions can be found (in Spanish) in ASF (2009, pp. 33- 35)

Table 3: *List of endpoint data to be collected*

DATA ITEM	DATA TYPE	OBS. UNIT	SOURCE
Data collected from routine audit reports			
Number of remarks per municipal audit	Count	Municipality	ASF 2011
Actions promoted (see Fig. 1)	Ordinal
Overall opinion of the auditor
Money refunded to the federal treasury
Reasons adduced by ASF and EFSL for re-funds to federal government	Categorical
Data collected from official sources			
FISM transfers in year prior to intervention and two subsequent years	Continuous	Municipality	Fed. Treasury
Quarterly geographic and programmatic allocation of municipal FISM investments	..	Investment item	..
Data collected by direct observation of researchers			
Municipal data access and transparency	Various	Municipality	Investigators
Data collected through Survey of Municipal Administrators[†]			
Subject data module	Various	Administrators	Tel. Survey
Career concerns module
FISM knowledge module
FISM priorities module
Audit Module
Administrative capacity module

[†]For a copy of the survey forms please contact the corresponding author

of the survey, as well the contact strategy. The survey was fielded over the phone between April 27, 2012 and June 7, 2012. We administered the survey to key personnel in each municipality, including: the Municipal President, Treasurer, Director of Public Services, Director of Public Works, and/or Director of Urban Planning. It was not always possible to contact the personnel, in which case we moved down the municipal hierarchy. Given the sample size of this study, strenuous efforts were made to ensure full response. Table 3 summarizes the kinds of data collected.

Analytical Plan

Because we have a small sample and are concerned about power, we begin by asking very little of the data and ask progressively more depending on the answers to previous queries. In particular we follow the inferential process below:

1. *Sharp null hypothesis test*: We begin by testing the sharp null of no effect on any unit against the alternative of some effect (e.g. change in location, scale, or distribution). These tests can tell us whether the treatment has an effect, but they are silent as to the magnitude and variability of the effect.
2. *Visual inspection of outcome distributions*: We plot histograms, box plots, and density plots, as relevant, for the outcomes of interest across treatment arms.
3. *Descriptive inference*: We describe measures of central tendency, like experimental group averages and their standard deviation, along with the difference across averages and their standard deviations.
4. *Modeling*: When appropriate, we generate estimates of causal effects and confidence intervals by invoking assumptions of non-interference. That is, we calculate Intent-to-treat estimates. Then, we discuss potential threats to causal inference associated with non-compliance.

Four features of this inferential process are worth highlighting. First, because the treatment was randomized with known probabilities we rely on randomization tests of the sharp null of no effect (see Keele, McConnaughey and White (2012) for details). Second, for testing the sharp null we typically choose test statistics with the most power in small to medium samples. For example, unlike differences in means, rank based statistics do not suffer from low power given outliers, skewed distributions, or fat tails (Keele et al. 2012, pp. 6 - 7).

Third, we classify null hypotheses into families of hypotheses, and use sequential testing procedures to control the family-wise Type-I error rate (Rosenbaum 2009; Small, Volpp and Rosenbaum 2011). This approach can reduce the experiment-wise Type-I error rate in a context of multiple hypothesis testing, and be more powerful than comparable Bonferroni procedures.

Finally, whereas the treatment was block randomized to municipalities several of our outcome variables are measured at the individual municipal administrator level. At this level of analysis, the treatment can be thought of as cluster randomized. To avoid bias in our estimates, we aggregate individual responses to the cluster level (i.e. the municipality), and base our test on the aggregate data.

An overview of the hypotheses and the corresponding tests is displayed in Table 4. A more detailed description of the outcome concepts, measures, measurement instruments, and analytical procedures can be found in the online Appendix.

Table 4: *Overview of hypothesis tests*[†]

Treatment status and probability of audits

Awareness of treatment status

The null is that median responses to the treatment status questionnaire are the same across ASF, EFSL, and control groups. The ordered alternative is that audited municipal administrators are aware of treatment status, with those audited by ASF more aware relative to those audited by EFSL

Short-run beliefs about probability of being audited

The null is that municipal administrators belief about the probability of being audited in the two years after the audit (years 2 and 3) is the same across treatment and control groups. The one sided alternative is that administrators in audited municipalities are more likely to believe the probability of being audited in year 2 is lower than in year 3

Long-run beliefs about the probability of being audited

The null is that municipal administrators belief about the probability of being audited in any given year during the the next three years, is the same across experimental arms. The one-sided alternative is that beliefs about the probability of being audited are higher in audited municipalities compared to control ones.

Continued on next page

Table 4 – continued from previous page

Differences between state and federal audits

Institutional effectiveness

The null is that the number of remarks per municipal inspection and the amount of refunds obtained for the federal treasury is the same across ASF and EFSL. The ordered alternative is that the ASF is more effective or stringent, yielding more remarks per audit and more refunds. We also test for heterogeneous effects by stratifying using party affinity and auditor independence measures.

