

Measuring Corruption Using Governmental Audits: A New Framework and Dataset*

Michael Denly[†]

August 16, 2021

For the most recent version of this paper, please visit:

www.mikedenly.com/files/audit-measurement.pdf

Abstract

For about 25 years, empirical scholarship on corruption has primarily relied on perceptions data, but the drawbacks of these measures are ample and well-known. More recently, analyses centered on Brazil have showcased the utility of randomly assigned audits as a more objective alternative to perception-based measures. However, Brazil is the only country with randomized audits and has many unique institutional features that limit the external validity of the numerous studies using the Brazil data. In this paper, I provide a new framework to assess the quality of audit data even when they are not randomly assigned. Specifically, I show that experimental or observational audit data are most credible to measure corruption when: 1) the auditing institution legally enjoys political/institutional independence, particularly from the executive branch; 2) the distribution of audits does not exhibit biases toward any particular group, especially political rivals following close elections; and 3) the implementation/dosage is consistent across similar types of audits. I demonstrate the utility of the framework by analyzing a massive new dataset of subnational audits from India, Mexico, Honduras, and Guatemala. The new data and framework proposed in this paper will help researchers undertake more objective analyses of governmental corruption around the world.

*For funding, I thank the Texas Politics Project, Innovations for Peace and Development, and the Wilson Endowment at the University of Texas at Austin. For extraordinary research assistance, I thank Maria Eugenia Blanco and the following current and former Research Affiliates and Graduate Research Fellows at UT Austin's Innovations for Peace and Development: Adityamohan Tantravahi, Akshat Gautam, Alejandra Zuniga, Alessandro Rey Acosta, Amila Lulo, Ana Ramirez, Ana Suarez, Anoushka Ramesh, Awab Ahmed, Blaine Finstein, Braden Holt, Brayan Salgado, Briana Rocha, Caleb Rudow, Calista LaMotte, Carmen Magana, Caroline Hastings, Caroline Woodman, Connor Hanna, Dalia Dabbagh, David Tibaquira, David Zulli, Debasmita Bhakta, Dorna Abdi, Elena Solimano, Emily Ibarra, Erin Eggleston, Ethan Tenison, Evan Shimek, Evelin Caro Gutierrez, Fabiola Ramirez, Felipa Mendez, Francesca D'Annunzio, Grace Thomas, Isabel Campos, Iulia Tothezan, Isabel Ayala, Ishani Pandya, Ivana Jelensky, Jenny Rodriguez, Joelean Hall, Jonah Isaac, Josh Meuth Alldredge, Jude Rios, Judy Lane, Kais Bhandari, Karan Kanatala, Kate Hopp, Katya Bandouil, Katherine Alonzo, Kelsia Adil, Lissette Almanza, Luca Venegoni, Luisa Venegoni, Magdalena Ibarra, Marco Mejia, Maria Fernanda Guerrero, Marianne Nader, Mary White, Mishan Kara, Mobin Piracha, Mohamed Abulfalgha, Monserrat Ramon, Nathan Stern, Nick Hoffman, Nicole Domingo, Nicky Pownall, Paige Johnson, Preston Saunders, Rachel Boles, Rachel Rosenburg, Ryan Jabra, Saamia Imtiaz, Samuel Gilette, Sara Lowe, Sterling Mosley, Tara Jackson, Temi Ajibola, Thomas Boswell, Vanessa Gonzales, Vanessa Lizcano, and Veronica Lulu Stracher. For advice or feedback, I thank Akshat Gautam, Arun Manuja, Ilona Wysmulek, Jan Pierskalla, John Gerring, John Ogallo, Lorenzo Crippa, Mike Findley, Ken Greene, Serge Nkuindja, and Xiaobo Lu. All errors are those of the author.

[†]PhD Candidate, Department of Government, University of Texas at Austin. mdenly@utexas.edu

Since the publication of [Mauro's \(1995\)](#) seminal study just over 25 years ago, empirical scholarship on governmental corruption has primarily relied on perceptions data.¹ As numerous studies have demonstrated, that is a grave problem: perceptions data suffer from myriad methodological challenges that make them undesirable for advancing new theory and knowledge (e.g., [Kurtz and Schrank, 2007a,b](#); [Andersson and Heywood, 2009](#); [Hollyer, 2018](#)). Nevertheless, some of the most prominent recent work on corruption continue to be based mainly off perceptions data (e.g., [Mungiu-Pippidi, 2015](#)).

Over the past 10-15 years, many objective measures of corruption have emerged as alternatives to perceptions data (e.g., procurement red flags, abnormal asset growth, tax evasion), but violations contained in governmental audit reports are perhaps the most promising. Corruption is a multifaceted, complex phenomenon, and audits capture that complexity—everything from the aforementioned objective measures to fraud, theft, and nepotism. Given that audits are specific to individuals or institutions, they also overcome the level-of-analysis problems that plague most country-based research using perceptions data (see [Gingerich, 2013a](#)).

To date, however, scholars using audit data have focused almost exclusively on the randomly assigned audits of Brazil's municipal transfer spending.² On the one hand, these data from Brazil are useful because they overcome selection/endogeneity problems and allow for causal estimation. On the other hand, scholars' constant use of these same data has generated knowledge with severe external validity challenges. Notably, Brazil is the only country in the world with randomized audits of its administrative units, and the randomized audit program only concerns around one percent of its municipalities each year ([Ferraz and Finan, 2018](#); [Colonnelli and Prem, 2021](#)). Fortunately, some other countries disseminate

¹ So many studies rely on corruption perceptions data that it is impossible to mention them all here, but some of the most pertinent ones include [Mauro \(1995\)](#) on growth; [Gerring and Thacker \(2004, 2005\)](#), [Kunicová and Rose-Ackerman \(2005\)](#), and [Lederman, Loayza and Soares \(2005\)](#) on institutions; and [Treisman \(2000, 2007\)](#) on culture, economic development, and democracy.

² See, for example, [Ferraz and Finan \(2008, 2011\)](#), [Ferraz, Finan and Moreira \(2012\)](#), [Melo, Pereira and Figueiredo \(2009\)](#), [Brollo et al. \(2013\)](#), [Pereira and Melo \(2015\)](#), [Brollo and Troiano \(2016\)](#), [Avis, Ferraz and Finan \(2018\)](#), [Cavalcanti, Daniele and Galletta \(2018\)](#), and [Colonnelli and Prem \(2021\)](#).

subnational audit data, but researchers need a framework for discerning the validity of these data given that they are not randomly assigned.

In this article, I develop a three-pillar framework to discern the validity of corruption data derived from audit findings. I validate my framework with observational data, but it is equally applicable to experimental and quasi-experimental research designs. The framework's first pillar asserts that the audit agency must not only have *de facto* political/institutional independence but a legal (*de jure*) basis for it. Second, given that most countries' audit agencies have discretion to perform risk-based audits, the frequency, selection, or distribution of audits must not exhibit bias toward any particular group. In most countries, political rivals—especially opposition-party politicians selected in close elections—will be the most salient group, but hometown or ethnic biases might be relevant for some countries as well. Third, the most credible audit data will also be able to empirically show fairness in the implementation or dosage of each audit. For example, opposition party politicians must not be subject to more stringent audits than those sharing the same party as the executive. Because relevant quantitative data will not always be available to assess implementation/dosage, it is at least necessary to qualitatively examine auditor professionalism and codes of ethics/conduct to prevent corrupt behavior.

I demonstrate the utility of my framework by collecting and analyzing a massive, new, micro-level dataset of subnational audit findings from India, Mexico, Honduras, and Guatemala. The specific data vary by country, but the panel dataset covers infractions committed, misappropriated or stolen money, follow-up on audit recommendations, whistleblower complaints received, and relevant information on the sector and sub-sector of the corruption activity. Overall, the new data and framework proposed in this paper will help researchers undertake more objective analyses of governmental corruption around the world.

The paper proceeds as follows. In Section 1, I lay out the challenges to measuring corruption and explain why perceptions-based indicators are not ideal for advancing new theory and knowledge. Section 2 describes the benefits of audit data in measuring corruption

and the external validity challenges associated with over-relying on the Brazilian audit data. Section 3 details the framework and data central to this paper. Using the novel data that I collected from India, Mexico, Honduras, and Guatemala, in Section 4 I analyze the extent to which the data from each country meet the framework. Section 5 discusses the lessons learned from the results, and Section 6 concludes.

1. Measuring Corruption with Observational, Perception-Based Data

Corruption entails “the misuse of public office for private gain”,³ and many analysts define the phenomenon to be much broader (e.g., Gingerich, 2013*b*; Mungiu-Pippidi, 2015). Irrespective of how one defines corruption, though, one thing is certain: by its very nature, corruption is a clandestine activity, so it is very difficult to measure.

The International Country Risk Guide (ICRG), Transparency International (Corruption Perceptions Index-CPI), and the World Bank (Control of Corruption score-Worldwide Governance Indicators) constructed the first widely-available corruption measures. Each of these indexes embarked on measuring corruption by aggregating and re-scaling survey data from businesspeople. Over time, as more data became available, Transparency International and the World Bank greatly diversified the data from which they constructed their corruption measures (e.g., Kaufmann, Kraay and Mastruzzi, 2011, 225). In the process, the measures gained significant conceptual intention,⁴ which is especially useful because corruption is multidimensional concept with experience-near and experience-distant meanings.⁵

³ This is probably the most-commonly accepted definition of corruption. For more on the definition of corruption, see, for example, Rose-Ackerman and Palifka (2016).

⁴ Intention is a synonym for connotation, meaning the “ensemble of characteristics and/or properties associated with, or included in, a given word, term, or concept” (Sartori, 1984, 24).

⁵ “Experience-distant concepts are ones that specialists of one sort or another... employ to forward their scientific, philosophical, or practical aims. Experience-near concepts, in contrast, are one[s] that someone might himself naturally and effortlessly use to define what he or his see, feel, think, imagine, and so on, which he would readily understand when similarly applied by others (Schaffer, 2016, 2).”

Although development of the perceptions indexes represented major steps forward in terms of measuring corruption, they never were ideal for theory development and testing (Hollyer, 2018). ICRG developed its index for the purposes of helping businesses make decisions about corruption risks related to foreign investment, so the measure is necessarily limited in terms of content validity.⁶ Among its many limitations, Transparency International CPI data are not suitable for over time comparisons, which severely limit scholars ability to use them for theory (Andersson and Heywood, 2009, 758). The World Bank's Control of Corruption score is perhaps the the most sophisticated of the three measures and is suitable for over time comparisons. Still, the measure uses questionable error structures,⁷ has limited construct validity,⁸ and suffers from information leakage,⁹ halo effects,¹⁰ and content opacity (Kurtz and Schrank, 2007a,b; Langbein and Knack, 2010; Thomas, 2010; Bersch and Botero, 2014; Gisselquist, 2014; Hollyer, 2018; Magnusson and Tarverdi, 2020).¹¹

Information leakage, halo effects, content opacity, questionable error structures, and low construct validity are very problematic from the perspective of theory and knowledge-building (Hollyer, 2018). All of these deficits are related, too. For example, when participants respond to survey questions about corruption, it is difficult for them to exclude extraneous information and concepts that may overlap such as democracy (Hollyer, 2018). Along these

⁶“Content validity assesses the degree to which an indicator represents the universe of content entailed in the systematized concept being measured” (Adcock and Collier, 2001, 537).

⁷ By “questionable error structures”, I mean that the Worldwide Governance Indicators aggregate multiple measures that are drawn from the same sources but assume that errors are independent across those sources. As Magnusson and Tarverdi (2020) demonstrate, allowing for cluster dependence among the errors leads to substantively different conclusions in at least two prominent studies.

⁸ Construct validity concerns whether the concepts measures what it is supposed to measure (Trochim, 2006). In particular, the Worldwide Governance Indicators have trouble with a component of construct validity called discriminant validity. It concerns whether measures are not associated with measures that they are not supposed to be associated with. The Worldwide Governance Indicators lack discriminant validity because the correlation among indicators is so high that one may ask whether they are really different (Langbein and Knack, 2010; Thomas, 2010).

⁹ Information leakage refers to when a survey taker’s perspective on one phenomenon is influenced by something else.

¹⁰ Halo effects specifically refer to when positive effects about something color or influence a person’s opinion in a positive way about something else.

¹¹ Kurtz and Schrank (2007a,b) also critique the Worldwide Governance Indicators for systematic measurement error, sampling bias, and cultural biases, but those critiques are less convincing nowadays, particularly since Kaufmann, Kraay and Mastruzzi (2011) have included more sources and countries.

lines, Kurtz and Schrank (2007a,b) show that countries' growth trajectories cloud survey respondents' answers on governance-related questions. Such problems compound as perceptions indexes include more sources, too, which not only complicates measure error structures but also makes the precise scope of perception-based measures necessarily opaque. With opaque measures, it is impossible to precisely verify the degree of construct validity: that is, whether concepts measure what they are supposed to measure (Trochim, 2006; Thomas, 2010). Which begs the question: how can a measure without construct validity be useful for theorizing? In short, its ability to do so is very limited, especially when investigating proximate and endogenous phenomena like democracy and corruption.

More recently, scholars have developed useful, Bayesian-based improvements to the perception indexes (Bersch and Botero, 2014; Standaert, 2015; Coppedge et al., 2020).¹² The most sophisticated of these measures is that of the Varieties of Democracy (V-Dem) Project, which relies mainly on expert coding and performs very highly in terms of reliability and validity (McMann et al., 2021). Nevertheless, any index relying on perceptions-based measures and expert coding cannot fully overcome the aforementioned mentioned trade-offs, because they still exist—albeit to a smaller degree (Hollyer, 2018, 118, 128).

2. Audit as Objectives Measures of Corruption

2.1. The Comparative Advantage of Audits

Generally, objective measures of any phenomenon, including corruption, do not capture as much conceptual intention as the perception-based ones, but audit-based measures are not as susceptible to such a drawback. The reason pertains to the diversity of violations that

¹² The benefits and drawbacks of Bayesian statistics as compared to their frequentist counterparts greatly exceed the scope of this paper. In brief, the Bayesian measurement models of corruption such as Bersch and Botero (2014), Standaert (2015), and McMann et al. (2021) allow for less missing data, incorporation of priors, and are more attuned to measuring intractable, unobservable concepts like corruption (Fariss, Kenwick and Reuning, 2020).

audits can reveal. More specifically, audits can capture all of the different objective corruption indicators that separate studies examine individually, including: bribes, procurement red flags, ghost firms, nepotism, spending misappropriations, tax evasion, and theft (e.g., Ferraz and Finan, 2008; Brollo et al., 2013). Consequently, audits generally provide a more comprehensive view of corruption than experience-based measures such as Transparency International's (2021) Global Corruption Barometer, websites like ipaidabribe.com, the World Bank's (2021) Enterprise Surveys, or sms-message based measures (Buntaine et al., 2018).

Because audit-based corruption measures, like any other objective measure, refer to actual instances of corruption, they also do not suffer from the same information leakage, halo effects, and construct validity challenges as perception-based measures. Accordingly, objective measures are more useful for theorizing and knowledge creation than perceptions-based measures (Hollyer, 2018). Perhaps equally as significant, objective measures such as audit violations can solve the level-of-analysis problem that plagues most country-level perceptions data: that is, the fact that within-country variation often overwhelms national-level variation (Gingerich, 2013a, 505, 538). Aside from equivalence considerations,¹³ the use of subnational data also tends to facilitate natural experiments and causal inferences—as opposed to correlational analyses with regression (Stanig, 2018, 150-151).

Finally, audits bring unique advantages in terms of cost and research feasibility. The fact that so many countries undertake audits on a routine basis means that researchers do not need to design custom-made, context-specific corruption measures or expensive field experiments (e.g., Olken, 2007; Duflo et al., 2013, 2018). As I explain below in further detail, not all audit data are perfectly comparable across countries, so audits are not a perfect substitute to perceptions data. However, the fact that researchers can obtain data from each respective country's audit agency provides some form of standardization for measuring corruption within countries.

¹³ For more on equivalence, see Locke and Thelen (1995) and Stegmueller (2011).

2.2. Brazilian Audit Innovations and External Validity Issues

On the subject of using audits to measure corruption within countries, Ferraz and Finan's (2008) study of randomized municipal audits of federal transfer funds in Brazil pioneered an entire research agenda.¹⁴ From the perspective of measurement, these objective data on corrupt acts like procurement fraud or over-invoicing from Brazil measure actual corruption, not perceptions of it (Ferraz and Finan, 2008, 710). Also, due to the fact that the audits are randomly assigned, they obviate the many endogeneity concerns that corruption typically engenders (Escaleras, Anbarci and Register, 2007). In turn, from the perspective of knowledge creation, these data enabled Ferraz and Finan (2008) to conclusively show—for the first time—that information about politicians' corruption levels can hurt their re-election prospects. Other important results stemming from the Brazilian audit infractions data include that: re-election incentives condition politicians' corruption levels (Ferraz and Finan, 2011); getting audited nudges subnational government entities to reduce their future corruption levels—although effects are sectorally heterogeneous (Avis, Ferraz and Finan, 2018; Zamboni and Litschig, 2018); anti-corruption audits spur beneficial economic activity (Colonnelli and Prem, 2021); malfeasance revelations affect the types of candidates that parties put forth on party lists (Cavalcanti, Daniele and Galletta, 2018); and female politicians are less corrupt than their male counterparts (Brollo and Troiano, 2016).

The above studies are all very precisely estimated with objective data, but relying only on one specific program in one country over and over again raises the familiar external validity problem: that is, whether results generalize to a broader population or transport to other target populations (Findley, Kikuta and Denly, 2021). The Brazilian anti-corruption program, in particular, yields significant external validity challenges: its annual audits only cover federal funds in 1% of municipalities, which must be non-capitals with fewer than 500,000 inhabitants (see Ferraz and Finan, 2018; Colonnelli and Prem, 2021). More broadly, Brazil is

¹⁴ In this article, I focus on audits. However, there were many other innovations to measuring corruption using, for example, objective data from procurement (e.g., Broms, Dahlström and Fazekas, 2019; Fazekas and Kocsis, 2020), asset declarations (e.g., Eggers and Hainmueller, 2009; Fisman, Schulz and Vig, 2014), and taxes/duties (e.g., Rijkers, Baghdadi and Raballand, 2017; Naritomi, 2019).

the only country that randomly audits administrative units,¹⁵ so the political and other biases that randomization helps solve in Brazil are likely relevant in other countries. Similarly, most countries' audit agencies need to use discretion to choose where to audit through risk-based audits: high-corruption areas or ones that receive whistleblower complaints need more attention than low-corruption areas. Essentially, the context of the Brazilian anti-corruption program is rather specific, so scholars need to move beyond Brazil and examine the cause and consequences of corruption with audit data from other settings.