Institutional severity

The null is that both the ASF and EFSL hand out similar judgements and opinions. The ordered alternative is that the ASF yields more stringent judgements and opinions. We also test for differences in the reasons given by ASF and EFSL for refunds.

Soft-touch strategies and compliance

Knowledge acquisition

The null is that knowledge about the rules and regulations of the FISM federal grant program is the same across experimental arms. The ordered alternative is that audited municipal administrators display more knowledge relative to control, with those audited by ASF doing at least as well as those audited by an EFSL.

Perception of FISM Priorities

The null is that the respondent's perception of how FISM grants ought to be allocated geographically between council seat and outlying localities, and programatically across public, club and private goods is the same across experimental arms. The ordered alternative is that audited respondents allocate more to outlying localities or public goods, with those audited by ASF doing at least as well as those audited by an EFSL.

Perception of municipal capacity

The null is that respondent's evaluation of municipal capacity and the need for training are the same across experimental arms. The ordered alternative is that audited respondents have a lower perception of municipal capacity and greater perception of need for training, with those audited by ASF doing at least as well as those audited by an EFSL.

Compliance with FISM reporting and data accessibility

The null is that the municipal scores and experimental arm indicator are independent. The ordered alternative is that there is a positive monotone relation between experimental arm and scores, with municipalities audited by the ASF scoring at least as high as those audited by EFSL and so on.

Career prospects

Subject's expectations about future political appointments

Continued on next page

Table 4 – continued from previous page

The null is that respondents in audited municipalities receiving a low opinion from any auditor have the same expectations about future political appointments as respondents in the control arm. The one-sided alternative is that respondents in audited municipalities receiving a low opinion from any auditor have lower expectations about political appointments. We look at the effect of receiving a high opinion analogously.

Subject’s expectations about future career prospects

The null is that respondents across all treated and control arms have the same distribution of career expectations, the alternative is that they differ in location and/or scale. If that null is rejected we test for differences in scale, which would suggest greater sorting by principals. Finally we test whether auditor scores of incumbents and their career expectations are aligned amongst the treated.

Subject’s rank ordering of principals

The dependent variable has three levels. Along with the 17 blocks this defines a $2 \times 3 \times 17$ contingency table when we pool treatments. Conditional on the blocking variables, the null is that the odds ratio of being in the treatment group is the same across columns. The one-sided alternative is that it exhibits a rising trend across columns. If we reject the null, we test the effect of each treatment against control separately.

Spending

Actual allocation of FISM grants

The null is that the geographic allocation of FISM grants between council seat and outlying localities, and programatically across public, club, and private goods, is the same across experimental arms. The ordered alternative is that audited municipalities allocate more to outlying localities or to goods that benefit low-income residents (i.e. increase compliance), with those audited by ASF doing at least as well as those audited by an EFSL.

Governor’s reaction

Federal FISM grants are distributed via state governors who exercise some discretion. The null is that being audited has no main effect on the change in gubernatorial FISM transfers. The alternative is that audited municipalities receive more funds to compensate federal refunds mandated by the auditors. If that null is rejected we also test for heterogeneous effects stratifying by party affinity, and test whether there is an association between the amounts refunded to the Federal government and changes in FISM transfers.

†When possible we use randomization inference for all tests of sharp null hypotheses. The specific test statistic is chosen to maximize power for the given test, often depending on the nature of the outcome variable and how it is measured (e.g. nominal, ordinal, interval or ratio scale).

Results

Table 5 presents the effects of our treatments on awareness of treatment status. We hypothesized that municipal administrators in treated municipalities are aware of their treatment status, with those audited by federal auditors more aware relative to those audited by state auditors. In our survey, we asked whether the municipality was audited by the EF or the ASF. We find that municipal administrators assigned to an EF audit are nine percentage points more likely than the control group to report that they were audited by the EF. This difference is significant at the ten percent level. Yet, assignment to an EF audit also leads municipal administrators to report at higher rates than the control that they were audited by the ASF (.14 percentage points significant at the five percent level).

On the other hand, municipal administrators assigned to an ASF audit report at higher rates than the control group that they were audited by the ASF (.23 percentage points significant at the one percent level), and they report at the same rate as the control group that they were audited by the EF. Comparing the two treatment groups, municipal administrators report at similar rates an EF audit, but administrators assigned to ASF do report at higher rates an AF audit (.10 percentage points significant at the ten percent level). A possible explanation for these results is that the frequency of state audits is higher than the frequency of a federal audit, which is consistent with the higher proportion of municipal administrators in all groups who report an EF audit. In addition, EF audits are commissioned by ASF, which could create confusion.

We also asked in our survey whether the administrator has heard of the ASF, the EF, and whether they were aware of ASF mandate to oversee federal resources allocated to municipalities. We created an index of awareness based on these questions, and the correct answers to treatment status. Both assignment to EF and ASF increase the index of awareness in a statistically significant way, but they are not different from each other.