3. New Audit-Based Corruption Data and A Framework for Assessing Them

There is a dearth of political corruption studies employing audit data outside of Brazil.¹⁶ One reason likely pertains to the lack of easily-accessible audit data. A second reason is that scholars do not have systematized ways of discerning the validity of audit data, especially when audits are not randomly assigned. In the absence of randomly assigned audits, it is feasible that presidents or prime ministers can use their control of the executive branch, which includes the bureaucracy, to unfairly audit political opponents. The risk is particularly high in poor, developing democracies and countries with low horizontal accountability¹⁷—that is, exactly the types of countries where corruption tends to be high as well.

In this section, I attempt to address both of the above gaps in the literature. First, I provide a new, three-pillar framework for assessing the utility of audit data. The framework is especially useful for when audits are not randomly assigned, but all three of its pillars are

¹⁵ Audits agencies across the world, however, randomly audit individual taxpayers and firms.

¹⁶ Bobonis, Cámara Fuertes and Schwabe (2016), Bo, Wu and Zhong (2020), Larreguy, Marshall and Snyder (2020), and Ajzenman (2021) are, to my knowledge, the only such studies outside of Brazil using administrative data—aside from those on one-off programs, such as Di Tella and Schargrodsky (2003), Olken (2007), and Nikolova and Marinov (2017).

¹⁷ Horizontal accountability concerns the ability to keep checks on itself—i.e., the executive branch. For new data on horizontal accountability, refer to Lührmann, Marquardt and Mechkova (2020). For more on the poor performance of poor democracies, see Keefer (2007a,b).

also relevant for experiments or natural experiments that randomly assign audits. Second, I provide a wealth of new, micro-level audit data to demonstrate the utility of the framework and facilitate further audit-based corruption research using these objective data.

3.1. A New Framework

Three pillars underpin my new framework for discerning the utility of audit-related corruption data—regardless of whether they are randomly assigned. The first pillar concerns the legal basis for the political/institutional independence of the audit agency. The second pillar focuses on the frequency, selection, and distribution of audits, and whether they are subject to political, ethnic, or other types of biases. The third and final pillar pertains to potential biases in audit implementation or dosage. I describe each of these pillars in turn.

3.1.1. Pillar 1: Legal Basis for Political/Institutional Independence

For audit agencies to produce credible corruption data, they must enjoy political/institutional independence to pursue unbiased auditing and have a legal basis for it. Political/institutional independence ostensibly has a *de facto* self-enforcing, norms component to it (see [Weingast, 1997](#)). However, it is *de jure* the legal basis for independence that makes an agency's commitment to unbiased auditing credible (see [North and Weingast, 1989](#); [Acemoglu, Johnson and Robinson, 2005](#)). An audit agency's independence is most credible when it enjoys protections from the executive branch and is its own institutional veto player (see [Tsebelis, 2002](#), 2).¹⁸ Of course, audit agencies can be subject to oversight from a partisan veto player (e.g., a legislature) or another bureaucratic agency as part of a larger system of horizontal accountability (see [Ferraz and Finan, 2018](#)). However, oversight institutions cannot unduly influence auditing outcomes if scholars are to use the respective data to accurately measure corruption.

¹⁸ [Tsebelis \(2002, 2\)](#) defines a veto player as the actor(s) who must agree for a policy to change. Institutional veto players generally enjoy protections from a country's constitution.

As highlighted by INTOSAI (2019) and as suggested by theory (e.g., North and Weingast, 1989; de Figueiredo and Weingast, 2005), the strongest legal provisions regarding audit agency independence are those in a country's constitution. In particular, an independent audit agency will have clear rules and tenure security for the agency head, budget autonomy, and discretion to audit without securing approval from a small group of political veto players. For example, if a country's President or Prime Minister can easily fire and replace the audit agency head at will, it will be difficult for the audit process to be unbiased. The same is true for budgets: if audit agencies receive smaller outlays after undertaking unpopular audits, then independence becomes more elusive. That is why there is greater risk if audit agencies are subject to the whims of a single president or prime minister. By contrast, there is less possibility for a single individual to unduly influence audit agency independence when it reports to a partisan veto player such as a legislature, which has a larger group of individuals overseeing it (see Tsebelis, 2002, 2).

3.1.2. Pillar 2: Frequency, Selection, and Distribution of Audits

Pillar 2 of the framework concerns the frequency or selection of audits, which must not exhibit biases toward any particular group. Among these potential biases, ethnic, hometown, and political rivalry biases are likely most relevant (Chu et al., 2021; Seim and Robinson, 2020). Political rivalry biases can refer to any faction or individual with interests that do not coincide with a country's veto players or electorate. However, in most countries, political rivalry biases pertain to opposition party politicians and their districts. That is especially the case when these opposition party politicians win their elections in close races.

Although many observers may argue that random assignment immediately solves this second pillar, some forms of random assignment are more credible than others. Complete random assignment,¹⁹ block randomization, or a well-justified, stratified sampling strategy²⁰

¹⁹ In complete random assignment, units are randomly assigned into equally-sized treatment and control groups.

²⁰ In stratified random sampling, the researcher devises a strategy to sample

can help meet the pillar for experimental studies. Simple randomization²¹ is generally sufficient but less credible.

For observational studies to meet this second pillar on audit distribution fairness, quantitative analysis is necessary. If all units (e.g., states, municipalities) in a sample do not receive the same number of audits for the given time interval, then it is necessary to estimate the conditional mean of the number of audits received.²² When doing so, it is useful to take into account demographics such as population (more populated places will generally have more vulnerable revenue), past corruption (either lags or cumulative sums), and revenue amounts.

Regardless of what estimations of the conditional means show, it is necessary to complement these analyses with a regression discontinuity design (RDD) when examining partisan bias. Specifically, the latter needs to follow the setup of [Brollo and Nannicini \(2012\)](#) and examine whether co-partisanship or party alignment between the executive and lower-level government units predicts audit frequency or selection after close elections. The RDD not only allows for quasi-causal estimation but is particularly relevant because countries' presidents may wish to use their control of the bureaucracy to target audits at electorally vulnerable opposition party politicians. If either the conditional mean estimates or those of the RDD consistently demonstrate targeting at a particular group, then it is clear that the data are biased.

3.1.3. Pillar 3: Implementation/Dosage

The framework's third pillar on implementation/dosage concerns whether all units in the sample—municipalities, politicians, etc.—receive audits that are similarly stringent. If, for example, politicians aligned with the president receive less stringent audits than those

²¹ By simple randomization, I mean independent random draws (think: coin flips) to decide whether a unit is assigned to the treatment or control group. In simple randomization, there is no guarantee that the treatment and control groups are equally-sized.

²² It is possible to estimate the conditional mean via numerous methods. I provide examples in Section 3.2 below.

given to opposition party politicians, then it is difficult to assert that the audit process is fair. By the same token, it is necessary to note that some more stringent types of audits, such as forensic audits, may be necessary to complement financial or compliance audits under some circumstances (e.g., a whistleblower complaint).

In many ways, the third criterion on implementation/dosage is the hardest to verify and might be the easiest to ignore. Not coincidentally, I did not find a single published study that attempted to examine audit implementation/dosage in detail. One reason is likely that not every country will have the requisite quantitative data on the number of auditors sent to a place, audit duration, the amount of money of audited as percent of the total, etc.

Qualitative data can be useful to assess audit implementation/dosage as well. In fact, given that quantitative data on audit implementation/dosage will not always be accessible, qualitative data will often provide the only recourse to assess audit implementation/dosage. Particularly relevant qualitative data include those related to professionalism (e.g., in hiring) and auditor codes of ethics/conduct (Gustavson and Sundström, 2018). The latter are most credible when their scope is quite broad, especially regarding conflicts of interest, and there are frequent internal or external audits with consequences for auditor noncompliance.²³ Some very skilled qualitative researchers may also undertake interviews of auditing staff, but adequately convincing readers that social desirability bias²⁴ is not driving auditors' responses may be challenging. Ethnographic studies of auditors may be the most credible qualitative method to assess audit implementation/dosage. However, the skills required to avoid Hawthorne effects²⁵ may be prohibitively high for non-ethnographers.²⁶

²³ External audits are especially important in instances of high state capture (Denly et al., 2019).

²⁴ Social desirability bias refers to when respondents are unwilling to admit to socially unacceptable behavior and thus provide inaccurate responses to enumerators.

²⁵ Hawthorne effects refer to the when participants change their behavior because they know that they are under study.

²⁶ The extensive set of skills and techniques required to undertake a credible ethnographic study fall outside this scope of the present one.

3.2. New Subnational Audit Data

3.2.1. Case Selection and Overview

Table 1: Summary of Original Data Collected

	New Data (My Contribution)				Others
	Honduras	Guatemala	Mexico	India	Brazil
Administrative Unit	Municipal	Municipal	Municipal	State	Municipal
Funds Covered	All	All	Federal only	All	Federal only
Years Covered	2002-2018	2004-2018	2007-2018	2004-2018	2003-2015
Number of Audits	900	3,350	3,211	1,100	2,200
Audited Annually %	21%	88%	11%	100%	1%
Sector	No	No	Yes	Yes	No
Sub-sector	No	No	No	Yes	No
Infractions	Yes	Yes	Yes	Yes	Yes
Details of Infractions	Yes	No	Yes	Partial	Yes
Money Stolen/Missing	No	Yes	Yes	Yes	Yes
Follow-up on Audit Recommendations	No	No	No	Yes	No
Whistleblower Complaints	No	Yes	No	No	No
Charges Filed	No	Yes	Yes	No	No
Money Audited	No	No	Yes	Partial	Yes
Money Outside Audit	No	No	Yes	No	Yes

Note: The codebooks for Honduras (Appendix A), Guatemala (Appendix C), Mexico (Appendix D), and India (Appendix B) contain more details on the new data. [Ferraz and Finan \(2018\)](#) explain the Brazil data.

I systematically selected the country cases for this study, focusing on members of the International Organization of Supreme Audit Institutions (INTOSAI), which has 195 full member countries as of 2021. I focused on INTOSAI members because it has specific provisions in its Mexico and Lima Declarations regarding audit independence and data transparency for its member countries ([INTOSAI, 1977, 2007](#)). By the same token, most INTOSAI member countries do not make their subnational audits easily accessible to citizens or researchers. For example, searches and emails to all of the SAIs from English-, French-, and Spanish-speaking African countries only yielded a limited amount of subnational audit data from Kenya and South Africa.²⁷ Against this backdrop of extremely limited data availability, I selected data

²⁷ The sample size of the Kenyan subnational audit data was both small and not available for multiple

from the SAIs of Guatemala (2004-2018), Mexico (2007-2018), Honduras (2002-2014), and India (2000-2017). Mexico, Honduras, and Guatemala correspond to the Spanish-speaking Latin American countries that make their data relatively easy to find. India is an Asian country that publishes its data in English.

Institutionally, Guatemala, Honduras, and Mexico are presidential democracies, whereas India is parliamentary democracy. Although Guatemala, Honduras, and Mexico are all located in Central America, they exhibit meaningful variation. For example, only Mexico has political variation at the state level, because the presidents of Guatemala and Honduras appoint governors from their own parties. Additionally, Mexican mayors were not eligible for reelection for the period under study (see [Motolinia, 2021](#)).

Table 1 summarizes the original audit data that I collected and cleaned with the help of a large team of research assistants. The audit data from Guatemala, Honduras, and Mexico correspond to the municipal level, whereas those from India correspond to the state level. The scope of the Guatemalan, Honduran, and Indian audits is quite broad, covering all potential expenses. By contrast, the Mexican municipal audits examined in this study are more limited in scope.

3.2.2. Mexican Audit Data

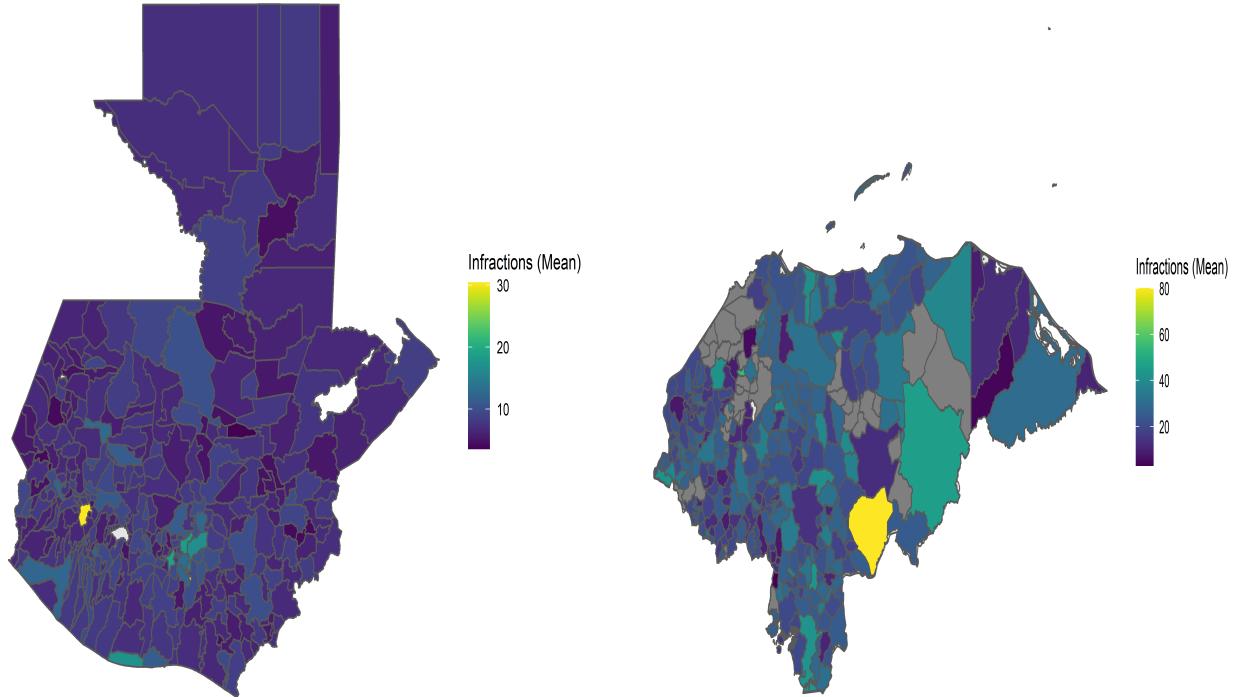
Each municipality in Mexico can receive an audit from either the federal supreme audit agency (ASF, *la Auditoría Superior de la Federación*) or the respective state-level auditing entity (EFSE, *las Entidades de Fiscalización Superior Estatales*).²⁸ However, the EFSEs often lack independence from powerful state governors, consistent budgets, standardized procedures, and the ability to trigger disciplinary proceedings, so the effectiveness EFSE audits in combating corruption is limited ([OECD, 2017](#); [Zachary and Spaniel, 2020](#)). Accordingly, the present study relies on the more neutral ASF audits (see also Section 4.4).

electoral periods, so I did not include these data in this study. South Africa released more of its audit data on its website during the period in which this study was completed.

²⁸ EFSEs are also called *las Entidades de Fiscalización Superior Locales* (EFSLS).

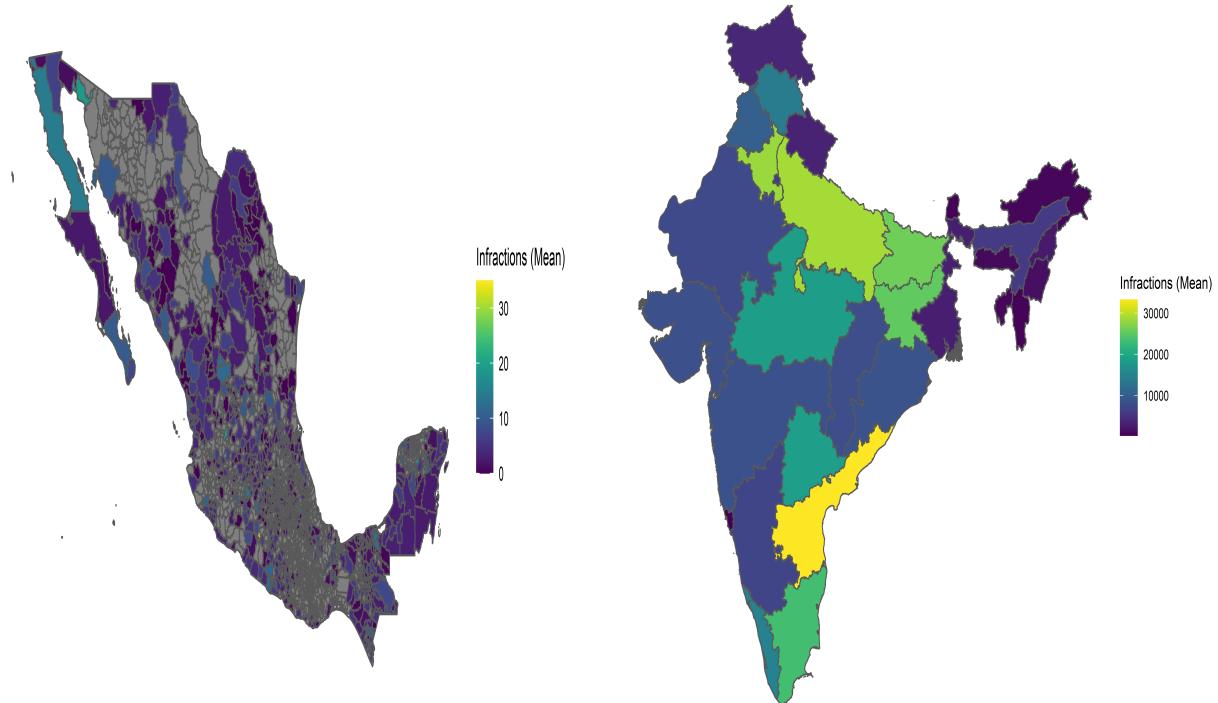
Figure 1: Maps of Subnational Corruption Data by Administrative Unit

(a) Mean Infractions in Guatemala, 2004-2019 (b) Mean Infractions in Honduras, 2002-2018



(c) Mean Infractions in Mexico, 2007-2018

(d) Mean Infractions in India, 2004-2018



Note: In India, infractions refer to Observations. Gray space in Mexico indicates that no audit was undertaken, and gray space in Honduras indicates that no audit was undertaken or no data were available.

Similar to the randomized Brazilian municipal audits first examined by [Ferraz and Finan \(2008\)](#) that now comprise the basis for much knowledge on patterns of local-level corruption, the ASF audits only pertain to the municipal use of federal funds. Table D2 provides a breakdown of these audits by fund type, and Figure 1c shows the territorial reach of these audits—with gray areas indicating that the municipality did not receive an audit. To date, existing studies using these data by [Larreguy, Marshall and Snyder \(2020\)](#) and [Ajzenman \(2021\)](#) focus exclusively on audits of the Fund for Municipal Social Infrastructure (FISM, *el Fondo de Infraestructura Social Municipal*), which comprise circa 26% of the audits in the present dataset.

Although the ASF audits in this dataset only pertain to the municipal use of federal funds, the ASF has a great diversity of audits and takes many different types of actions after these audits (see Tables D3 and D4). Appendix D provides a full codebook of the Mexican audits dataset.

3.2.3. Guatemalan Audit Data

The Guatemalan audit data cover all municipality expenses for circa 320 of the country's 338 municipalities from 2004-2019. As Table 1 shows, the audits yield data on infractions (*sancciones*), whistleblower complaints (*denuncias*), and charges filed (*informes/pliegos de cargos*). For each of these categories, the data provide not just a count of these categories but also the amount of stolen/misappropriated money associated with them, which I deflated to constant 2013 Quetzales. Appendix C provides a full codebook of these data.

3.2.4. Honduran Audit Data

Unlike the other SAIs described above, Honduras' SAI (*El Tribunal Superior de Cuentas*) does not provide detailed lists of audit infractions. Instead, it only provides the actual municipal audits reports. Accordingly, a large team of research assistants and I inspected

each infraction in every available report for corruption using a typology loosely based on the World Bank’s (2016) Anti-Corruption Guidelines (see Table A2). Given that the World Bank’s (2016) Anti-Corruption Guidelines are primarily designed for sanctioning companies, we supplemented their primary categories of bribery, collusion, fraud, obstruction, and coercion with the additional categories of theft, nepotism, cronyism, and grand corruption (see Table A2).

Not all mismanagement is corruption, though. To distinguish between corruption and mismanagement, the team and I only classified instances under any of the above categories of corruption when there was clear intent from the perpetrators. Especially in developing country contexts, bureaucrats often lack training and equipment (e.g., computers), and these impediments can lead to mismanagement and clerical errors that are distinct from corruption.²⁹ Against this backdrop, the dataset entails a count of the total number of infractions as well as the number of corrupt infractions meeting the definitions in Table A2. Appendix A provides further details about these data and their coding procedure.

3.2.5. Indian Audit Data

In contrast to the SAIs of Guatemala and Mexico, the Comptroller and Auditor General (CAG) of India does not provide comprehensive audit information at the municipality or Gram Panchayat level each year. However, the CAG does make available a wealth of panel audit data at the state level. These data notably include disaggregated information at the sector, receipt, and department levels (see Table 1).

The main challenge with the Indian audit data is that, similar to Honduras, the CAG does not provide the data in tabular format. A large team and I thus extracted the data from the individual audits reports, which usually span hundreds of pages in length. As

²⁹ In interviews that I conducted in Honduras in 2016, I learned from comptrollers in multiple municipalities that they often had to drive somewhere else to enter required audit information into computers connected to the Internet. In turn, the cumbersome process also yielded clerical errors that would sometimes show up in audits as accounting inconsistencies.

the Codebook in Appendix B describes, the data contain information on the number of infractions and amounts of stolen/misappropriated money. The CAG disaggregates these data into larger Inspection Reports (IR) and more specific paragraphs (infractions) within those IRs.

Perhaps of most interest to researchers is that the data include information on the extent to which state governments resolve the infractions and return the stolen/misappropriated money from year-to-year (see Appendix B). Past research already indicates that audits are perhaps the most effective tool in reducing corruption (Gans-Morse et al., 2018), but these new data will help researchers better understand the conditions under which audits themselves are most effective at reducing corruption. Such information will be crucial for advancing scholarly knowledge about corruption, because anti-corruption tools are only effective if corrupt actors take remedial action in response to them.

4. Analysis of the Country Cases

In this Section, I apply the framework in Section 3.1 to analyze the extent to which researchers can profitably use the non-randomized, subnational audit data from India, Mexico, Guatemala, and Honduras. As above, I analyze the legal independence of the audit agency from the executive branch, the partisan distribution of the audits, and their implementation/dosage.

4.1. India

4.1.1. Legal Basis for Political/Institutional Independence (India)

There is a plethora of legislation that legally protects the independence of the CAG and its Auditor General, who serves a term of six years. The relevant legislation protecting the Auditor General from political interference dates back to Audits and Accounts Order of

1936, which the Companies Act of 1956 and Duties, Powers, and Conditions (DPC) Act of 1971 reinforce. Articles 148-151 of the Indian Constitution further protect the independence of the CAG (see [Elkins et al., 2014](#)). Although the president appoints the Auditor General after nomination from the Prime Minister, the Parliament determines the salary and office requirements of the Auditor General. Additionally, the process for removing the Auditor General is akin to that of a Supreme Court Justice, and the Auditor General is not eligible for further political office after his or her term at the CAG ends. Given that the constitution grants the Auditor General even more autonomy in terms of day-to-day operations, it is clear that the Auditor General has significant legal independence from the executive branch to carry out unbiased auditing.

4.1.2. Frequency, Selection, and Distribution of Audits (India)

Each year, India's CAG publishes audit reports for all of India's 29 states on its website. Accordingly, analysts can use these state-level audits data from India without concern of distributional bias along political, ethnic, or hometown lines.

4.1.3. Implementation/Dosage (India)

CAG audits take place on a fiscal year timetable that starts in April and ends the following in March. Because all audits follow this timetable, there is no obvious concern regarding the potential time implementation/dosage of the audits.

There are similarly no ostensible concerns regarding the scope of implementation/dosage of the audits. The CAG also has an extensive auditor code of ethics, emphasizing integrity, independence, objectivity, impartiality, confidentiality, and competence as key ethical principals ([Comptroller and Auditor General of India, 2012](#)). When combined with the similarly broad CAG Regulations on Audits and Accounts, Auditing Standards, and Audit Quality Management Framework ([Comptroller and Auditor General of India, 2015, 2017, 2020](#)), it

is clear that professional auditors can conduct audits to their full extent without issues of missing information or other hindrances.

4.2. Honduras

4.2.1. Legal Basis for Political/Institutional Independence (Honduras)

The Supreme Tribunal of Accounts (TSC, *Tribunal Superior de Cuentas*) is the institution responsible for government audits in Honduras. Articles 205, 222-227, and 240 of the Honduran Constitution and the Organic Law of the TSC (2002, revised 2011) provide the legal basis of the TSC (see [Elkins et al., 2014](#)). Overall, its legal basis is strong. Notably, the TSC reports to Congress, not the President, so the chance that any one individual can act as a veto player for the TSC is low. That is especially the case because the Organic Law of the TSC specifically stipulates that the TSC head is not eligible for reelection after serving the initial term; staff are prohibited from political participation, except voting; and numerous politicians, including the President, are not eligible to be a member of the TSC ([El Congreso Nacional de Honduras, 2002](#)).³⁰ From an operational autonomy perspective, the TSC's independence is similarly robust. The Organic Law of the TSC supersedes all other laws, except those in the Constitution, so the TSC can engender compliance with its operations, and the Organic Law grants the TSC broad scope to do so ([El Congreso Nacional de Honduras, 2002](#)).

The one area where the autonomy of the TSC is weaker pertains to its budget. Per the Law of Municipalities, the TSC is supposed to receive a budget akin to one percent of municipal revenues in the previous year ([El Congreso Nacional de Honduras, 2011](#)). In 2011, the government revised the Organic Law of the TSC in order to allow for contributions from

³⁰ Other politicians who are not eligible to be part of the TSC include the Secretaries of State, the members of the Board of Directors of the National Congress, the Judges of the Supreme Court of Justice, the Attorney General of the Republic, the members of the National Elections Court, the Superintendent of Concessions and Licenses, the Attorney General of the Republic, and the National Commissioner of Human Rights ([El Congreso Nacional de Honduras, 2002](#), Article 16).

various other sources as well, including foreign aid. However, according to author interviews with various mayors as well as a TSC transparency request response, the budget is a recurring a challenge in Honduras more broadly, and the TSC is no exception.

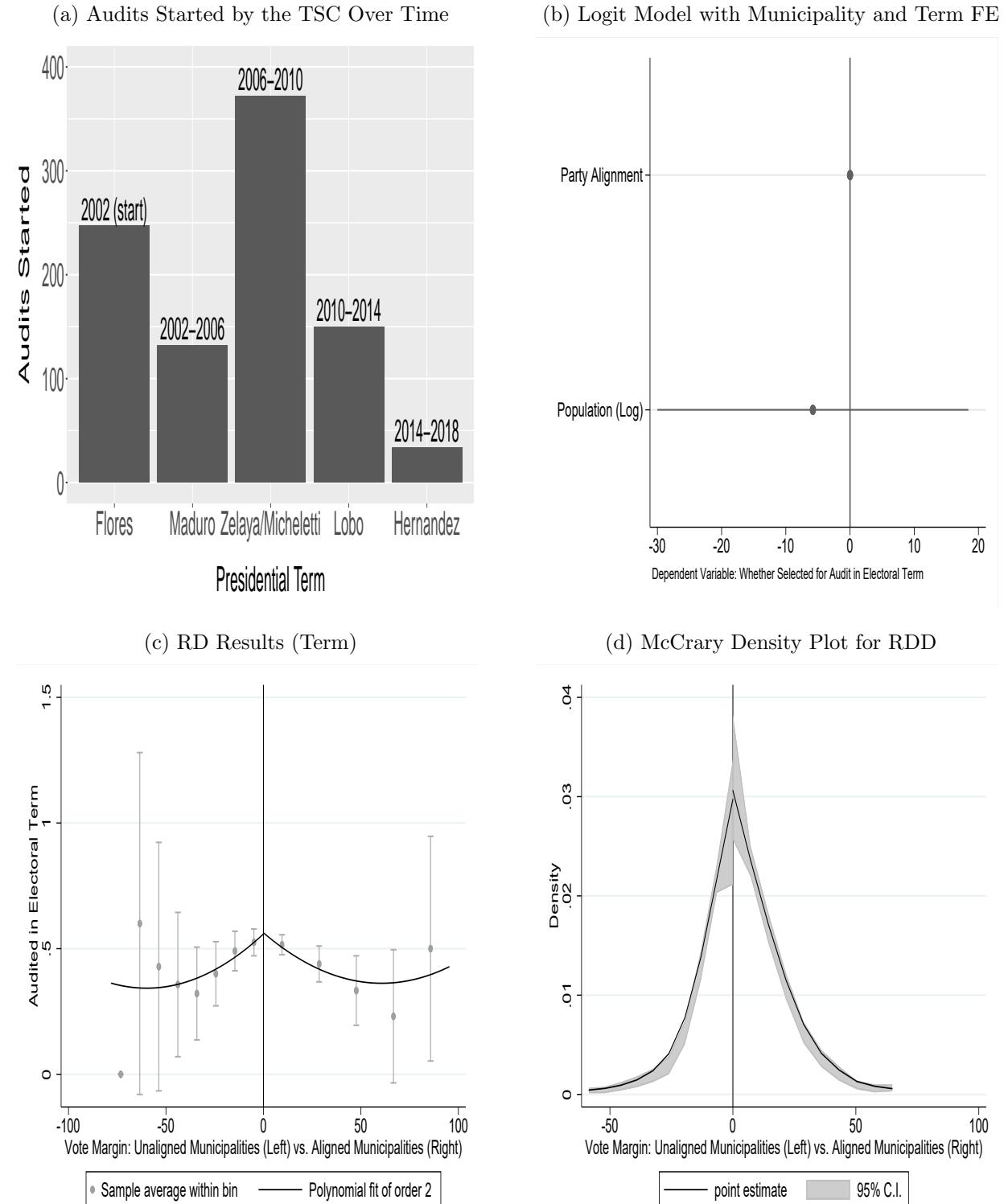
4.2.2. Frequency, Selection, and Distribution of Audits (Honduras)

The TSC performs numerous audits of both municipalities and state-owned enterprises each year. However, as compared to the SAIs from other countries examined in this paper, the TSC undertakes relatively few municipal audits. Honduras has 298 municipalities but only in 2002—when the TSC became a formal SAI—did the country start more than 200 audits in a year (see Figure 2a). Since then—and particularly over the course of Juan Orlando Hernandez’s two-term presidency (2014-present)—the number of municipal audits undertaken by the TSC has dropped steadily. A caveat is that many audits take place over the course of multiple years, and the final reports from the audits only become available years after the audits ends. Accordingly, the number of audits during the Hernandez Presidency will increase slightly as the TSC releases more multi-year audits reports on its website. Nevertheless, Honduras’ relative paucity of municipal audits compared to Guatemala and Mexico is still significant.

Partisanship. In terms of the relationship between political rivalry and auditing decisions, the logit model in Figure 2b shows that partisan motivations are not driving which municipalities receive audits each electoral (presidential) term. I focus on results by electoral term, not the year, given both the low audit frequency and the fact that Honduras’ General Elections take place every four year on the same timetable for most positions. I also only examine mayor-president party alignments because Honduras’ president appoints governors from his/her own party, so there is no relevant political variation at the department (state/province-equivalent) level.

Consistent with the framework detailed in Section 3.1, I run a regression discontinuity design along the lines of [Brollo and Nannicini \(2012\)](#) as well. It uses random variation

Figure 2: Main Results: Honduras' Municipal Audit Distribution



in close elections—using an automatically derived bandwidth following Calonico, Cattaneo and Titiunik (2014)—to assess whether party alignment between the president and mayors is driving auditing decisions. Overall, I find the same pattern in Figure 2b and Figure 2c: party alignment does not affect auditing decisions. As shown in Figure 2d, the regression discontinuity analysis passes the McCrary (2008) density test, too, so random variation in close elections is smoothly distributed, and there are no signs of electoral fraud that could skew the results.

Ethnicity. The inclusion of ethnicity data by municipality in the 2013 Honduran census allows for the testing of whether municipalities with higher percentages of indigenous peoples are more likely to receive an audit. As Appendix F.1 demonstrates, such municipalities are indeed more likely to receive an audit. By the same token, interviews with Honduran government officials consistently indicated that areas with higher shares of indigenous populations tend to have lower capacity and were less likely to submit required paperwork on time.³¹ It is thus feasible that these municipalities with higher shares indigenous populations have higher corruption risks, making the ASF’s risk-based auditing decisions justifiable—especially in the context of a limited budget.

4.2.3. Implementation/Dosage (Honduras)

Unlike the audit data from other countries examined in this paper, Honduras’ TSC does not have fixed timetables for its audits, but it does provide start and end dates for each audit. It is thus possible to analyze whether party alignment affects whether opposition municipalities receive audits that are longer and, presumably, more stringent. A potential challenge is that the start and end dates often only differ by one day, and the audits often last for multiple years, indicating that only the audit start dates are meaningful for analysis.

Irrespective of these potential challenges, Table 2 examines potential partisanship-related duration biases in audit intensity through Cox, Exponential, Weibull, Gompertz,

³¹ These interviews were conducted prior to undertaking the relevant regression analyses.

Table 2: Survival Analysis: Does Party Alignment Affect the Implementation of Audits?

	(1) Cox	(2) Exponential	(3) Weibull	(4) Gompertz	(5) Log-Normal	(6) Log-Log
Party Alignment	0.440 (0.316)	0.587 (0.381)	0.497 (0.347)	0.470 (0.332)	-2.632 (1.957)	-3.170 (2.260)
Population (log)	0.505*** (0.0961)	0.607*** (0.131)	0.546*** (0.111)	0.522*** (0.103)	-3.869*** (0.647)	-3.959*** (0.743)
Constant		-19.22*** (1.662)	-11.63*** (1.437)	-10.24*** (1.334)	79.08*** (6.620)	79.60*** (8.312)
p (log)			-1.907*** (0.0275)			
Gamma				-0.953*** (0.0201)		
Sigma (log)					2.565*** (0.0670)	
Gamma (log)						1.854*** (0.0477)
Constant		-19.22*** (1.662)	-11.63*** (1.437)	-10.24*** (1.334)	79.08*** (6.620)	79.60*** (8.312)

Clustered, robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Note: All specifications contain term fixed effects.

Log-Normal, and Log-Logistic regression. I use these various functional forms because it is too difficult to theorize about the correct functional form of survival models (Blossfeld, Golsch and Rohwer, 2007). Regardless of what the most accurate functional form may be, all of the models yield the same result: partisanship does not affect the duration of audits, as evidenced by the lack of statistical significance and coefficient for party alignment switching between positive and negative. Appendix E provides further technical details behind the analysis.

On the qualitative side, the TSC has a very comprehensive staff code of ethics, containing hundreds of articles regarding professionalism, conflict of interests, bribery, nepotism, collusion, and other impediments to unbiased auditing (Tribunal Superior de Cuentas de Honduras, 2018). All violators of the code of ethics need to answer to the TSC's Probitry and Ethics Committee, which has the authority to conduct investigations. These investigations can also be quite broad given that the code of ethics specifically references the Code of

Ethical Conduct of Public Servants, the Code of Conduct Regulations Ethics of the Public Servants, and the Regime of the Career of the Officials and Employees of the TSC ([Tribunal Superior de Cuentas de Honduras, 2018](#)).