Table 6 displays the results for municipal administrators' knowledge of FISM rules, perceptions of their own capacity, and plans for training. We hypothesized that audited municipal administrators display more knowledge about the rules of federal transfer schemes such as FISM relative to the control group, and that administrators audited by ASF displayed more knowledge than administrators audited by EF. In our survey we asked 24 items about FISM regulation, including what types of projects can be funded with these resources, how often the money is allocated to the municipality, how to report expenses, etc. We created an index based on the sum of correct responses to such items. Overall, municipal administrators in the three experimental groups responded 62 percent of the knowledge items correctly. We find that assignment to EF audits leads to very minor increases in knowledge, and assignment to ASF leads to no increases compared to the control group.

We also hypothesized that audited municipal administrators have a perception of lower municipal capacity and greater need for training relative to the control group, with those audited by ASF doing at least as well as those audited by EF. In our survey, we asked whether the municipality had updated information about the needs of the population, staff to implement and oversee public works, mechanisms to evaluate investments and public works. We added responses to these questions to create an index of capacity that ranges from 0 to 4. As expected, we find that EF and ASF audits lead to a 0.25 and 0.42 percentage points lower perception of capacity relative to the control group, respectively. In a parametric test, both of these effects are statistically significant at least at the five percent level. In a non-parametric test, the effect of assignment to EF is close to being statistically significant at the 10 percent level, and assignment to ASF is statistically significant at the one percent level. Comparing the two treatment arms, the effects are not different from each other. Our survey also included a couple of questions about training opportunities available to municipal administrators. We added responses to these survey items, and created an index of training plans. We find that EF and ASF audits decrease municipal administrators' reports of future training relative to the control group. The treatment groups are not different to each other.

Thus, audits lead to a perception of lower capacity. Yet, audits also lead to a decrease in administrators' reports of future training.

In Table 7, we include the results regarding the differences between state and federal audits. We hypothesized that the ASF is more effective and stringent, yielding more remarks per audit and more refunds compared to EF. We collected these outcomes directly from the audit reports. We find that EF reports 14 remarks per audit and ASF reports 22, in average. The difference of 7.4 remarks per audit is statistically significant at the 1 percent level. Out of the total remarks per audit, EF resolves 2.22 whereas ASF resolves 17.67, in average. The difference of 15.4 is statistically significant at the 1 percent level. Thus, ASF yields more remarks per audit and resolves more of these remarks relative to EF.

Based on their remarks, the auditors produce an overall assessment of the municipality, which can be: abstention, clean, negative or with exceptions. EF and ASF produce a similar proportion of abstentions, 7 and 6 percent, respectively. EF produces 32 percent of clean assessments and ASF produces none. The difference is statistically significant at the one percent level. EF produces 34 percent of negative assessments and ASF produces 63 percent of negative assessments. The 28 percentage points difference is statistically significant at the one percent level. Finally, EF produces 27 percent of assessments with exceptions, and ASF produces 31 percent. The 4 percentage points difference is not statistically significant. In terms of refunds, both EF and ASF report in average a potential refund of 8.5 and 8.3 million pesos, respectively. The 175,000 pesos difference is not statistically significant. However, EF actually recovers 179,000 pesos, whereas ASF recovers 2.3 million pesos. Thus, ASF recovers 2.1 million pesos more than the EF per audit (these difference is statistically significant at the 5 percent level in a parametric test, and at the 1 percent level in a non-parametric test).

Most of the municipal administrators that responded to our survey do not hold an elected position, with the exception of the municipal president. Who do they feel accountable to? Do audits change their perception of who their principal is? We hypothesized that municipal

administrators assigned to a ASF audit perceive the ASF as a more important principal. In our survey, we asked administrators whose opinion is most important to them when thinking about being accountable. Table 8 shows that, in average, 44, 43 and 60 percent of the administrators in control, EF, and ASF municipalities, respectively, are most concern with the ASF's opinion. The group assigned to EF is not different to the control group, but the group assigned to ASF reports caring about ASF's opinion at higher rates than the control and EF group. The differences of 16 and 17 percentage points are statistically significant at the one percent level.

Remarkably, very few municipal administrators across the three experimental groups (2 percent, in average) report caring about the opinion of municipalities' residents who live outside the capital of the municipality. More municipal administrators report caring about the opinion of municipalities' residents who live in the capital, in the control group 28 percent, in the EF group 23 percent, and in the ASF group 23 percent. The audits have no effect on how much administrators care about voter's opinions.

Some municipal administrators care about the opinion of the municipal president. In the control group, in average, 7 percent of administrators mentioned the mayor, in the EF group 15 percent of administrators mentioned the mayor, and in the ASF group six percent of administrators mentioned the mayor. The difference 8 and 9 percentage points difference between the EF group and the control and ASF group, respectively, are statistically significant. Thus, whereas an ASF audit leads administrators to see the ASF as a principal, an EF audit leads administrators to see the mayor as principal.

We hypothesized that audits lead municipal administrators to have different expectations about future political appointment and, more generally, career prospects. Table 9 includes our findings on career prospects. In our survey, we asked whether working in the current municipal government improves administrators' career prospects, whether administrators expect to have a higher wage in their next job, and whether they consider public servants to

be well-regarded by their community. We also asked where administrators expect to work once the municipal administration exhausts its time in office.

In the control group, a high percentage of administrators (80 percent) see their work in the municipality as an asset. Yet assignment to an EF and ASF audit decreases the share of administrators who think that working in the municipal government improves their career prospects by 16 and 13 percentage points. These differences are statistically significant at the one percent level. The difference between the EF and ASF group is not statistically significant. Similarly, a high percentage of administrators in the control group (90 percent) report that public servants are well-regarded by their communities. Yet assignment to either an EF or an ASF audit leads to a 5 and 7 percentage points decrease in such reports. The difference between the EF and ASF group is not statistically significant. About expectations of future wages, assignment to a EF audits decreases by almost 50 percentage points the expectation of a higher wage in the next job, whereas assignment to an ASF audit has no statistically significant effect.

When asked about their future jobs, 33 percent of administrators in control municipalities expect they will work in the next municipal government. Assignment to an EF or an ASF audit has no effect on this expectation. 40 percent of administrators in the control group expect to have their own business once their time in office is over. Audits do not change that expectation either. 10 percent of administrators in the control group expect to move up to either the state or federal government. Assignment to an EF audit leads to a 5 percentage point decrease in the expectation of moving up to a higher level of government (statistically significant at the five percent level). Assignment to an ASF audit has no effect on such expectation. The six percentage points difference between the EF and ASF group is statistically significant in the non-parametric test at the ten percent level. Finally, in the control group 14 percent of administrators expect to either retire, move to the education sector, or become employees in the private sector. In the EF and ASF groups, 19 and 10

percent of administrators expect so, respectively. Only the difference of 9 percentage points between the EF and ASF group is statistically significant at the 10 percent level in the non-parametric test. Thus, most administrators expect to have their own business, or continue their work in the municipal government. A few of them expect to advance in their careers to a higher level of government, and a few other expect to be out of government without a business of their own. While an ASF audit has little impact on career expectations, and EF audit leads to less administrators expecting to move up and more expecting to be retired, or employed in the education or private sectors.

We surveyed municipal administrators who are in some way or another responsible for the provision of public services in the municipalities. We hypothesized that audited respondents allocate more of FISM resources to outlying localities, and more resources to public goods compared to club or private goods, with those audited by ASF doing at least as well as those audited by EF. Tables 10 and 11 include these results. We find that administrators in the three experimental groups report that about 40 percent of resources should be spent inside the council seat. We also find that administrators in the three experimental groups report that about 50 percent of resources should be spent in public goods, and about 21 percent in club goods. Neither type of audit had an effect on these expressed preferences. Administrators in the control group said that 27 percent of resources should be spent in private goods, administrators in the ASF group said 29 percent of resources should be spent in these category, and administrators in the EF group said 31 percent of resources should be spent in private goods. The difference of 4 percentage point increase in private goods funding between the EF and the control is statistically significant at the ten percent level in the parametric test, and in the one-sided non-parametric test. Thus, audits have little effect on what administrators think should be allocated to public, club, or private goods.

We also asked in our survey, administrators perceptions of what people want in terms of funding public, club and private goods. We find that in the ASF group, administrators

think people want 50 percent of resources spent in public goods, compared to 41 and 42 percent in the EF and control groups, respectively. The ASF group is different to the other two groups in a statistically way, at the 5 percent level. The three groups think that people want around 22 percent of resources spent in club goods. And, administrators in the control group think people want about 32 percent of resources spent in private groups, administrators in the EF group think people want 37 percent of resources going to private goods, and administrators in the ASF group say 28 percent of resources. Assignment to an EF audit leads to more administrators thinking that people want private goods, and assignment to a ASF leads to less. These differences are statistically significant in the non-parametric one-sided tests. Finally, we asked administrators what percentage should their neighboring municipality spend in public, club or private goods if they wanted to win an election. In accordance to their answers to what they think people want, assignment to an EF audit decreases the percent of resources that administrators think should be spent in public goods to win an election, and increases the percentage of private goods, relative to both the control and the ASF group (these results are statistically significant at least at the ten percent level in the one-sided non-parametric test). Assignment to an ASF audit increases the percentage of public goods relative to the control and EF group (statistically significant at least at the 10 percent level in the one-sided non-parametric test) and decreased the percentage of private goods relative to the EF group (statistically significant at least at the 10 percent level in the one-sided non-parametric test). Thus, although audits have no effect on administrators reports of what should be funded, they do have an effect on what administrators think people want and what they think should be funded to win elections.