4.3. Guatemala

4.3.1. Legal Basis for Political/Institutional Independence (Guatemala)

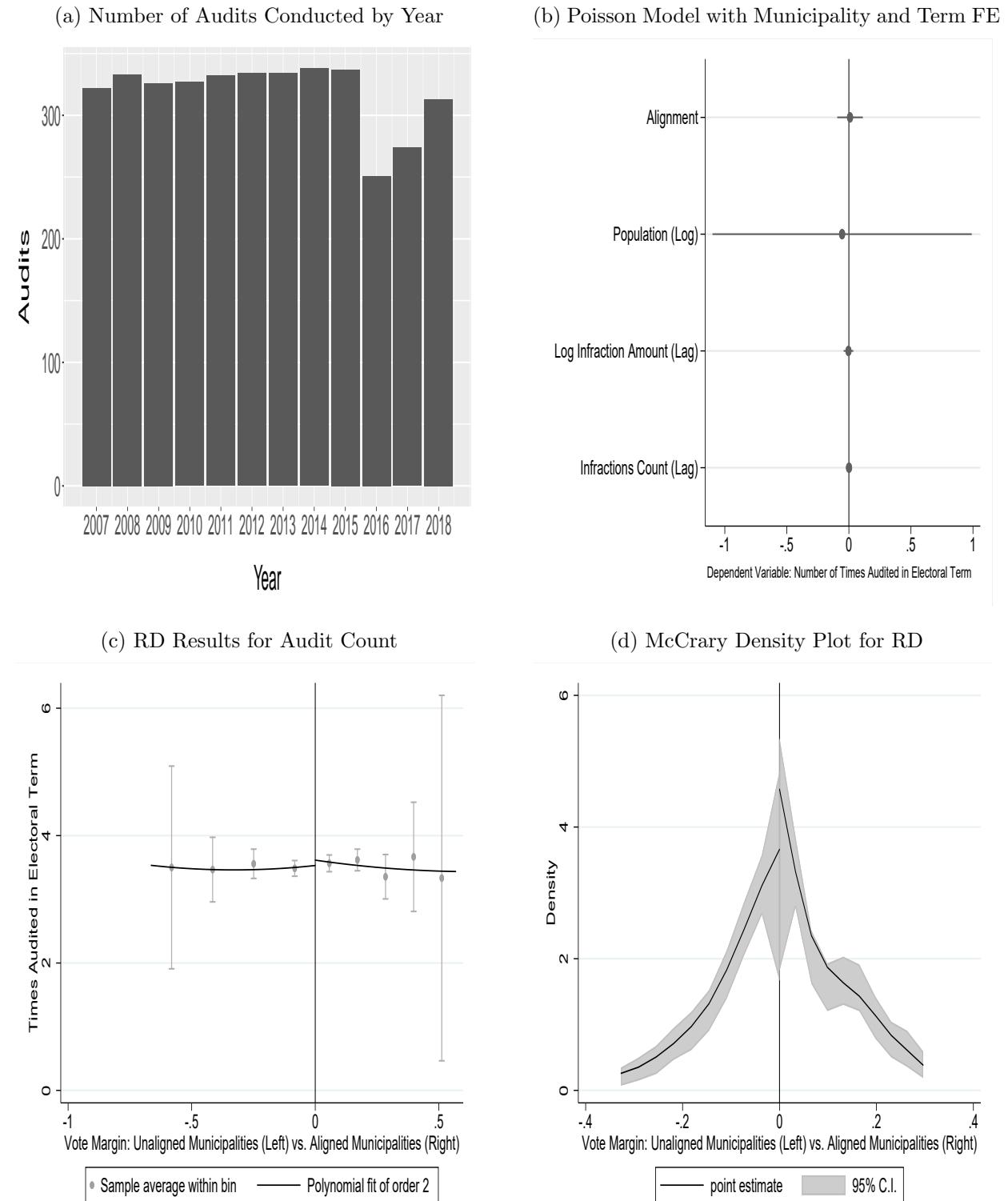
Guatemala's SAI, the Comptroller General of Accounts (CGC, *Contraloría General de Cuentas*), is the institution responsible for audits in Guatemala. Articles 232-236 of Guatemala's 1985 Constitution provides the basis for the office's independence and ability to audit all uses of public funds throughout the country (see [Elkins et al., 2014](#)). Notably, these articles stipulate that Congress, not the President, elects the Comptroller General of Accounts (*Controlador de Cuentas*) to non-renewable, four-year terms. Removing the Comptroller General of Accounts is also uniquely within the purview of the Congress. It can only remove the Comptroller General of Accounts by majority vote only for reasons pertaining to "negligence, crime, and lack of aptitude." The Organic Law of the CGC further supports the provision elaborated in the Constitution, too ([Contraloría General de Cuentas de Guatemala, 2002](#)).

4.3.2. Frequency, Selection, and Distribution of Audits (Guatemala)

Each year, the CGC audits circa 318 of Guatemala's 340 municipalities (see Figure 3a).³² Because there is not much variation in which municipalities the CGC audits each year, it is more appropriate to assess the potential political bias with a count model such as a poisson model. As shown in Figure 3b, past corruption (lagged infractions and log amounts attached to those infractions), partisanship, and demographics (population) are all similarly poor predictors of how many audits a municipality receives in a given year. Like Honduras,

³² This pattern starts in 2007. From 2004-2006, the Comptroller Generall undertook much fewer audits.

Figure 3: Main Results: Guatemala's Municipal Audit Distribution



Guatemala also does not have political variation at the department (state-equivalent) level, so it is possible to analyze mayor-president party alignments directly. When switching to a regression discontinuity design in Figure 3c, the predictive power of alignment is similarly poor. Because the regression discontinuity estimates pass the McCrary (2008) density test as well (see Figure 3d), it is safe to conclude that the audit distribution is fair.

4.3.3. Implementation/Dosage (Guatemala)

Guatemala's CGC does not provide quantitative data for assessing audit implementation/dosage, but it does have an extensive code of ethics/conduct for all employees (Contraloría General de Cuentas de Guatemala, 2018). It regulates personal and political conflicts of interest, bribery, and favors of any kind. Articles 40, 41, and 73 of the Organic Law further specify the relevant sanctions for staff (Contraloría General de Cuentas de Guatemala, 2002), so it is clear that the CGC has relevant regulations and laws to facilitate unbiased auditing.

4.4. Mexico

4.4.1. Legal Basis for Political/Institutional Independence (Mexico)

Mexico's SAI, the Supreme audit agency of the Federation (ASF, *Auditoría Superior de la Federación*), receives its charter directly from Articles 74, 79, and 113 of the Mexican Constitution (see Elkins et al., 2014). Technically, the ASF is part of the Chamber of Deputies, so it does not report to the President. Despite being part of the Chamber of Deputies, the “Constitution grants the ASF technical, managerial, and functional autonomy” (OECD, 2017, 18). The one more challenging area pertains to the budget, which prevents the ASF from conducting more audits (OECD, 2017).

4.4.2. Frequency, Selection, and Distribution of Audits (Mexico)

The ASF does not audit all of Mexico's 2,400 municipalities each year (see Figure 1c). Instead, the ASF conducts risk-based audits, which are decided upon by the ASF planning commission in consultation with the Chambers of Deputies. According to an analysis by the [OECD \(2017\)](#), the decision process is robust, but it is still useful to undertake relevant quantitative tests of potential biases. In this Latin American context, partisanship is the most relevant bias to examine.

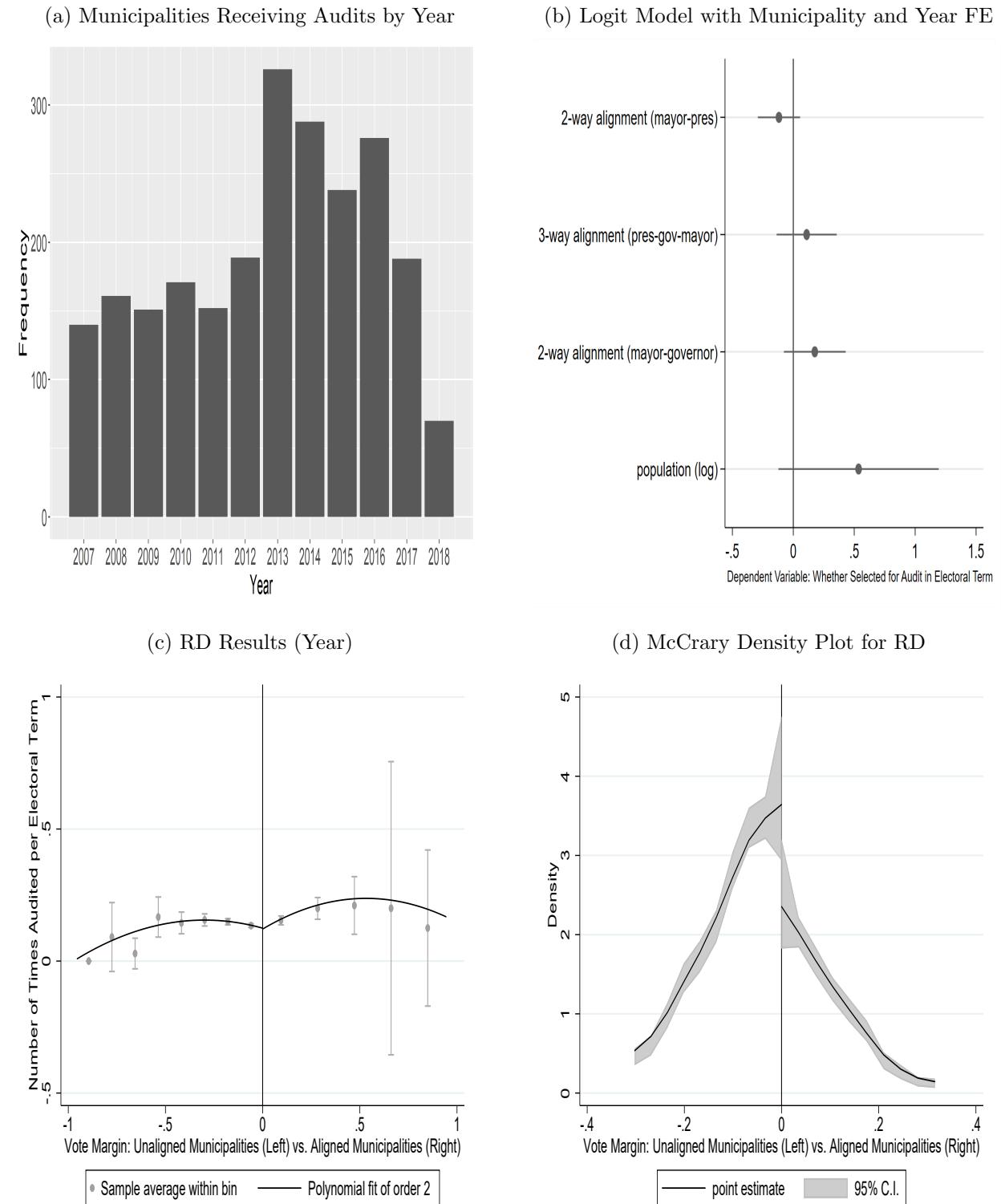
In contrast to Guatemala and Honduras, for which electoral term regressions are more appropriate to discern potential patterns of partisan biases in audits, year-wise regressions are more appropriate for Mexico. The reason is that, at least until 2018,³³ Mexico's mayoral, gubernatorial, and presidential elections did not take place on the same calendar. Essentially, the electoral term is not as meaningful to examine alignment patterns in countries without general elections on the same calendar, which included Mexico until 2018.

Also unlike Guatemala and Honduras, Mexico has governors from different political parties. It is necessary to account for that state-level political variation in quantitative analyses, because it could confound empirical estimates. Against this backdrop, the logit model depicted in Figure 4b controls for all types of coalitional party alignment configurations, including those of mayor-president, mayor-governor, governor-president, and mayor-governor-president.³⁴ Additionally, because of the limited territorial reach of the ASF audits (see Figure 1c), I supplement the municipal-level regressions with ones at the state level. As shown in Figure F.1, none of the alignment configurations predict greater or lower numbers of audits.

³³ For more on the relevant electoral reform, see [Motolinia \(2021\)](#).

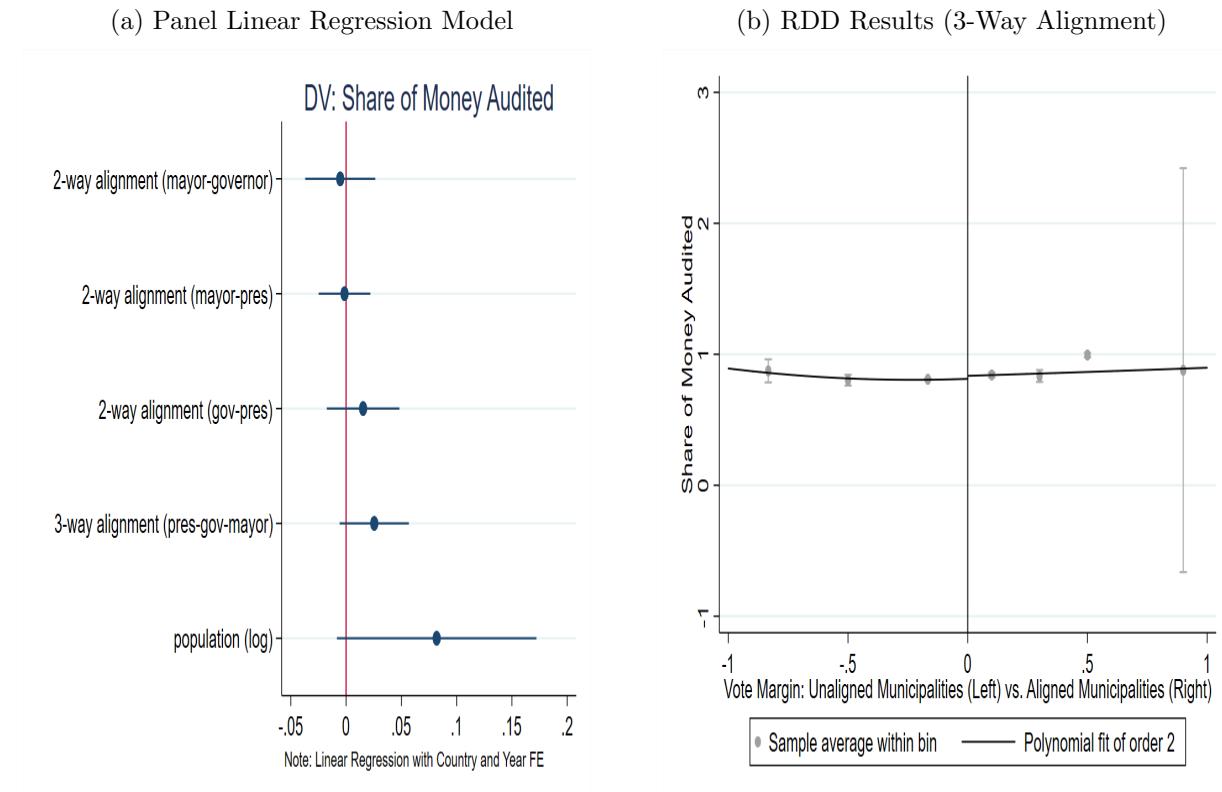
³⁴ All of these alignment patterns are coalitional, not always direct, because many mayors run on multi-party coalitions in Mexico. Following [Benton \(2019\)](#), I assign the alignment status on the basis of whether one party in the coalition is aligned, which is a germane coding decision because the aligned party is generally the most powerful one in each coalition.

Figure 4: Main Results: Mexico's Municipal Audit Distribution



In the corresponding regression discontinuity design model using three-way party alignment in Figure 4c,³⁵ there is a slight jump on the right-side of the plot, indicating that aligned municipalities are more likely to receive an audit. That jump, however, is not statistically significant, and the McCrary (2008) density test in Figure 4d does not pass. It is thus difficult to conclude that there is any partisan bias in the audit distribution.

Figure 5: Mexico's Audit Implementation/Dosage



4.4.3. Implementation/Dosage (Mexico)

Mexico provides credible quantitative data from which to assess the implementation/dosage of audits. For each audit, Mexico's ASF provides the amount of money in Mexican pesos in the audit sample as well as the overall amount of money that it could have audited. If membership in an opposition party predicts a higher share of funds audited, it suggests that

³⁵ Note: results are similar for the other alignment configurations.

opposition parties are receiving more stringent audits. As results from Figure 5 suggest, that does not appear to be the case for Mexico. Both the linear regression model and regression discontinuity approach reach the same conclusion.

Mexico's ASF also has an extensive Integrity Policy, comprising a code of ethics, code of conduct, and directives on conflicts of interest ([Auditoría Superior de la Federación de México, 2013](#)). In future iterations, I will analyze the Integrity Policy in further detail.

5. Discussion and Broader Applicability of the Framework

Table 3: Overview of Results

Country	Independence	Distribution	Dosage	Caveats
Guatemala	Yes	Yes	Yes	
Honduras	Yes	Yes	Yes	Low budget and <i>potential</i> ethnic bias.
India	Yes	Yes	Yes	
Mexico	Yes	Yes	Yes	Limited ASF mandate and budget.

Table 3 provides an overview of the results discussed in the previous section. All of the countries in the sample have sufficient legal independence from the executive branch on paper, thereby fulfilling the first pillar of the framework. The SAIs of Mexico and Honduras, however, have limited budgets, and Mexico's ASF also has a limited remit. Empirical analyses suggest potential ethnic biases in Honduras, though interview evidence suggest that the patterns are likely justifiable given capacity challenges. With respect to the second pillar on the audit distribution, quantitative analysis does not uncover any issues of partisan bias. The framework's third pillar on audit implementation/dosage is perhaps the most difficult to analyze. Nevertheless, the relevant quantitative measures—e.g., audit duration and share of money audited—as well as qualitative analysis of codes of ethics do not suggest any potential biases. By the same token, it is worth noting that the present analysis focuses only on available data, none of the SAIs examined provide accessible data on auditor sanctions

or biases (see Chu et al., 2021),³⁶ and there may be some additional implementation/dosage challenges that only a skilled interviewer or ethnographer could uncover.

As more SAIs release their subnational data and thereby fulfill their INTOSAI membership obligations (see INTOSAI, 1977, 2007), the above framework should prove particularly useful for governmental audit data from other countries with democratic institutions—including South Africa and the Philippines, both of which currently release sufficient audit data for analyses. Single-party authoritarian regimes such as China also conduct subnational audits (e.g., Bo, Wu and Zhong, 2020), and the framework’s pillars relating to independence and implementation/dosage remain relevant for such countries. The same is true for the analysis of state-owned enterprise (SOE) audits. However, future research is likely necessary to discern if additional pillars or considerations are necessary for both audits relating to SOEs and single-party regimes.

6. Conclusion

Audit data do not provide the only objective alternative to measuring corruption with perceptions data. For example, scholars have convincingly measured corruption using data from the stock market,³⁷ procurement,³⁸ asset declarations,³⁹ taxes,⁴⁰ and customs duties.⁴¹ However, audit data standout from the above measures for a simple reason: audits reveal a greater diversity of corrupt activities. Ghost firms, theft, nepotism, fraud, violations related to all of the above measures, and many other types of corruption are all within the purview of audits.

³⁶ Even transparency requests to India and Honduras’ SAIs did not yield any data on auditor sanctions.

³⁷ See, for example, Fisman (2001), Faccio (2006), and Fisman and Wang (2015)

³⁸ See, for example, Bandiera, Prat and Valletti (2009), Mironov and Zhuravskaya (2016), Fazekas (2017), Fazekas, Cingolani and Tóth (2018), Baltrunaite (2019), Broms, Dahlström and Fazekas (2019), and Fazekas and Kocsis (2020).

³⁹ See, for example, Eggers and Hainmueller (2009) and Fisman, Schulz and Vig (2014).

⁴⁰ See, for example, Kleven et al. (2011), Khan, Khwaja and Olken (2016), Bilicka (2019), and Naritomi (2019).

⁴¹ See, for example, Fisman and Wei (2004, 2009), Sequeira and Djankov (2014), and Rijkers, Baghdadi and Raballand (2017).