Next, we explore whether audits have an effect on public spending. We collected municipal budgets and finance data from two sources. First, the National Institute of Statistics and Geography (INEGI) collects yearly data on municipal income and expenditures. The information comes from municipal and state treasurers. From this source, we collected spending figures, which are categorized into broad concepts, including municipal spending allocated

to public officials wages, office supplies, operational costs (i.e. general services), social infrastructure, transfers to government agencies (including health and education), subsidies to private sector and social assistance (including scholarships, in-kind food subsidies, humanitarian aid, temporal work, etc.). This data can give us a first cut at the effects of audits.

We computed the average spending per category between 2006 and 2010 to create a baseline spending measure. This “normal” baseline spending helps us reduce measurement error and missing values in the time series. Municipalities in our study spent between 2006 and 2010, in average, 25% of their income in wages, 40% in social infrastructure, 5% in office supplies, 10% in operational costs, and 9% in transfers, subsidies and social assistance.¹¹ Then, we compute the change in spending per category from the baseline to 2012, that is the year after the audits were completed.

We find that state and federal audits have no effect on changes in expenditures allocated to wages, office supplies or general services. Nor do state audits have any effect on spending. However, federal audits lead to an increase of 1.83 percentage points on expenditures allocated to social assistance, and a decrease of 0.20 percentage point on expenditures allocated to social infrastructure. Both results are significant at the 10%.¹²

Discussion

So far, we find that, as expected, federal auditors produce more remarks per audit, recover substantially more money (two million pesos more per audit) and their reports are harsher compared to the state auditors. We also find that federal audits increase the share of

¹¹The rest of the resources are allocated across various categories, all of which receive less than 1% of the budget.

¹²We are also collecting from the Ministry of Finance quarterly expenditure data, which will allow us to test whether audited municipalities spend more resources on poor localities compared to the control group.

expenditures allocated to social assistance the year after audits are completed, and decrease the share of expenditures allocated to social infrastructure. Why do federal audits have such an effect on spending?

“Soft-touch” strategies do not explain these spending patterns because audits have no effect on municipal officers’ knowledge about the rules of federal transfer schemes, or on public officials’ preferences for public spending. Moreover, audits decrease the perceived capacity of the municipality, but do not increase municipalities plan to build up their capacity through, for example, training of its administrators. Indeed, audits decrease municipalities’ offer of training.

Noticeably, social assistance is a category of expenditures that is harder to audit compared to social infrastructure. While the former leaves little trail, the latter gives auditors more inputs to evaluate the performance of governments, and potential wrongdoings. For example, a physical audit on a school would involve an assessment of the financial records of the project, an inspection of the building, measurement of the quality of construction materials, and a comparison of the quality and quantity of materials at market rate with the reported expenditures. In contrast, an audit of a subsidy for school breakfasts program would involve little more than the self-reported financial operation involved. Hence, by reducing infrastructure spending and increasing social assistance, local governments are displacing resources to categories of spending that are harder to oversee.

For now, an open question that is what is the overall welfare effects of different types of audits. In an ongoing project, we are collecting quarterly expenditure data from the Mexican Ministry of Finance. In the near future, we will be able to say something more specific about the effects of audits on the welfare of poor people. In the meantime, the evidence in this paper reveals important angles of the workings of intrastate accountability in Mexico. As O’Donnell (2003) and Mainwaring (2003) explained, building intrastate accountability is crucial to the functioning of democracies, yet the task is far from simple. Not only because increased

oversight comes with transaction costs, but because institutional design choices (central vs. local auditors) matter, and audited public officials can displace their wrongdoings.

References

- Allingham, Michael G. and Agnar Sandmo. 1972. "Income tax evasion: a theoretical analysis." *Journal of Public Economics*, 1(3-4):323- 338.
- ASF. 2009. Informe del Resultado de la Revisión y Fiscalización Superior de la Cuenta Pública 2007. Technical Report. Auditoría Superior de la Federación.
- Auditoría Superior de la Nación. 2011. Informe del Resultado de la Fiscalización Superior de la Cuenta Pública 2009: Marco de Referencia. Technical Report Volume V, Title 4, Section 1. Auditoría Superior de la Federación.
- Baron, D.P. and D. Besanko. 1984. "Regulation, asymmetric information, and auditing." *The Journal of Economics*, pp. 447-470.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *The Journal of Political Economy* , 76(2):169-217.
- Bobonis, G.J., L.R.C. Fuertes and R. Schwabe. 2009. "Does Exposing Corrupt Politicians Reduce Corruption?". Unpublished manuscript.
- Chong, A. A.L. De La O, D. Karlan, L. Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification" *Journal of Politics*, January 2015, Vol. 77, No. 1, pp. 55-71.
- Di Tella, Rafael and Ernesto Schargrodsky. 2003. "The Role of Wages and Auditing during