The challenge with audit-related corruption measures to date is that, with very few exceptions, scholars have focused almost exclusively on the municipal audit data from Brazil. The numerous Brazil studies using these data have produced some very useful causal findings, notably due to the random assignment of audits to individual municipalities. Nevertheless, the external validity of these studies remains a concern due to the unique nature of Brazil's anti-corruption program and the fact that Brazil is the only country in the world with randomized audits.

Overall, the new data and framework advanced in this paper will help analysts undertake sophisticated analyses of corruption across the world. In the process, researchers will use perceptions-based data less, move beyond Brazil, account for (especially political) biases in corruption data and, in turn, better understand the diverse causes and consequences of corruption.

References

- Acemoglu, Daron, Simon Johnson and James A. Robinson. 2005. "Institutions as a Fundamental Cause of Long-Run Growth." *Handbook of Economic Growth* 1(SUPPL. PART A):385–472.
- Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3):529–546.
- Ajzenman, Nicolás. 2021. "The Power of Example: Corruption Spurs Corruption." *American Economic Journal: Applied Economics* 13(2):230–257.
- Andersson, Staffan and Paul M. Heywood. 2009. "The Politics of Perception: Use and Abuse of Transparency International's Approach to Measuring Corruption." *Political Studies* 57(4):746–767.
- Auditoría Superior de la Federación de México. 2013. "Política Institucional de Integridad de la ASF.".
- Avis, Eric, Claudio Ferraz and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126(5):1912–1964.
- Baltrunaite, Audinga. 2019. "Political Contributions and Public Procurement: Evidence from Lithuania." *Journal of the European Economic Association* 18(2):541–582.
- Bandiera, Oriana, Andrea Prat and Tommaso Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review* 99(4):1278–1308.
- Bauhr, Monika and Nicholas Charron. 2018. "Insider or Outsider? Grand Corruption and Electoral Accountability." *Comparative Political Studies* 51(4):415–446.
- Benton, Allyson Lucinda. 2019. "Party Leader or Party Reputation Concerns? How Vertical Partisan Alignment Reins in Subnational Fiscal Profligacy." *Journal of Politics* 81(1):201–214.
- Bermeo, Nancy. 2016. "On Democratic Backsliding." *Journal of Democracy* 27(1):5–19.
- Bersch, Katherine and Sandra Botero. 2014. "Measuring Governance: Implications of Conceptual Choices." *European Journal of Development Research* 26(1):124–141.
- Bilicka, Katarzyna Anna. 2019. "Comparing UK Tax Returns of Foreign Multinationals to Matched Domestic Firms." *American Economic Review* 109(8):2921–2953.
- Blossfeld, Hans-Peter, Katrin Golsch and Götz Rohwer. 2007. *Event History Analysis with Stata*. New York: Psychology Press: Taylor and Francis Group.

- Bo, Shiyu, Yiping Wu and Lingna Zhong. 2020. “Flattening of Government Hierarchies and Misuse of Public Funds: Evidence from Audit Programs in China.” *Journal of Economic Behavior and Organization* 179(71673174):141–151.
- Bobonis, Gustavo J., Luis R. Cámara Fuertes and Rainer Schwabe. 2016. “Monitoring Corruptible Politicians.” *American Economic Review* 106(8):2371–2405.
- Brollo, Fernanda and Tommaso Nannicini. 2012. “Tying Your Enemy’s Hands in Close Races: The Politics of Federal Transfers in Brazil.” *American Political Science Review* 106(4):742–761.
- Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti and Guido Tabellini. 2013. “The Political Resource Curse.” *American Economic Review* 103(5):1759–1796.
- Brollo, Fernanda and Ugo Troiano. 2016. “What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil.” *Journal of Development Economics* 122:28–45.
- Broms, Rasmus, Carl Dahlström and Mihály Fazekas. 2019. “Political Competition and Public Procurement Outcomes.” *Comparative Political Studies* 52(9):1259–1292.
- Buntaine, Mark T., Ryan S. Jablonski, Daniel L. Nielson and Paula M. Pickering. 2018. “SMS Texts on Corruption Help Ugandan Voters Hold Elected Councillors Accountable at the Polls.” *Proceedings of the National Academy of Sciences* 115(26):6668–6673.
- Calonico, Sebastian, Matias D. Cattaneo and Rocío Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82(6):2295–2326.
- Cavalcanti, Francisco, Gianmarco Daniele and Sergio Galletta. 2018. “Popularity Shocks and Political Selection.” *Journal of Public Economics* 165:201–216.
- Chu, Jian, Raymond Fisman, Songtao Tan and Yongxiang Wang. 2021. “Hometown Ties and the Quality of Government Monitoring: Evidence from Rotation of Chinese Auditors.” *American Economic Journal: Applied Economics* 13(3):38.
- Colonelli, Emanuele and Mounu Prem. 2021. “Corruption and Firms.” *Review of Economic Studies* .
- Comptroller and Auditor General of India. 2012. “Code of Ethics for the IAAD Office of the Comptroller and Auditor General of India.” .
- Comptroller and Auditor General of India. 2015. “Audit Quality Management Framework.” .
- Comptroller and Auditor General of India. 2017. “CAG’s Auditing Standards.” .
URL: https://cag.gov.in/sites/default/files/cag_pdf/Auditing_Standards_new.pdf
- Comptroller and Auditor General of India. 2020. “Regulations on Audits and Accounts.” .
- Contraloría General de Cuentas de Guatemala. 2002. “Ley Orgánica de la Contraloría General de Cuentas.” .

Contraloría General de Cuentas de Guatemala. 2018. “Código de Ética de Contraloría General de Cuentas.”.

Coppedge, Michael, John Gerring, Adam N. Glynn, Carl Henrik Knutsen, Staffan I. Lindberg, Daniel Pemstein, Brigitte Seim, Svend-Erik Skaaning and Jan Teorell. 2020. *Varieties of Democracy: Measuring Two Centuries of Political Change*. Cambridge: Cambridge University Press.

de Figueiredo, Rui and Barry R. Weingast. 2005. “Self-Enforcing Federalism.” *Journal of Law, Economics, and Organization* 21(1):103–135.

Denly, Michael, Michael G. Findley, Vepa Rejepov and Rachel L. Wellhausen. 2019. *Integrity Systems and the Rule of Law in Armenia: An Evidence Review for Learning, Evaluation and Research Activity II (LER II)*. Washington, DC: USAID.

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 and Economics* 46(1):269–292.

Duflo, Esther, Michael Greenstone, Rohini Pande and Nicholas Ryan. 2013. “Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India.” *Quarterly Journal of Economics* 128(4):1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande and Nicholas Ryan. 2018. “The Value of Regulatory Discretion: Estimates from Environmental Inspections in India.” *Econometrica* 86(6):2123–2160.

Eggers, Andrew C. and Jens Hainmueller. 2009. “MPs for Sale? Returns to Office in Postwar British Politics.” *American Political Science Review* 103(4):513–533.

El Congreso Nacional de Honduras. 2002. “Ley Orgánica del Tribunal Superior de Cuentas.”.

El Congreso Nacional de Honduras. 2011. “Ley de Municipalidades.”.

Elkins, Zachary, Tom Ginsburg, James Melton, Robert Shaffer, Juan F. Sequeda and Daniel P. Miranker. 2014. “Constitute: The World’s Constitutions to Read, Search, and Compare.” *Journal of Web Semantics* 27-28:10–18.

Escaleras, Monica, Nejat Anbarci and Charles A. Register. 2007. “Public Sector Corruption and Major Earthquakes: A Potentially Deadly Interaction.” *Public Choice* 132(1-2):209–230.

Faccio, Mara. 2006. “Politically Connected Firms.” *American Economic Review* 96(1):369–386.

Fariss, Christopher J., Michael R. Kenwick and Kevin Reuning. 2020. Measurement Models. In *SAGE Handbook of Research Methods in Political Science and International Relations*, ed. Luigi Curini and Robert J. Franzese. London: Sage Publications chapter 20, pp. 353–370.

- Fazekas, Mihály. 2017. "Red Tape, Bribery and Government Favouritism: Evidence from Europe." *Crime, Law and Social Change* 68(4):403–429.
- Fazekas, Mihály and Gábor Kocsis. 2020. "Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data." *British Journal of Political Science* 50(1):155–164.
- Fazekas, Mihály, Luciana Cingolani and Bence Tóth. 2018. Innovations in Objectively Measuring Corruption in Public Procurement. In *Governance Indicators: Approaches, Progress, Promise*, ed. Helmut Anheier, Matthias Haber and Mark Andreas Kayser. Oxford: Oxford University Press chapter 7, pp. 154–185.
- Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's 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.
- Ferraz, Claudio and Frederico Finan. 2018. Fighting Political Corruption: Evidence from Brazil. In *Institutions, Governance, and the Control of Corruption*, ed. Kaushik Basu and Tito Cordella. Cham, Switzerland: Palgrave Macmillan chapter 9, pp. 253–284.
- Ferraz, Claudio, Frederico Finan and Diana B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96(9-10):712–726.
- Findley, Michael G., Kyosuke Kikuta and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* 24:365–93.
- Fisman, Raymond. 2001. "Estimating the Value of Political Connections." *American Economic Review* 91(4):1095–1102.
- Fisman, Raymond, Florian Schulz and Vikrant Vig. 2014. "The Private Returns to Public Office." *Journal of Political Economy* 122(4):806–862.
- Fisman, Raymond and Shang-Jin Wei. 2004. "Tax Rates and Tax Evasion: Evidence from Missing Imports in China." *Journal of Political Economy* 112(2):471–496.
- Fisman, Raymond and Shang-Jin Wei. 2009. "The Smuggling of Art, and the Art of Smuggling: Uncovering the Illicit Trade in Cultural Property and Antiques." *American Economic Journal: Applied Economics* 1(3):82–96.
- Fisman, Raymond and Yongxiang Wang. 2015. "Corruption in Chinese Privatizations." *Journal of Law, Economics, and Organization* 31(1):1–29.
- Gans-Morse, Jordan, Mariana Borges, Alexey Makarin, Theresa Mannah-Blankson, Andre Nickow and Dong Zhang. 2018. "Reducing Bureaucratic Corruption: Interdisciplinary Perspectives on What Works." *World Development* 105:171–188.

- Gerring, John and Strom C. Thacker. 2004. "Political Institutions and Corruption: The Role of Unitarism and Parliamentarism." *British Journal of Political Science* 34(2004):295–330.
- Gerring, John and Strom C. Thacker. 2005. "Do Neoliberal Policies Deter Political Corruption?" *International Organization* 59(01):233–254.
- Gingerich, Daniel W. 2013a. "Governance Indicators and the Level of Analysis Problem: Empirical Findings from South America." *British Journal of Political Science* 43(July 2013):505–540.
- Gingerich, Daniel W. 2013b. *Political Institutions and Party-Directed Corruption in South America: Stealing for the Team*. New York: Cambridge University Press.
- Gisselquist, Rachel M. 2014. "Developing and Evaluating Governance Indexes: 10 Questions." *Policy Studies* 35(5):513–531.
- Greene, Kenneth F. 2007. *Why Dominant Parties Lose: Mexico's Democratization in Comparative Perspective*. New York: Cambridge University Press.
- Gustavson, Maria and Aksel Sundström. 2018. "Organizing the Audit Society: Does Good Auditing Generate Less Public Sector Corruption?" *Administration and Society* 50(10):1508–1532.
- Hollyer, James R. 2018. Measuring Governance: Objective and Perceptions-Based Governance Indicators. In *Governance Indicators: Approaches, Progress, Promise*, ed. Helmut Anheier, Matthias Haber and Mark Andreas Kayser. Oxford: Oxford University Press chapter 5, pp. 103–133.
- INTOSAI. 1977. "The Lima Declaration." .
- INTOSAI. 2007. "Mexico Declaration on Supreme Audit Institutions (SAI) Independence." .
- INTOSAI. 2019. "Independence of Supreme Audit Institutions." .
URL: https://www.intosai.org/fileadmin/downloads/focus_areas/independence/EN_0219_Handout_Independence.pdf
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal on the Rule of Law* 3(2):220–246.
- Keefer, Philip. 2007a. "Clientelism, Credibility, and the Policy Choices of Young Democracies." *American Journal of Political Science* 51(4):804–821.
- Keefer, Philip. 2007b. The Poor Performance of Poor Democracies. In *Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan C. Stokes. Oxford: chapter 36, pp. 886–909.
- Khan, Adnan Q., Asim Ijaz Khwaja and Benjamin A. Olken. 2016. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics* pp. 219–271.

- Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen and Emmanuel Saez. 2011. “Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark.” *Econometrica* 79(3):651–692.
- Kunicová, Jana and Susan Rose-Ackerman. 2005. “Electoral Rules and Constitutional Structures as Constraints on Corruption.” *British Journal of Political Science* 35(May):573–606.
- Kurtz, Marcus J. and Andrew Schrank. 2007a. “Growth and Governance: A Defense.” *Journal of Politics* 69(2):563–569.
- Kurtz, Marcus J. and Andrew Schrank. 2007b. “Growth and Governance: Models, Measures, and Mechanisms.” *Journal of Politics* 69(2):538–554.
- Langbein, Laura and Stephen Knack. 2010. “The Worldwide Governance Indicators: Six, One, or None?” *Journal of Development Studies* 46(2):350–370.
- Larreguy, Horacio, John Marshall and James M. Snyder. 2020. “Publicizing Malfeasance: When Media Facilitates Electoral Accountability in Mexico.” *Economic Journal* 130(631):2291–2327.
- Lederman, Daniel, Norman V. Loayza and Rodrigo R. Soares. 2005. “Accountability and Corruption: Political Institutions Matter.” *Economics and Politics* 17(1):1–35.
- Locke, Richard M. and Kathleen Thelen. 1995. “Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics.” *Politics and Society* 23(3):337–367.
- Lührmann, Anna, Kyle L. Marquardt and Valeriya Mechkova. 2020. “Constraining Governments: New Indices of Vertical, Horizontal, and Diagonal Accountability.” *American Political Science Review* 114(3):811–820.
- Magnusson, Leandro M. and Yashar Tarverdi. 2020. “Measuring Governance: Why Do Errors Matter?” *World Development* 136:105061.
- Mauro, Paolo. 1995. “Corruption and Growth.” *Quarterly Journal of Economics* 110(3):681–712.
- McCrory, Justin. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 142(2):698–714.
- McMann, Kelly, Daniel Pemstein, Brigitte Seim, Jan Teorell and Staffan I. Lindberg. 2021. “Assessing Data Quality: An Approach and An Application.” *Political Analysis*.
- Melo, Marcus André, Carlos Pereira and Carlos Mauricio Figueiredo. 2009. “Political and Institutional Checks on Corruption: Explaining the Performance of Brazilian Audit Institutions.” *Comparative Political Studies* 42(9):1–28.
- Mironov, Maxim and Ekaterina Zhuravskaya. 2016. “Corruption in Procurement and the Political Cycle in Tunneling: Evidence from Financial Transactions Data.” *American Economic Journal: Economic Policy* 8(2):287–321.

- Motolinia, Lucia. 2021. "Electoral Accountability and Particularistic Legislation: Evidence from an Electoral Reform in Mexico." *American Political Science Review* 115(1):97–113.
- Mungiu-Pippidi, Alina. 2015. *The Quest for Good Governance: How Societies Develop Control of Corruption*. New York: Cambridge University Press.
- Naritomi, Joana. 2019. "Consumers as Tax Auditors." *American Economic Review* 109(9):3031–3072.
- Nikolova, Elena and Nikolay Marinov. 2017. "Do Public Fund Windfalls Increase Corruption? Evidence From a Natural Disaster." *Comparative Political Studies* 50(11):1455–1488.
- North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49(4):803–832.
- OECD. 2017. *Mexico's National Auditing System: Strengthening Accountable Governance*. Paris: OECD.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(21):200–249.
- Pereira, Carlos and Marcus André Melo. 2015. "Reelecting Corrupt Incumbents in Exchange for Public Goods: Rouba mas faz in Brazil." *Latin American Research Review* 50(4):88–115.
- Rijkers, Bob, Leila Baghdadi and Gael Raballand. 2017. "Political Connections and Tariff Evasion: Evidence from Tunisia." *World Bank Economic Review* 31(2):459–482.
- Rose-Ackerman, Susan and Bonnie Palifka. 2016. *Corruption and Government: Causes, Consequences, and Reform*. Second ed. New York: Cambridge University Press.
- Sartori, Giovanni. 1984. *Social Science Concepts: A Systematic Analysis*. Beverly Hills: Sage Publications.
- Schaffer, Frederic Charles. 2016. *Elucidating Social Science Concepts: An Interprevist Guide*. New York: Routledge.
- Seim, Brigitte and Amanda Lea Robinson. 2020. "Coethnicity and Corruption: Field Experimental Evidence from Public Officials in Malawi." *Journal of Experimental Political Science* 7(1):61–66.
- Sequeira, Sandra and Simeon Djankov. 2014. "Corruption and Firm Behavior: Evidence from African Ports." *Journal of International Economics* 94(2):277–294.
- Standaert, Samuel. 2015. "Divining the Level of Corruption: A Bayesian State-Space Approach." *Journal of Comparative Economics* 43(3):782–803.

- Stanig, Piero. 2018. Considerations on the Method of Constructing Governance Indices. In *Governance Indicators: Approaches, Progress, Promise*, ed. Helmut Anheier, Matthias Haber and Mark Andreas Kayser. Oxford: Oxford University Press chapter 6, pp. 134–153.
- Stegmueller, Daniel. 2011. “Apples and Oranges? The Problem of Equivalence in Comparative Research.” *Political Analysis* 19(4):471–487.
- Thomas, Melissa A. 2010. “What Do the Worldwide Governance Indicators Measure?” *European Journal of Development Research* 22(1):31–54.
- Transparency International. 2021. “Global Corruption Barometer.”
- Treisman, Daniel. 2000. “The Causes of Corruption: A Cross-National Study.” *Journal of Public Economics* 76(3):399–457.
- Treisman, Daniel. 2007. “What Have We Learned About the Causes of Corruption from Ten Years of Cross-National Empirical Research?” *Annual Review of Political Science* 10(1):211–244.
- Tribunal Superior de Cuentas de Honduras. 2018. “Código de Conducta Ética del Tribunal Superior de Cuentas de Honduras.”
- Trochim, William. 2006. “The Research Methods Knowledge Base, 2nd Edition.” .
URL: <http://www.socialresearchmethods.net/kb/>
- Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, New Jersey: Princeton University Press.
- Weingast, Barry R. 1997. “The Political Foundations of Democracy and the Rule of Law.” *American Political Science Review* 91(2):245–263.
- World Bank. 2010. “Fraud and Corruption Awareness Handbook.”
- World Bank. 2016. “Guidelines on Preventing and Combating Fraud and Corruption in Projects Financed by IBRD Loans and IDA Credits and Grants.”
- World Bank. 2021. “Enterprise Surveys.” .
URL: <https://www.enterprisesurveys.org/en/enterprisesurveys>
- Zachary, Paul and William Spaniel. 2020. “Getting a Hand by Cutting Them Off: How Uncertainty over Political Corruption Affects Violence.” *British Journal of Political Science* 50(2):457–480.
- Zamboni, Yves and Stephan Litschig. 2018. “Audit Risk and Rent Extraction: Evidence from a Randomized Evaluation in Brazil.” *Journal of Development Economics* 134(April):133–149.

Appendix A Honduran Audits Codebook	App-3
A.1 Overview and Scope	App-3
A.2 Variables in the Honduran Audits Dataset	App-3
A.3 Finding and Cataloging the Municipal Audit Reports	App-4
A.4 Coding Process for the Municipal Audit Reports	App-6
 Appendix B Indian Audits Codebook	App-10
B.1 Overview and Scope	App-10
B.2 Variables in the Indian Audit Datasets	App-12
B.3 Coding Procedure	App-14
B.3.1 Revenue and All Sectors	App-14
B.3.2 Assessment, Levy and Collection of Major and Minor Mineral Receipts	App-16
B.3.3 Civil Sector, Commercial Sector, Civil and Commercial Sector . . .	App-17
B.3.4 Collection of Motor Vehicles Taxes, Fees, and Fines	App-18
B.3.5 Economic Sector	App-19
B.3.6 Mahatma Gandhi National Rural Employment Guarantee Scheme .	App-20
B.3.7 Non-Public Sector Undertakings	App-20
B.3.8 State Finance Sector	App-21
B.3.9 Taxes and Duties Sector	App-24
B.3.10 Transport and Infrastructure	App-25
B.3.11 Union Territory Finances	App-26
 Appendix C Guatemalan Audits Codebook	App-28
C.1 Overview and Scope	App-28
C.2 Variables in the Dataset	App-28
C.3 Coding Process	App-28

Appendix D Mexican Audits Codebook	App-30
D.1 Overview and Scope	App-30
D.2 Variables in the Dataset	App-31
 Appendix E Technical Appendix: Survival Analysis of the Honduran Audit Distribution	 App-34
E.1 Data for Survival Analysis	App-34
E.2 Survival/Event History Methods	App-34
 Appendix F Additional Regressions	 App-36
F.1 Honduras Ethnicity	App-36
F.2 Mexico Regressions	App-36

Appendix A Honduran Audits Codebook

A.1 Overview and Scope

The present codebook details the coded Honduran audits. The only audits in the dataset pertain to those of municipalities. Municipal-owned enterprises are not a part of this dataset.

A.2 Variables in the Honduran Audits Dataset

Table A1: Variables in the Dataset

Variable	Definition
<code>unique_number</code>	Unique identifier for the municipal audit. On occasions where an audit is divided into an “A” and “B”, we merge the “A” and “B” into one audit based on the unique list of infractions without double counting. The dataset also only includes municipal audits, not audits of state-owned enterprises taking place in specific municipalities.
<code>municipio</code>	Municipality where the audit took place.
<code>departamento</code>	Department (state/province equivalent) where the audit took place.
<code>audit_start_date</code>	Date when the audit started.
<code>audit_end_date</code>	Date when the audit ended.
<code>audit_duration</code>	Duration in days of the audit.
<code>total_infractions</code>	Total number of infractions in the audit.
<code>corrupt_infractions</code>	Total number of infractions in the audit.
<code>severity_infractions</code>	Severity of infractions in the audit. These qualitative assessments take the form of an ordinal score from 1 (not severe) to 7 (high severity).

A.3 Finding and Cataloging the Municipal Audit Reports

To access the PDF files with the municipal audit reports, visit <https://www.tsc.gob.hn/web/>, and click on “Informes de Auditorías”.



Next, click on “Informes de Auditorías”, and select the respective *departamento* (department).



The screenshot shows a web browser displaying the website of the Tribunal Superior de Cuentas of Honduras. The header features the national emblem of Honduras and the text "República de Honduras" and "Tribunal Superior de Cuentas". Below the header, there is a section titled "INFORMES DE AUDITORÍAS MUNICIPALES" which lists 11 departments: Atlántida, Colón, Comayagua, Copán, Cortés, Choluteca, El Paraíso, Francisco Morazán, Gracias a Dios, Intibucá, and Islas de La Bahía.

Now, go to to the respective audit and save it in the Google Drive. Note: the *Tribunal Superior de Cuentas* (Supreme audit agency) organizes the audits by their end date.

The screenshot shows a web browser displaying the website of the Tribunal Superior de Cuentas for the Atlántida department. The header features the national emblem of Honduras and the text "República de Honduras" and "Tribunal Superior de Cuentas". Below the header, there is a section titled "ATLÁNTIDA" which lists several audit reports:

- [INFORME N° 077-2014-DAM-CFTM-AM-A PROCESADORA MUNICIPAL DE CARNE \(PROMUC\) LA CEIBA](#)
- [INFORME N° 076-2014-DAM-CFTM-AM-B DIMATELA](#)
- [INFORME N° 076-2014-DAM-CFTM-AM-A DIMATELA](#)
- [INFORME N° 073-2014-DAM-CFTM-AM-A ESPARTA](#)
- [INFORME N° 066-2014-DAM-CFTM-AM-A JUTIAPA](#)
- [INFORME N° 062/2014-DAM-CFTM-AM-B LA CEIBA](#)
- [INFORME N° 052/2014-DAM-CFTM-AM-A ARIZONA](#)
- [INFORME N° 044/2014-DAM-CFTM-AM-B TELA](#)
- [INFORME N° 059/2012-DAM-CFTM-AM-A SAN FRANCISCO](#)

Finally, before entering the audit on the Google Sheet with the relevant details, make sure

that the audit is not already there. We do not want any duplicates. You can spot any duplicates, in particular, by the audit start and end dates.

unique_number	missing link?	link	municipio	departamento	audit_start_date	audit_end_date
19	missing from MTURK		la ceiba	atlantida	12/19/2007	5/31/2014
20		http://www.tsc.gob.hn/Auditorias/Informes_de_Auditoria/Sector_municipal/Atlantida/032-2008-DASM.pdf	la masica	atlantida	11/4/2005	7/8/2008
21		http://www.tsc.gob.hn/Auditorias/Informes_de_Auditoria/Sector_municipal/Atlantida/013-2012-DAM-CFTM-AM-A.pdf	la masica	atlantida	7/9/2008	1/31/2012

A.4 Coding Process for the Municipal Audit Reports

Generally, each audit report opens up with a declaration of what the auditor covers in his report. A table of contents follows and can sometimes be helpful in pinpointing where infractions are located. However, it's better to review each individual page to ensure each infraction is accounted for.

Para las categorías de control interno mencionadas anteriormente, se obtuvo una comprensión de su diseño y funcionamiento; y se observó hechos que se dan a conocer debido al efecto adverso que pueden tener para las operaciones de esa entidad y se detallan a continuación:

1. Ingresos recaudados y no depositados 24 horas después de su recaudación;
2. El personal que labora en la procesadora Municipal de Carnes (PROMUC) no está debidamente identificado;
3. La Administración de la Procesadora Municipal de Carnes (PROMUC) no cuenta con un Reglamento o Manual de Funciones;
4. Diferencias determinadas al analizar los ingresos mensuales registrados en detalle del Estado de Resultado con relación a los valores cuantificados según documentación soporte;
5. La administración de Procesadora Municipal de Carnes (PROMUC) no cuenta con formato de orden de pago;
6. Algunos Funcionarios y Empleados no cuentan con expediente de personal y otros están incompletos;

Skimming through the document, you will start to notice numbered lists. These are often the infractions that the auditor will provide further detail on in the coming pages. Said infractions are then supplemented with a recommendation from the auditor. The screenshot below captures what an infraction looks like within a report:

The screenshot shows a web page from tsc.gob.hn with a dark header bar. Below it, the main content area has a white background. At the top left, there's a bold section heading "B. DEFICIENCIAS DE CONTROL INTERNO". Underneath it, a numbered list item "1. INGRESOS RECAUDADOS Y NO DEPOSITADOS 24 HORAS DESPUÉS DE SU RECAUDACIÓN" is shown in bold. To its right is a detailed paragraph about uncollected funds. Below this is a table with three columns: "Fecha de Recaudación", "Fecha de Depósito", and "Valor del Depósito". The table contains five rows of data. Further down, there are two more paragraphs of explanatory text. At the bottom, a bold section "RECOMENDACIÓN N°1" is visible.

Copy and paste the Caps locked infraction alongside the brief paragraph description into the “Original Spanish Infractions” column on the Google Sheet.

The screenshot shows a Google Sheets document titled "honduran_audits". The top menu includes File, Edit, View, Insert, Format, Data, Tools, Add-ons, Help, and "All changes saved in Drive". The sheet itself has several rows of data. Row 1 contains the header "list infractions and page numbers (avoid duplicates)" and "Original Spanish infractions". Rows 20 and 21 contain detailed audit findings. Row 20 spans across both columns and includes a large red box highlighting the text. Row 21 also spans across both columns and includes a large red box highlighting the text. The bottom of the sheet shows a navigation bar with tabs for "original audits", "Pivot Table 2", "Pivot Table 3", and "short_infractions".

After pasting the infraction and its description into the Google Sheet, include the page number of the infraction at the end of the pasted statement. Please translate each infraction, and classify each infraction according to the typology of Table A2. Note that it is important to consider whether the audit report indicates corrupt intent. Often, intent is very difficult discern. When the intent is unclear, we generally do not classify the infraction as corrupt. The only exception is if the infraction is so egregious that it could not be construed as other than corrupt behavior. Regardless, we separately count the overall number of infractions as well as the number of corrupt infractions. Additionally, we make a ordinal, 1-7 assessment of the severity of the infractions in each audit.

L	M	N	O
total infractions	corrupt infractions	severity of infractions (1=not, 7=very)	list infractions and page numbers (avoid duplicates)
43	5	3	Monthly bank statements were not produced (p11). Municipal staff did not prepare budget reports (p11). Taxpayers arrears were not registered in the accounts (p11). Revenues and expenditures were not updated (p12). Municipal revenues were not deposited intact (p12). Many municipal transactions were paid in cash (p12). There was no small municipal account for minor expenses (p13). No regular cash balance reviews (p13). Revenues from land sales were placed in the checking account for covering functioning costs (p13). Insufficient documentation of subsidy receipts (p13). Interest earned on savings accounts not documented in revenue accounting (p14). Treasurer works on issues outside of her official charge (p15). Municipality did not submit tax receipt income documents to the Tribunal Superior (p16). Municipal Secretary cannot show attendance lists of aldermen allowance meetings (p16). Municipal bulletin on local government actions was not published (p17). Tax documents were not prenumbered as required (p17). Taxpayer IDs not updated because treasurer did not send income statements to Cadastre Dept. in a timely manner (p17). Taxpayer IDs were not reviewed or reconciled with taxable values (p18). Taxpayer IDs had been marked up with pencil (p18). Cadastre Dept. did not send information to tax collectors about taxpayer debt (p18). Municipality did not have formal register

Because of the assessments regarding the corrupt behavior are challenging, all reports were reviewed by a senior coder. In some more difficult cases, a third coder performed a final review.

Table A2: Typology of Corrupt Activities

Term	Definition
Bribery	“Offering, giving, receiving or soliciting, directly or indirectly, of anything of value to influence improperly the actions of another party” (World Bank, 2016 , 3). Also, we can think of corruption simply as the abuse of public office for private gain. Don’t forget to code the use of agents/intermediaries or shell/suspicious companies here.
Fraud	“Any act or omission, including a misrepresentation, that knowingly or recklessly attempts to mislead, a party to obtain a financial or other benefit or to avoid an obligation” (World Bank, 2016 , 3). This includes fraudulent contract management: that is, “fraudulent implementation including misrepresentation of goods, works, and services as having been delivered according to specifications” (World Bank, 2010). You see these types of things in fraudulent invoices (e.g., overcharging), not completing the work that you say you did, government supervising officials signing-off on poor quality work.
Obstruction	(i) deliberately destroying, falsifying, altering or concealing of evidence material to the investigation or making false statements to investigators in order to materially impede an audit or investigation into allegations of a corrupt, fraudulent, coercive or collusive practice; and/or threatening, harassing or intimidating any party to prevent it from disclosing its knowledge of matters relevant to the investigation or from pursuing the investigation, or (ii) acts intended to materially impede the exercise of the state or an auditor’s contractual rights of audit or access to information” (World Bank, 2016 , 3).
Coercion	“impairing or harming, or threatening to impair or harm, directly or indirectly, any party or the property of the party to influence improperly the actions of a party” (World Bank, 2016 , 3).
Collusion	“an arrangement between two or more parties designed to achieve an improper purpose, including to influence improperly the actions of another party” (World Bank, 2016 , 3). Here we must count specification rigging: tailoring a procurement tender so that only 1 firm is qualified to win. Another important thing to count here is bid-rigging: this is usually when firms work together in an oligopolistic manner to keep bid prices low. Alternatively, what they do is one firm agrees to not bid on the contract, but one firm agrees to subcontract out arrangements downstream in the contract cycle. This is what we call splitting. This includes insider information sharing.
Theft	embezzlement of funds not covered under “fraud”.
Nepotism	Hiring of family members.
Cronyism	Hiring of friends.
Grand corruption	“Collusion among the highest levels of government that involves major public sector projects, procurement, and large financial benefits among high-level public and private elites” (Bauhr and Charron, 2018).

Appendix B Indian Audits Codebook

B.1 Overview and Scope

This section provides an overview of the dataset. To obtain objective data on corruption, we coded audit reports undertaken by the Comptroller and Auditor General (CAG) of India. Although the CAG undertakes yearly audits of states, local bodies, and union territories, the present dataset only covers state-level audits. By state, the CAG also conducts various types of audits each year: compliance, financial, and performance. We attempted to code all of these types of audits, but compliance and performance audits do not have similar formats to facilitate consistent coding, so the present dataset focuses on financial audits.

The first year for which audit reports are publicly available on the [CAG website](#) is 2001. The present dataset thus starts in 2001. To date, we have coded all available audit reports through 2019.

The CAG conducts its annual state-level audits by sector. Accordingly, the primary unit of analysis for this dataset is the state-sector-year. By each state-sector-year, the **main dataset** of the dataset provides users with data on the:

1. Number of outstanding Inspection Results (IRs). This is an aggregate-level measure of the number of cases/investigations related to a set of irregularities that are not immediately resolved.
2. Number of outstanding audit observations/paragraphs. This is a disaggregated measure of the total number of irregularities in the outstanding IRs. “Observations” and “paragraphs” are used interchangeably throughout the reports.
3. Amount of revenue involved in crore. This is a measure of the monetary value of the irregularities. A crore corresponds to 10,000,000 rupees.

The dataset also provides for a more disaggregated view of the audit data at the subsector,

department, and receipts levels. Similar to the data on the main dataset, the department-level data consists of data on the number of outstanding Inspection Results (IRs), number of outstanding audit observations, and amount of revenue involved in crore. The receipts data correspond to a more disaggregated view of the department-level data.

The subsector level provides the most disaggregated view of the data, including the ability to track the evolution of audit infractions over time. As Appendix B.2 describes, there are data on the:

1. Number of Inspection Results (IRs) at *opening*.
2. Number of audit observations/paragraphs at *opening*.
3. Amount of revenue involved in crore at *opening*.
4. Number of **additional** Inspection Results (IRs).
5. Number of **additional** audit observations/paragraphs.
6. Amount of **additional** revenue involved in crore.
7. Number of Inspection Results (IRs) at *clearance*.
8. Number of audit observations/paragraphs at *clearance*.
9. Amount of revenue involved in crore at *clearance*.
10. Number of Inspection Results (IRs) at *closing*.
11. Number of audit observations/paragraphs at *closing*.
12. Amount of revenue involved in crore at *closing*.

Some of the subsector-level audits do not provide the same level of detail. Instead, they simply provide data on the:

13. Number of **pending** Inspection Results (IRs).
14. Number of **pending** audit observations/paragraphs.

B.2 Variables in the Indian Audit Datasets

Table B1: Variables in the Main Dataset

Variable	Definition
state	State where the audit took place.
year	Year that the audit finalized.
IR_pending	Total number of Inspection Reports (IRs)/investigations pending for settlement.
outstanding_observations	Total number of infractions in the audit.
revenue_involved	Amount of revenue involved in the irregularities in crore. Note: one crore is 10,000,000 rupees.

Table B2: Variables in the Department Dataset

Variable	Definition
state	State where the audit took place.
department	Department within the state where the audit took place.
year	Year that the audit finalized.
outstanding_IRs	Total number of Inspection Reports (IRs)/investigations pending for settlement.
outstanding_observations	Total number of infractions in the audit.
revenue_involved	Amount of revenue involved in the irregularities in crore. Note: one crore is 10,000,000 rupees.

Table B3: Variables in the Department Dataset

Variable	Definition
state	State where the audit took place.
department	Department within the state where the audit took place.
nature_receipt	Area/subject of the receipt.
year	Year that the audit finalized.
outstanding_IRs	Total number of Inspection Reports (IRs)/investigations pending for settlement.
outstanding_observations	Total number of infractions in the audit.
revenue_involved	Amount of revenue involved in the irregularities in crore. Note: one crore is 10,000,000 rupees.

Table B4: Variables in the Subsector Dataset

Variable	Definition
<code>state</code>	State where the audit took place.
<code>sector</code>	Sector where the audit took place.
<code>subsector</code>	Economic sector where the audit took place.
<code>year</code>	Year that the audit finalized.
<code>IR_opening</code>	Total number of Inspection Reports (IRs)/investigations pending for settlement at audit opening.
<code>paragraphs_opening</code>	Total number of paragraphs/irregularities pending for settlement at audit opening.
<code>moneyvalue_opening</code>	Money value of the paragraphs/irregularities in crore at audit opening. Note: one crore is equivalent to 10,000,000 rupees.
<code>IR_addition</code>	Total number of additional Inspection Reports (IRs)/investigations pending for settlement during this audit.
<code>paragraphs_addition</code>	Total number of additional paragraphs/irregularities pending for settlement during this audit.
<code>moneyvalue_addition</code>	Money value of the additional paragraphs/irregularities in crore at audit opening. Note: one crore is equivalent to 10,000,000 rupees.
<code>IR_clearance</code>	Total number of cleared Inspection Reports (IRs)/investigations that were pending settlement at audit opening.
<code>paragraphs_clearance</code>	Total number of cleared paragraphs/irregularities were pending settlement at audit opening.
<code>moneyvalue_clearance</code>	Money value of the cleared paragraphs/irregularities in crore that were pending settlement at audit opening. Note: one crore is equivalent to 10,000,000 rupees.
<code>IR_closing</code>	Total number of Inspection Reports (IRs)/investigations pending for settlement at audit closing.
<code>paragraphs_closing</code>	Total number of paragraphs/irregularities pending for settlement at audit closing.
<code>moneyvalue_closing</code>	Money value of the paragraphs/irregularities in crore at audit closing. Note: one crore is equivalent to 10,000,000 rupees.
<code>IR_pending</code>	Total number of Inspection Reports (IRs)/investigations pending for settlement at audit closing.
<code>paragraphs_pending</code>	Total number of paragraphs/irregularities pending for settlement at audit closing.