- a Crackdown on Corruption in the City of Buenos Aires.” *Journal of Law & Economics*, 46(1):269 - 92.
- Fehr, Ernst and Armin Falk. 2002. “Psychological foundations of incentives.” *European Economic Review*, 46: 687- 724.
- Ferraz, Claudio and Frederico Finan. 2008. “Exposing Corrupt Politicians: The Effects of Brazils Publicly Released Audits on Electoral Outcomes.” *Quarterly Journal of Economics*, 123(2):703-745.
- Ferraz, Claudio and Frederico Finan. 2011. “Electoral Accountability and Corruption: Evidence from the Audits of Local Governments.” *American Economic Review* 101(4):1274 - 1311.
- García, Mariana. 2008. Cómo ejercen recursos y rinden cuentas los municipios? El Caso del Fondo para la Infraestructura Social Municipal del Ramo 33. Technical report Centro de Investigación para el Desarrollo (CIDAC).
- Gneezy, Uri and Aldo Rustichini. 2000. “A Fine Is a Price.” *The Journal of Legal Studies*, 29(1):117.
- Gobierno De Los Estados Unidos Mexicanos. 2007. Plan Nacional de Desarrollo 2007-2012. Technical report Presidencia de la República.
- Keele, Luke, Corrine McConnaughy and Ismail White. 2012. “Strengthening the Experimenters Toolbox: Statistical Estimation of Internal Validity.” *American Journal of Political Science*, pp. 1-16.
- Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, Sren Pedersen and Emmanuel Saez. 2010. “Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark.” Working Paper 15769 National Bureau of Economic Research.

- Lopes, Marcos Felipe Mendes and Guilherme Finkelfarb Lichand. 2008. "Random Audit Programs and Game-Theoretic Models: establishing ex-ante corruption control." *Anais do XXXVI Encontro Nacional de Economia [Proceedings of the 36th Brazilian Economics Meeting]* 200807182137430 ANPEC - Associao Nacional dos Centros de Psgradauo em Economia [Brazilian Association of Graduate Programs in Economics].
- Mainwaring, S. 2003. "Introduction: Democratic Accountability in Latin America" in S. Mainwaring and C. Welna (Eds.) *Democratic Accountability in Latin America*. Oxford University Press.
- Merino, Mauricio and Michiko Aramburo. 2009. "Informe sobre la evolución y el desempeño de la Auditoría Superior de la Federación". Technical report. Auditoría Superior de la Federación.
- Moreno, E., B. Crisp, M. S. Shugart. 2003. "The Accountability Deficit in Latin America" in S. Mainwaring and C. Welna (Eds.) *Democratic Accountability in Latin America*. Oxford University Press.
- O'Donnell, G. 2003. "Horizontal Accountability: The Legal Institutionalization of Mistrust" in S. Mainwaring and C. Welna (Eds.) *Democratic Accountability in Latin America*. Oxford University Press.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 115(2): 200-249.
- Olken, Benjamin A. and Rohini Pande. 2011. "Corruption in Developing Countries." Working Paper 17398. National Bureau of Economic Research.
- Pardinas, Juan E. 2010. "Índice de Competitividad Estatal 2010: La Caja Negra del Gasto Público". Technical report Instituto Mexicano para la Competitividad (IMCO).

Pereira, Carlos, Marcus Andre Melo and Carlos Mauricio Figueiredo. 2009. “The Corruption-Enhancing Role of Re-Election Incentives?: Counterintuitive Evidence from Brazils Audit Reports.” *Political Research Quarterly*, 62(4):731-744.

Rosenbaum, P.R. 2009. *Design of Observational Studies*. Springer Verlag.

Small, Dylan S., Kevin G. Volpp and Paul R. Rosenbaum. 2011. “Structured Testing of 22 Factorial Effects: An Analytic Plan Requiring Fewer Observations.” *The American Statistician*, 65(1):11-15.

Yang, Dean. 2008. “Can Enforcement Backfire? Crime Displacement in the Context of Customs Reform in the Philippines.” *Review of Economics and Statistics*, 90(1): 1-14.