B.3 Coding Procedure

The procedure for coding the audits differs slightly by sector. Most audits, however, follow the Revenue sector.

B.3.1 Revenue and All Sectors

main dataset. Code the three values of the column of the most recent date. In the table below, code the values in the June 2010 column:

	June 2008	June 2009	June 2010
Number of outstanding IRs	1362	1364	1518
Number of outstanding audit observations	3710	3685	4033
Amount involved (₹ in crore)	683.71	711.53	767.23

Department Dataset. Code the name of the state, report number, page number, department names found in column 2, year of the report audited, and the three numerical values found in columns 4-6. For department rows that have multiple numerical values, add the values together. For example, the Finance (Taxation) department has three values for the Number of Outstanding IRs, so add the values of 152, 8, and 68 to get the total.

SL No.	Name of the Department	Nature of receipts	Number of outstanding IRs	Number of outstanding audit observations	Money value involved (₹ in crore)
1.	Finance (Taxation)	(a) Taxes/VAT on sales, trade, etc.	152	697	299.94
		(b) Agriculture income tax	8	30	2.02
		(c) Entry tax, Electricity duty, Entertainments tax, and luxury tax, etc.	68	76	2.62
2.	Excise	State excise	61	167	11.42
3.	Revenue	Land revenue	605	1442	208.97
4.	Transport	Taxes on motor vehicles	192	493	20.01
5.	Stamps and registration	Stamps and registration fees	154	308	16.13
6.	Mines and geology	Non-ferrous mining and metallurgical industries	9	37	79.71
7.	Forest and environment	Forestry and wild life	269	783	126.41
Total			1,518	4,033	767.23

Subsector dataset. Code the state, the report number, page number table, year the report was published, the general sector (in this case the Revenue sector), the economic sector (indicated in the below sentence - Finance (Taxation) Department), year audited (column 1), and all of the numerical values. For the year audited (column 1), we only input the second year out of the range.

Year	Opening balance			Addition during the year			Clearance during the year			Closing balance during the year			(₹ in crore)
	IRs	Para-graphs	Money value	IRs	Para-graphs	Money value	IRs	Para-graphs	Money value	IRs	Para-graphs	Money value	
2000-01	214	800	6,448.21	55	249	2,625.61	19	108	1,169.33	250	941	7,904.49	
2001-02	250	941	7,904.49	39	115	1,704.89	18	58	477.85	271	998	9,131.53	
2002-03	271	998	9,131.53	57	162	1,958.23	11	102	437.99	317	1,058	10,651.77	
2003-04	317	1,058	10,651.77	70	237	5,607.70	93	350	1,503.31	294	945	14,756.16	
2004-05	294	945	14,756.16	37	216	4,099.05	49	140	1,442.43	282	1,021	17,412.78	
2005-06	282	1,021	17,412.78	67	224	9,252.86	67	134	355.22	282	1,111	26,310.42	
2006-07	282	1,111	26,310.42	42	189	6,251.86	41	317	4,740.79	283	983	27,821.49	
2007-08	283	983	27,821.49	65	300	11,135.80	47	100	8,659.96	301	1,183	30,297.33	
2008-09	301	1,183	30,297.33	45	335	3,336.28	103	469	4,246.29	243	1,049	29,387.32	
2009-10	243	1,049	29,387.32	46	254	8,799.31	13	75	2,113.75	276	1,228	36,072.88	

Notes. The table below would be coded in the subsector dataset. The relevant information found in the main dataset would fill columns A-G. For column H, you would put the “year of issue of Inspection Reports”. The 3rd column in this table would go in IR_opening (Column I) and paragraphs_opening (Column J). The 4th column in this table would be coded as IR_addition and paragraphs_addition. The “Total” column in the table below would not be coded. The 6th column in the table below would be coded as IR_clearance and paragraphs_clearance. The last column in this table would be coded as IR_pending and paragraphs_pending. There may be multiple of these tables in the report, so code all of them.

Table-13
Position of pending IRs/ Paras

Sl. No.	Year of issue of Inspection Reports	Outstanding IRs/ Paras as on 31 March 2014		Addition		Total		No. of IRs/ paras settled during 2014-15		No. of outstanding IRs/Paras as on 31 03.2015.	
		IRs	Paras	IRs	Paras	IRs	Paras	IRs	Paras	IRs	Paras
1.	Upto 2010-11	99	673	-	-	99	673	-	48	99	625
2.	2011-12	14	138	-	-	14	138	-	20	14	118
3.	2012-13	15	175	-	-	15	175	1	36	14	139
4.	2013-14	17	218	-	-	17	218	-	-	17	218
5.	2014-15	-	-	14	144	14	144	-	-	14	144
Total		145	1,204	14	144	159	1,348	1	104	158	1,244

B.3.2 Assessment, Levy and Collection of Major and Minor Mineral Receipts

The table below would be coded in the subsector dataset. The only column that would be coded is “Amount Accepted” because we are only interested in the number of cases that were recognized as legitimate by the government. The cases would be coded as paragraphs_opening (Column I) and the amount would be coded as moneyvalue_opening (Column J). If necessary, convert the money amount to crore.

Year of Inspection Report	No. of units audited	(₹ in crore)					
		Amount objected		Amount accepted		Amount recovered	
		Cases	Amount	Cases	Amount	Cases	Amount
2006-07	11	21	335.00	16	221.00	4	0.42
2007-08	13	640	68.09	470	56.62	5	0.29
2008-09	12	764	20.09	473	1.45	1	0.14
2009-10	7	396	4.64	335	2.33	45	4.83
2010-11	9	302	23.71	149	6.14	61	0.06
Total	52	2123	451.53	1443	287.54	116	5.74

B.3.3 Civil Sector, Commercial Sector, Civil and Commercial Sector

main dataset. This sector does not have a table suited for the main dataset.

Departments and Receipts Tabs. The table below would be coded under the Department Dataset.

Sl.No.	Department	Inspection Reports	Audit Paragraphs
1	Health and Family Welfare	116	690
2	Law	16	44
3	Industries	2	4
4	Fisheries & Animal Husbandry	13	44
5	Home, Prohibition and Excise	10	100
6	Commercial Taxes	1	2
7	Town and Country Planning	1	9
8	Higher Education	28	225
9	Tourism & Culture	11	36
10	Archaeology	4	15
11	Transport	6	9
12	Youth and Sports Development	1	3
13	Planning & Development	4	24
14	Tamil Development and Religious Endowment	12	26
15	Handlooms, Handicrafts, Textiles and Khadi	3	11
16	School Education	21	108
17	Revenue	22	165
18	Labour & Employment	7	22
19	Agriculture	104	373
20	Co-operation & Consumer Protection	32	115
21	Social Welfare	28	135
22	Backward Classes, Most Backward Classes & Minority Welfare	3	20
23	Adi Dravidar & Tribal Welfare	9	102
24	Public Works Department	5	10
Total		459	2,292

Department	Position of IRs issued upto September 2005 but not settled at the end of March 2006			Position of IRs and paragraphs not settled for more than 10 years			(Rupees in crore) Position of IRs in respect of which first reply not received from March 1992 to March 2006		
	No. of IRs	No. of paragraphs	Money value	No. of IRs	No. of paragraphs	Money value	No. of IRs	No. of paragraphs	Money value
Finance									
a) Sales Tax	79	258	10.68	37	140	1.24	16	81	3.41
b) Professions Tax	3	4	0.06				1	3	0.06
c) Stamp Duty and Registration Fees	4	4	0.13				1	1	0.01
d) Electricity Duty	182	728	28.73	34	139	1.18	43	274	4.98
e) Agricultural Income Tax	1	2					NIL	NIL	NIL
f) Amusements Tax	4	10	0.13	1	1	0.02	5	11	0.16
Forest									
Forest Receipts	95	260	16.01	32	169	2.11	27	140	3.60
Land and Land Reforms									
Land Revenue	2	2	0.04				1	1	0.02
Excise									
State Excise	9	12	0.70	1	1	1.38	3	5	0.09
Transport									
Motor Vehicles	14	63	12.99				9	46	10.69
Total	393	1343	69.47	105	450	5.93	106	562	23.02

The above table would be coded in both the departments and the receipts tabs. The only columns that would be coded are those under “Positions of IRs not settled at the end of March 2006” because we are only interested in outstanding IRss. For the department tabs, the name of the department would be the bolded term and the subterms would be added up. For the receipts dataset, each individual row (sales tax, professions tax, etc.) would be

a receipt.

Subsector dataset. Below are examples of tables that would be coded under the subsector dataset.

Year	Inspection Reports	Audit Paragraphs
Up to 2004-05	119	186
2005-06	394	782
2006-07	691	1,490
2007-08	1,175	3,679
2008-09	1,104	5,258
Total	3,483	11,395

Year	Opening balance		Addition during the year		IRs/paragraphs settled during the year		IRs/paragraphs outstanding at the close of the year	
	IRs	Paragraphs	IRs	Paragraphs	IRs	Paragraphs	IRs	Paragraphs
2004-05	66	124	37	75	11	22	92	177
2005-06	92	177	50	105	04	10	138	272
2006-07	138	272	12	29	-	-	150	301
2007-08	150	301	10	41	19	48	141	294
2008-09	141	294	07	23	10	17	138	300

B.3.4 Collection of Motor Vehicles Taxes, Fees, and Fines

main dataset. The table below would be coded under the main dataset.

	June 2008	June 2009	June 2010
Number of outstanding IRs	121	128	100
Number of outstanding audit observations	347	349	267
Amount involved (Rupees in crore)	12.52	22.24	9.44

Subsector dataset. In the table below, code the values in the IR_opening, paragraphs_opening,

and moneyvalue_opening columns respectively. If necessary, convert lakhs or rupees to crore.

Year	Number of outstanding IRs	Number of outstanding paragraphs	Money value involved (Rupees in lakh)
1998-99	2	2	1.92
1999-2000	5	8	13.88
2000-01	9	16	130.81
2001-02	11	24	165.07
2002-03	13	37	130.54
2003-04	14	49	121.70
2004-05	16	41	51.52
2005-06	10	47	196.71
2006-07	8	19	20.42
2007-08	5	9	10.91
2008-09	5	9	98.04
2009-10	2	6	2.58
Total	100	267	944.10

B.3.5 Economic Sector

main dataset. This sector does not have a table suited for the main dataset.

Department Dataset. Below is a unique table that could arguably be coded in either the department or subsector dataset. As a team, we collectively decided to code it in the department dataset. More specifically, we decided to code 15 rows for the 15 departments with the year 2004 as the year audited. Then, we repeated those 15 rows five more times to input the data from years 2005, 2006, 2007, 2008, and 2009.

SL No.	Department	Year upto 2003-04			2004-05			2005-06			2006-07			2007-08			2008-09		
		IRs	Paras	Amt.	IRs	Paras	Amt.	IRs	Paras	Amt.	IRs	Paras	Amt.	IRs	Paras	Amt.	IRs	Paras	Amt.
1	Public Works	467	1514	801.94	63	264	148.86	79	417	177.42	77	335	662.66	82	423	1471.14	83	378	1811.68
2	Water Resources	557	1322	503.7	89	298	284.8	118	414	1085.46	114	452	516.72	129	478	485.26	119	425	1705.7
3	Narmada Valley Development Authority	92	192	143.7	17	42	129.09	21	50	127.4	18	42	33.14	29	69	135.4	29	62	233.88
4	Panchayat & Rural Development (MPRRDA)	18	61	34.82	14	105	84.07	19	51	107.05	31	108	86.36	45	163	252.07	46	177	493.09
5	Forest	0	0	0	0	0	0	0	0	59	56	23.24	15	20	3.58	12	14	15.92	
6	Farmer Welfare & Agriculture Development	98	168	43.8	37	42	6.74	25	60	33.07	38	76	45.7	71	153	42.3	39	65	71.11
7	Co-operative	136	379	61.34	5	11	1.54	9	21	10.14	1	4	3.15	6	11	1.01	2	5	6.44
8	Animal Husbandry	145	276	20.89	6	12	0.37	20	66	22.24	24	62	0.22	21	66	19.02	39	119	18.04
9	Fisheries	10	11	0.39	1	2	0.08	5	12	2.46	0	0	0	2	3	0.11	2	4	0.44
10	Rural Industries	49	107	69.68	12	27	63.82	5	19	23.92	10	52	56.07	15	46	18.07	9	32	47.02
11	Commerce, Industries & Employment	42	85	114.79	1	6	14.21	11	24	53.23	63	134	249.55	3	12	21.34	8	17	35.27
12	Energy	10	13	211.86	1	1	0.04	0	0	0	0	0	0	0	0	0	0	0	0
13	Tourism	2	2	0.59	1	5	7.08	1	5	6.92	1	2	2.2	0	0	0	1	7	9.23
14	Civil Aviation	4	27	27.24	1	4	3.57	0	0	0	0	0	0	0	0	0	1	4	3.37
15	Horticulture and Food Processing	28	62	52.26	11	27	27.23	25	134	79.97	3	8	12.67	9	21	21.42	3	8	13.73
	Total	1658	4219	2087	259	846	771.5	338	1273	1729.28	439	1331	1691.68	427	1465	2470.72	393	1317	4464.92

Receipts and Subsector datasets. There are no tables suited for the receipts and sub-

sector coding.

B.3.6 Mahatma Gandhi National Rural Employment Guarantee Scheme

These two tables would not be coded because we are only interested in departments, not districts.

Name of the district	2008-09		Open with 2009-10		2010-11		2011-12		Total	
	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount
Bidar	17	0	-	-	5	0	-	-	22	0
Chikkaballapur	-	-	1	9.10	-	-	-	-	1	9.10
Dharwad	-	-	1	3.58	-	-	-	-	1	3.58
Koppal	-	-	-	-	-	-	14	0	14	0
Raichur	1	1.39	-	-	-	-	-	-	1	1.39
Tumkur	-	-	3	40.18	-	-	-	-	3	40.18
Total	18	1.39	5	52.86	5	0	14	0	42	54.25

(b) Details of civil cases

Name of the district	2008-09		2009-10		2010-11		2011-12		Total	
	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount	No. of cases	Amount
Bangalore (Rural)	-	-	-	-	8	12.01	3	1.22	11	12.23
Bangalore (Urban)	-	-	6	12.86	-	-	-	-	6	12.86
Belgaum	-	-	-	-	-	-	1	0	1	0
Bidar	12	0	6	0	28	0	-	-	46	0
Chickmagalur	-	-	-	-	4	0.80	-	-	4	0.80
Chitradurga	1	4.21	-	-	-	-	-	-	1	4.21
Dharwad	-	-	-	-	1	12.91	58	1.80	59	14.71
Gadag	-	-	-	-	5	1.14	3	8.37	8	9.51
Hassan	-	-	-	-	7	20.80	1	0	8	20.80
Haveri	-	-	1	0.01	5	9.55	1	0	7	9.56
Uttara Kannada	-	-	-	-	1	0	-	-	1	0
Kodagu	-	-	2	0.16	5	1.11	-	-	7	1.27
Koppal	-	-	5	0	17	0	14	0	36	0
Mandy	-	-	2	15.47	2	3.96	1	0	5	19.43
Mysore	-	-	1	4.50	3	74.10	0	-	4	79.60
Shimoga	-	-	3	4.93	1	0.39	0	-	4	5.32
Yadgir	-	-	-	-	2	23.59	3	4.67	5	28.26
Total	13	4.21	26	37.93	89	160.36	85	16.06	213	218.56

Source: Information furnished by the RDPR department

B.3.7 Non-Public Sector Undertakings

The table below would be coded in subsector. The year audited would be the year in the very top row.

Table 1.4: Outstanding Inspection Reports/paragraphs

Particulars	Prior to April 2009	2009-10	2010-11	2011-12	2012-13	2013-14	Total
Inspection Reports	13221	1175	1232	1342	1583	1676	20229
Paragraphs	17953	2489	2800	4051	5674	8008	40975

B.3.8 State Finance Sector

Department Dataset. The tables below would be coded in the departments dataset. As before, make sure to convert all currency values into crore.

Table 3.4: Department-wise break-up of misappropriations, losses, defalcations, etc.

(Rupees in lakh)

Sl. No.	Name of the Department	Number of cases	Amount
1.	Forest	13	196.93
2.	Revenue	1	0.85
3.	Animal Resources Development	1	0.73
4.	Health and Family Welfare	3	198.38
5.	Agriculture	2	1.45
Total :		20	398.34

(Reference : Paragraph 3.5; Page 52)

Department/category wise details in respect of cases of loss to Government due to theft, misappropriation/loss of Government material

Name of Department	Theft Cases		Misappropriation/ Loss of Government Material		Total	
	Number of Cases	Amount (₹ in lakh)	Number of Cases	Amount (₹ in lakh)	Number of Cases	Amount (₹ in lakh)
PWD	-	-	5	1749.75	5	1749.75
Medical	-	-	3	17.90	3	17.90
Employment & Craftsman	-	-	1	1.92	1	1.92
Education	-	-	6	2960.42	6	2960.42
Treasuries & Accounts	-	-	1	460.89	1	460.89
Rural Development	-	-	1	126.82	1	126.82
Irrigation & Flood	-	-	1	386.00	1	386.00
Development Authority of Nagaland	-	-	1	100.00	1	100.00
Urban Development	-	-	1	545.00	1	545.00
DUDA	-	-	1	47.50	1	47.50
Food & Civil Supplies	-	-	1	266.25	1	266.25
Registrar of Co-operative Societies	-	-	1	10.81	1	10.81
Transport Commissioner	-	-	2	146.57	2	146.57
Vigilance Commission	1	1.20	-	-	1	1.20
Total	1	1.20	25	6819.83	26	6821.03

Similarly, the table below would be coded in the department dataset. They are relevant because they show paragraphs/infringements as well as money amounts. When coding the table below, you would disregard rows 1-20: they are too low-level. Only focus on the last row, total.