Table 5: Awareness of treatment status

	Treatment assigninment:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	(3)>(2)>(1)			
Dependent variables:							
Audited by EF	0.72 (0.32)	0.80 (0.24)	0.77 (0.23)		0.09 (0.05)	0.51 (0.05)	-0.04 (0.05)
				0.18	0.09 (0.07)	0.32 (0.32)	0.45 (0.86)
Audited by ASF	0.39 (0.33)	0.53 (0.29)	0.62 (0.24)		0.14 (0.06)	0.23 (0.05)	0.09 (0.05)
				0.00	0.01 (0.01)	0.00 (0.00)	0.09 (0.07)
Awareness index	3.04 (0.75)	3.72 (0.71)	3.36 (0.44)		0.68 (0.13)	0.32 (0.12)	-0.36 (0.12)
				0.00	0.00 (0.00)	0.01 (0.00)	0.00 (0.99)

Table 6: Knowledge and capacity

	Treatment assigninment:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	(3)>(2)>(1)			
Knowledge Index	15.02	16.02	15.02		1.00	0.00	1.00
	(3.37)	(2.07)	(2.49)		(0.53)	(0.54)	(0.47)
				0.96	0.06	0.99	0.04
					0.13	0.49	0.05
Capacity Index	3.27	3.02	2.85		-0.25	-0.42	-0.17
	(0.70)	(0.85)	(0.86)		(0.13)	(0.13)	(0.18)
				0.00	0.05	0.00	0.34
					0.11	0.00	0.30
Training Index	1.61	1.45	1.40		-0.16	-0.21	-0.04
	(0.39)	(0.44)	(0.38)		(0.07)	(0.07)	(0.09)
				0.00	0.02	0.00	0.62
					0.02	0.00	0.50

Table 7: Audit results

	Treatment assignment:		Differences in means and std. errors (p-value from t test) (p-value from jonter test)
	EF (Means and std. dev.)	ASF (Means and std. dev.)	
	(2)	(3)	(3) vs (2)
Remarks	14.56 (10.85)	21.98 (7.51)	7.42 (1.96) 0.00 0.00
Resolved remarks	2.22 (5.27)	17.67 (6.66)	15.45 (1.29) 0.00 0.00
Overall assessment: Abstention	0.07 (0.26)	0.06 (0.24)	-0.01 (0.05) 0.84 0.84
Clean	0.32 (0.47)	0.00 (0.00)	-0.32 (0.07) 0.00 0.00
Negative	0.34 (0.48)	0.63 (0.49)	0.28 (0.10) 0.01 0.01
W/ Exception	0.27 (0.45)	0.31 (0.47)	0.04 (0.10) 0.65 0.65
Money planned to recover (in thousand pesos)	8507.56 (11041.44)	8331.96 (10803.15)	-175.60 (2320.82) 0.94 0.84
Money recovered (in thousand pesos)	179.99 (518.33)	2328.77 (6428.78)	2148.78 (1007.62) 0.04 0.00

Table 8: Who you are accountable to?

	Treatment assignment:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF (std. dev.)		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	$ (3)>(2)>(1) $			
ASF	0.44 (0.37)	0.43 (0.25)	0.60 (0.27)		-0.01 (0.06)	0.16 (0.06)	0.17 (0.05)
				0.03	0.83 0.97	0.01 0.01	0.00 0.01
Residents outside the council seat	0.02 (0.09)	0.02 (0.08)	0.02 (0.08)		0.00 (0.01)	0.00 (0.01)	0.00 (0.02)
				0.77	0.81 0.81	0.81 0.81	1.00 1.00
Residents inside the council seat	0.28 (0.30)	0.23 (0.23)	0.23 (0.26)		-0.05 (0.05)	-0.05 (0.05)	0.00 (0.05)
				0.32	0.33 0.42	0.34 0.42	1.00 0.94
Municipal president	0.07 (0.17)	0.15 (0.24)	0.06 (0.13)		0.08 (0.03)	0.00 (0.03)	-0.09 (0.04)
				0.19	0.01 0.01	0.94 0.60	0.04 0.10

Table 9: Career Prospects

	Treatment assignment:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	$(3) > (2) > (1)$			
Working in the mpio improves your career prospects	0.80 (0.30)	0.64 (0.31)	0.67 (0.30)		-0.16 (0.05)	-0.13 (0.05)	0.03 (0.06)
				0.00	0.00	0.01	0.65
					0.00	0.00	0.72
In your next job, your wage will be higher	0.70 (0.30)	0.49 (0.35)	0.65 (0.33)		-0.21 (0.05)	-0.06 (0.05)	0.16 (0.07)
				0.72	0.00	0.27	0.03
					0.00	0.35	0.02
Public servants are well-regarded	0.90 (0.24)	0.85 (0.20)	0.83 (0.29)		-0.05 (0.04)	-0.07 (0.04)	-0.02 (0.05)
				0.01	0.18	0.10	0.73
					0.01	0.04	0.75
Where do you expect to work: Municipality	0.33 (0.27)	0.34 (0.28)	0.35 (0.25)		0.01 (0.05)	0.03 (0.04)	0.14 (0.06)
				0.70	0.79	0.56	0.80
					0.66	0.76	0.82
State/Federal government	0.10 (0.16)	0.04 (0.11)	0.10 (0.21)		-0.05 (0.03)	0.01 (0.03)	0.06 (0.03)
				0.31	0.04	0.78	0.07
					0.04	0.79	0.15
Own Business	0.40 (0.30)	0.43 (0.29)	0.42 (0.26)		0.03 (0.05)	0.02 (0.05)	-0.01 (0.06)
				0.87	0.62	0.74	0.88
					0.96	0.87	0.92
Other (i.e. teacher, retire, employee)	0.14 (0.24)	0.19 (0.29)	0.10 (0.16)		0.05 (0.04)	-0.04 (0.04)	-0.09 (0.05)
				0.87	0.27	0.28	0.07
					0.50	0.62	0.34

Table 10: Priorities for public spending

	Treatment assignmnet:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF (dev.)		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	$ (3)>(2)>(1) $			
What % should be spent inside council seat	37.98 (20.06)	36.81 (39.55)	39.55 (19.13)		-1.17 (3.38)	1.57 (3.41)	2.74 (3.87)
				0.72	0.73 0.68	0.65 0.57	0.48 0.35
What % should be spent in: public goods	47.23 (19.43)	45.74 (12.67)	49.68 (16.40)		1.48 (3.07)	2.45 (3.21)	3.94 (3.02)
				0.67	0.63 0.64	0.45 0.50	0.20 0.30
club goods	21.51 (10.90)	23.09 (9.74)	20.96 (10.35)		1.57 (1.82)	-0.56 (1.85)	-2.13 (2.07)
				0.80	0.39 0.19	0.76 0.79	0.31 0.17
private goods	27.06 (13.13)	31.17 (12.64)	29.36 (12.80)		4.11 (2.24)	2.30 (2.25)	1.81 (2.62)
				0.32	0.07 0.13	0.31 0.58	0.49 0.42
What % people want spent in: public goods	42.53 (19.52)	41.45 (16.87)	49.66 (18.91)		-1.09 (3.24)	7.13 (3.33)	8.21 (3.70)
				0.07	0.74 0.76	0.03 0.02	0.03 0.03
club goods	22.24 (11.46)	22.00 (12.09)	22.40 (14.02)		-0.24 (2.00)	0.16 (2.10)	0.40 (2.70)
				0.49	0.90 0.95	0.94 0.39	0.88 0.61
private goods	31.89 (15.92)	36.55 (18.32)	27.94 (13.80)		4.66 (2.86)	-3.96 (2.64)	-8.62 (3.35)
				0.26	0.11 0.25	0.14 0.07	0.01 0.01

Table 11: Priorities for public spending

	Treatment assigninment:			Non-parametric test (p-value)	Differences in means and std. errors (p-value from t test) (p-value from jonter test)		
	Control (Means and std. dev.)	EF	ASF		(2) vs (1)	(3) vs (1)	(3) vs (2)
	(1)	(2)	(3)	$ (3)>(2)>(1) $			
To win elections, what % should be spent in:							
public goods	36.65 (23.44)	30.92 (21.63)	42.60 (21.52)		-5.73 (4.27)	5.95 (4.04)	11.68 (4.72)
					0.18	0.14	0.02
				0.39	0.12	0.09	0.01
club goods	20.04 (13.73)	18.00 (13.56)	21.60 (13.21)		-2.04 (2.54)	1.56 (2.39)	3.60 (2.93)
					0.42	0.51	0.22
				0.98	0.41	0.74	0.28
private goods	28.27 (19.58)	33.13 (21.22)	29.13 (18.19)		4.86 (3.72)	0.87 (3.38)	3.99 (4.30)
					0.19	0.80	0.36
				0.63	0.10	0.87	0.17

Table 12: Changes in public spending

	Treatment assigninment:			Differences in means and std. errors	
	Control (Means and std. dev.)	EF	ASF	(p-value from t test)	(p-value from RI)
	(1)	(2)	(3)	(2) vs (1)	(3) vs (1)
Change from baseline to 2012:					
Wages	0.05 (0.43)	0.11 (0.42)	0.05 (0.54)	0.05 (0.13) 0.66 1.00	0.01 (0.13) 0.95 0.91
Office supplies	0.29 (0.80)	0.36 (0.63)	0.23 (0.73)	0.08 (0.22) 0.71 1.00	-0.08 (0.22) 0.71 0.73
General services	0.09 (0.45)	0.23 (0.57)	0.30 (0.76)	0.10 (0.16) 0.56 1.00	0.18 (0.16) 0.24 0.32
Social Infrastructure	0.08 (0.49)	0.10 (0.49)	-0.20 (0.54)	0.08 (0.17) 0.99 1.00	-0.29 (0.16) 0.06 0.06
Transfers	0.08 (0.88)	-0.20 (0.76)	0.06 (1.12)	-0.28 (0.27) 0.84 0.99	0.05 (0.26) 0.29 0.85
Social Assistance	0.02 (1.32)	0.77 (1.16)	1.83 (5.20)	0.30 (0.96) 0.75 0.99	1.63 (0.90) 0.07 0.08

Notes: Spending data comes from INEGI's Municipal finances data set.