Appendix-3.7
(Reference: Paragraph 3.5; Page 49)

Statement showing irregularities commented upon in the outstanding Inspection Reports and paragraphs as on 31 March 2011

Sl. No	Type of irregularities	Education				Cooperation		Industry		Police		Total	
		Secondary Education		HPU		Para	Amount	Para	Amount	Para	Amount	Para	Amount
		Para	Amount	Para	Amount								
1	Drawal of funds in advance of requirement	29	243.99	7	54.06	0	0	4	81.26	15	1642.78	55	2022.09
2	Non-adjustment of contingent advances	24	8.59	24	973.04	1	0.12	0	0	0	0	49	981.75
3	Excess/irregular expenditure for want of sanctions	1169	10041.12	81	1414.35	12	28.07	7	1394.71	11	45.99	1280	12924.24
4	Wasteful/ infructuous/ unfruitful expenditure	85	1169.11	13	1262.09	2	3.70	8	2318.70	3	188.55	111	4942.15
5	Diversion of funds	3	2.03	2	46.74	0	0	0	0	1	73.63	6	122.40
6	Overpayment, non-recovery of rent, advances/ miscellaneous recoveries	1915	1699.08	43	199.92	19	226.89	20	366.96	30	3256.73	2027	5749.58
7	Non-production of actual payees' receipts	61	220.50	6	17.46	6	8.25	0	0	2	13.12	75	259.33
8	Outstanding loans	1	0.09	3	171.13	5	345.18	19	1011.73	0	0	28	1528.13
9	Idle machinery/ equipment including vehicles	30	34.06	10	248.45	1	0.37	1	8.54	5	144.68	47	436.10
10	Non-accounting/ shortage of stores/cash, etc.	235	342.88	8	11.82	3	0.30	4	1.07	3	1109.89	253	1465.96
11	Non-recoupment of expenditure	3	0.08	8	200.05	0	0	1	107.94	1	24.04	13	332.11
12	Misappropriation of stores/ cash/ funds	83	86.73	6	50.36	2	0	3	0.32	1	1.93	95	139.34
13	Incomplete/ abandoned works	25	1331.50	7	292.71	5	97.56	12	548.36	13	5622.99	62	7893.12
14	Loss/theft embezzlement/ defalcation, etc.	228	102.84	16	365.55	4	49.58	4	309.94	7	7.94	259	835.85
15	Non-production of UCs	12	73.78	6	37.03	1	0	8	1719.86	0	0	27	1830.67
16	Non-disposal of unserviceable articles of stores	163	586.98	8	2.74	3	0.55	1	5.15	2	19.99	177	615.41
17	Non-reconciliation with treasuries/ banks	154	339.86	2	0.01	1	0.01	0	0	0	0	157	339.88
18	Non-utilization of Grants-in-aid	72	994.31	5	58.29	2	22.85	8	844.18	1	1.50	88	1921.13
19	Non-deposit of interest in treasuries	39	11.36	10	28.19	1	1.14	8	947.72	2	51.29	60	1039.70
20	Miscellaneous irregularities	804	1444.11	58	2632.18	14	299.71	8	233.99	12	304.22	896	4914.21
Total		5135	18733.00	323	8066.17	82	1084.28	116	9900.43	109	12509.27	5765	50293.15

Sl. No.	Name of Departments	Defalcation, misappropriation loss, theft, fraudulent and excess payment	
		Number	Amount
1	Education	04	1.50
2	Energy	02	0.05
3	Panchayati Raj	01	0.21
4	Registration, Excise and Prohibition	02	0.56
5	Rural Development	03	0.12
6	Rural Works	02	0.54
7	Urban Development and Housing	03	0.60
Total		17	3.58

(Reference: Paragraph 3.5; Page 48)

Department/category-wise details of loss to Government due to theft, misappropriation, shortage etc.

Sl.No.	Name of the department	Theft		Misappropriation/ shortage of Government material		Total	
		Number of cases	Amount (₹ in lakh)	Number of cases	Amount (₹ in lakh)	Number of cases	Amount (₹ in lakh)
1.	Agriculture	8	1.53	139	477.36	147	478.89
2.	Animal Husbandry	9	0.25	5	88.77	14	89.02
3.	Higher Education	9	1.47	16	37.69	25	39.16
4.	Commercial Tax	3	127.68	3	127.68
5.	Co-operation	1	0.02	1	0.14	2	0.16
6.	Health and Family Welfare	4	4.27	36	237.12	40	241.39
7.	Elementary Education	1	..	6	540.17	7	540.17
8.	Election	1	0.23	1	0.23
9.	Labour and Employment	6	4.04	9	58.47	15	62.51
10.	Energy	1	0.07	1	68.14	2	68.21
11.	Finance	1	0.90	5	66.86	6	67.76
12.	Home	1	0.97	9	33.01	10	33.98
13.	Horticulture	6	14.68	6	14.68
14.	Inspector of factories	1	1.43	1	1.43
15.	Museum	1	14.57	1	14.57
16.	Public	2	1.04	2	1.04
17.	Rural Development and Panchayati Raj	1	3.60	7	11.69	8	15.29
18.	Revenue	144	108.34	144	108.34
19.	School Education	1	0.02	22	74.89	23	74.91
20.	Sericulture	1	0.05	3	1.75	4	1.80
21.	Social Welfare	2	2.93	9	3.08	11	6.01
22.	Planning and development	1	0.09	1	0.09
23.	Transport	1	1.97	1	12.80	2	14.77
24.	Environment and Forests	1	0.01	3	2.31	4	2.32
25.	Highways	2	16.79	2	16.79
26.	Industries	1	1	..
27.	Public Works	10	4.26	10	4.26
28.	Personnel and Administrative Reforms	1	1.91	1	1.91
Total		50	22.19	443	2005.18	493	2027.37

Subsector. The table below would be coded under the subsector dataset. The IRs and Paragraphs would go in “IR_pending” and “paragraphs_pending”, respectively.

Appendix-3.6
(Reference: Paragraph 3.5; Page 49)

Year-wise break up of outstanding Inspection Reports/Paras upto March 2011 of selected DDOs

Period	Education				Cooperation		Industry		Police		Total	
	Secondary Education		Himachal Pradesh University									
	IRs	Paras	IRs	Paras	IRs	Paras	IRs	Paras	IRs	Paras	IRs	Paras
Upto March 2001	776	1634	6	103	14	17	16	20	5	6	817	1780
2001-02	88	234	1	4	2	3	1	2	1	2	93	245
2002-03	104	269	1	11	3	5	1	1	4	6	113	292
2003-04	94	266	1	11	3	10	4	7	5	6	107	300
2004-05	83	260	2	42	3	3	3	8	4	15	95	328
2005-06	201	648	1	31	1	1	4	5	3	8	210	693
2006-07	249	814	1	21	4	6	5	12	9	28	268	881
2007-08	210	606	1	14	6	16	4	10	6	13	227	659
2008-09	96	278	1	64	3	11	5	13	4	6	109	372
2009-10	22	58	2	12	1	5	1	3	1	3	27	81
2010-11	25	68	1	10	1	5	6	35	6	16	39	134
Total	1948	5135	18	323	41	82	50	116	48	109	2105	5765

B.3.9 Taxes and Duties Sector

main dataset. The table below would be coded under the main dataset with the totals added up. For IR_pending the number would be 2732 (919 + 891 + 922), for outstanding_observations the number would be 5865 (1936 + 1944 + 1985), and for revenue_involved the number would be 540.78 (178.58 + 173.54 + 188.66).

Details of IRs	June 2013	June 2014	June 2015
Number of IRs pending for settlement	919	891	922
Number of outstanding audit observations	1,936	1,944	1,985
Amount of revenue involved (₹ in crore)	178.58	173.54	188.66

Department Dataset. For departments, code the name of the state, report number, page number, department names found in column 2, year of the report audited, and the three numerical values found in columns 4-6. For department rows that have multiple numerical values, add the values together.

Sl. No	Name of the Department	Nature of receipts	Numbers of outstanding IRs	Numbers of outstanding audit observations	Money value involved
1.	Finance	Taxes on Sales, Trade etc. and luxury tax etc. Entertainment	445 10	1,131 15	111.75 0.15
2.	Excise	State Excise	67	114	29.58
3.	Transport	Taxes on motor vehicles	103	266	35.63
4.	Stamp and Registration	Stamp and registration fees	297	459	11.55
		Total	922	1,985	188.66

Receipts Dataset. Using the same table above, code the name of the state, report number, page number, department names found in column 2, nature of receipts found in column 3, year of the report audited, and the three numerical values found in columns 4-6. The receipts dataset is meant to indicate the breakdown of the departments' values that we added together in the department dataset. If the department and receipts dataset's values are the same, code it as is.

Subsector dataset. Code the state, the report number, page number table, year the report was published, the general sector, the economic sector, year audited , and all of the numerical values. For the year audited (column 1), we only input the second year out of the range.

Year	Opening Balance			Addition during the year			Clearance during the year			Closing balance during the year		
	IRs	Para graphs	Money value	IRs	Para graphs	Money value	IRs	Para graphs	Money value	IRs	Para graphs	Money value
2010-11	312	751	50.01	54	158	89.37	6	32	0.70	360	877	138.68
2011-12	360	877	138.68	54	192	8.56	12	46	65.94	402	1,023	81.30
2012-13	402	1,023	81.30	37	151	23.79	27	78	11.20	412	1,096	93.89
2013-14	412	1,096	93.89	35	140	21.26	36	156	10.38	411	1,080	104.77
2014-15	411	1,080	104.77	44	135	8.83	17	100	13.12	438	1,115	100.48

B.3.10 Transport and Infrastructure

Code the table below under the department dataset.

5.2 Results of audit

The Revenue and Land Reform Department collected ₹ 83.54 crore during 2014-15. During the period 2014-15 we test checked the records of 20 units out of 307 units of Land Revenue with revenue collection of ₹ 5.69 lakh, revealed non/short levy of cesses and/or interest on arrears of cess, non/short fixation of *salami* and commercial rent, non-settlement of vested lands etc. involving ₹ 3.89 crore in 178 cases. This indicates the near abdication of duty of collection of Land Revenue by 20 units as detailed in **Table – 5.2**.

Table – 5.2

Sl. No.	Categories	Number of cases	(₹ in crore)	
			Amount	
1	Non-settlement of vested lands	16	0.10	
2	Non-settlement of <i>sairats</i>	9	0.02	
3	Other cases	153	3.77	
	Total	178	3.89	

B.3.11 Union Territory Finances

Department Dataset. The table below would be coded under the department dataset. We would only code the total number of cases and the total amount for each department. The number of cases would go in outstanding_observations and the amount would go in money_involved. Make sure to convert the lakhs to crore, if necessary.

Open with ▾ Department/category-wise pending cases of misappropriation, loss, theft, defalcation etc., of Government material						
SL. No.	Name of the Department	Theft and Loss cases		Misappropriation/ Loss of Government Material		Total
		Number of cases	Amount (₹ in lakh)	Number of cases	Amount (₹ in lakh)	
1	Adi-Dravidar Welfare	3	3.26	1	0.43	4
2	Agriculture	6	1.89	2	8.15	8
3	Animal Husbandry	3	0.06	1	41.57	3
4	Art and Culture	3	0.37	Nil	Nil	3
5	Education	18	1.70	3	43.81	21
6	Electricity	265	4,380.73	1	0.17	266
7	Health and Family Welfare	13	1.37	Nil	Nil	13
8	Information Technology	1	1.15	Nil	Nil	1
9	Labour	2	0.70	Nil	Nil	2
10	Local Administration	5	2.03	1	0.47	6
11	Police	Nil	Nil	2	0.50	2
12	Public Works	2	0.20	2	5.01	4
13	Revenue	2	0.84	Nil	Nil	1
14	Tourism	1	*	Nil	Nil	1
15	Women and Child Development	1	3.75	Nil	Nil	1
	Total	325	4,398.05	13	100.10	338
						4,498.15

Similarly, the table below would be coded under the department dataset. The number of IRs would go in outstanding_IR and the number of paragraphs would go in outstanding_observations.

Details of IRs issued upto March 2016 and paragraphs pending as on September 2016

Sl.No	Name of the Department/ Directorate/Societies	Inspection Reports	Paragraphs
(1)	(2)	(3)	(4)
1	Accounts and Treasuries	20	55
2	Adi-Dravidar Welfare	12	73
3	Agriculture	50	238
4	Animal Husbandry and Animal Welfare	10	71
5	Art and Culture	9	48
6	Civil Supplies and Consumer Affairs	14	54
7	Collegiate and Technical Education	77	353
8	Commercial Taxes	63	353
9	Co-operation	16	70
10	Economics and Statistics	3	5
11	Election	3	13
12	Electricity	32	144
13	Fire Service	2	12
14	Fisheries and Fishermen Welfare	34	151
15	Forest and Wild Life	4	13
16	Heads of State	13	46
17	Health and Family Welfare	73	286
18	Hindu Religious Institutions	—	27
19	Industries and Commerce	36	143

Subsector dataset. Code the table below under subsector. Code the state, the report number, page number table, year the report was published, the general sector, the economic sector, year audited, and all of the numerical values. For the year audited (column 1), we only input the second year out of the range.

TABLE 4.10 - POSITION OF INSPECTION REPORTS

Year	(₹ in crore)											
	Opening balance			Additions during the year			Clearance during the year			Closing balance		
	IRs	Paras	Money value	IRs	Paras	Money value	IRs	Paras	Money value	IRs	Paras	Money value
2006-07	25	54	6.04	2	3	0.03	2	5	Nil	25	52	6.07
2007-08	25	52	6.07	5	10	Nil	3	4	Nil	27	58	6.07
2008-09	27	58	6.07	5	6	0.21	4	7	0.25	28	57	6.03
2009-10	28	57	6.03	2	3	0.16	6	8	0.21	24	52	5.98
2010-11	24	52	5.98	1	2	0.09	1	8	Nil	24	46	6.07
2011-12	24	46	6.07	7	21	132.78	10	20	3.27	21	47	135.58
2012-13	21	47	135.58	6	11	0.41	1	4	0.56	26	54	135.43
2013-14	26	54	135.43	8	23	0.72	Nil	3	0.02	34	74	136.13
2014-15	34	74	136.13	Nil	Nil	Nil	1	3	0.02	33	71	136.11
2015-16	33	71	136.11	3	10	0.82	Nil	Nil	Nil	36	81	136.93

(Source: As per data maintained in the office of the AG (E&RSA), Tamil Nadu)

Appendix C Guatemalan Audits Codebook

C.1 Overview and Scope

This section provides an overview of the Guatemalan audit dataset. It covers data on infractions (*sancciones*), lists of charges (*pliegos de cargos/formulación de cargos*), and whistleblower complaints (*denuncias*). For each of these variables, there is both a count and amount of money in Guatemala.

C.2 Variables in the Dataset

Table C1: Variables in the Guatemalan Audits Dataset

Variable	Definition
<code>department</code>	Department (province/state equivalent) where the audit took place.
<code>municipality</code>	Municipality where the audit took place.
<code>year</code>	Year that the audit finalized.
<code>infractions</code>	Number of infractions in that year's audit(s).
<code>infractions_amount</code>	Amount of stolen or misappropriated money associated with the infractions in that year.
<code>charges_filed</code>	Number of charges filed.
<code>charges_filed_amount</code>	Amount of stolen or misappropriated money associated with the charges filed.
<code>whistleblower_complaints</code>	Number of whistleblower complaints filed in that year.
<code>whistleblower_complaints_amount</code>	Amount of stolen or misappropriated money associated with the whistleblower complaints filed in that year.

C.3 Coding Process

The Guatemalan audits do not require a complicated coding process. Coders can simply access the audits from the website of the Comptroller General of Accounts (Contraloría

General de Cuentas), which is the supreme audit agency of Guatemala. Here is a screenshot of the yearly 2007 report.

No.	Municipalidad	SANCIONES		PLIEGOS DE CARGOS		DENUNCIAS	
		Cant	Monto	Cant	Monto	Cant	Monto
1	ACATENANGO	16	280,000.00	3	120,507.76	0	0.00
2	AGUA BLANCA	3	28,000.00	0	0.00	0	0.00
3	AGUACATAN	1	4,000.00	0	0.00	0	0.00

Some of the yearly reports, such as that of 2014 (below), are organized by department.

DEPARTAMENTO DE IZABAL

No.	Departamento	Denuncias		Sanciones		Formulaciones	
		Cantidad	Monto Q	Cantidad	Monto Q	Cantidad	Monto Q
1	EL ESTOR	0	0.00	5	85,355.00	0	0.00
2	LIVINGSTON	0	0.00	2	24,000.00	0	0.00
3	LOS AMATES	0	0.00	3	25,600.00	0	0.00
4	MORALES	0	0.00	4	32,686.07	0	0.00
5	PUERTO BARRIOS	0	0.00	8	80,700.91	0	0.00
TOTALES		0	0.00	22	248,341.98	0	0.00

Other years, including 2016 (below), have two reports per year. To obtain yearly totals by municipality, I simply added the infractions, charges filed, whistleblower complaints, and their corresponding amounts. As stipulated above, I do not include

No.	MUNICIPALIDADES Y SUS EMPRESAS	SANCIONES		FORMULACIÓN DE CARGOS		DENUNCIAS	
		Cantidad	Monto en Quetzales	Cantidad	Monto en Quetzales	Cantidad	Monto en Quetzales
1	Empresa Municipal de Agua de la Ciudad de Guatemala	10	125,437.50				
2	Empresa Eléctrica Municipal de Guastatoya, El Progreso					2	794,125.06
3	Municipalidad de Aguacatán, Huehuetenango					1	2,163,791.55
4	Municipalidad de Antigua Guatemala, Sacatepéquez			2	605,727.40	2	0.00
5	Municipalidad de Cabricán, Quetzaltenango					1	4,015.00

Appendix D Mexican Audits Codebook

D.1 Overview and Scope

Mexico has a two-tiered auditing system. State-level auditing entities (EFSE, *las Entidades de Fiscalización Superior Estatales*) conduct routine municipal audits and report to the respective state governor. The supreme audit agency (ASF, *la Auditoría Superior de la Federación*), which reports to the Mexican Chamber of Deputies, undertakes audits of federal transfers to municipalities. Given that EFSEs tend to be not very independent and effective (OECD, 2017), the present dataset focuses on the ASF audits.

Table D2: ASF Audits Conducted by Fund/Expenditure Category (2007-2018)

	Fund/Expenditure Category	Count	Percent
1	Fund for Strengthening the Boroughs of Mexico City's Social Infrastructure	1313	27.84
2	Fund for Municipal Social Infrastructure	1263	26.78
3	Fund for Strengthening the Boroughs of Mexico City	789	16.73
4	Fund for Subsidizing Public Security for the Boroughs of Mexico City	524	11.11
5	Federal Investments in Municipalities	492	10.43
6	Evaluation System for Performance of Federal Spending	81	1.72
7	Fund for Strengthening of Municipal and State Infrastructure	78	1.65
8	Fund for Paving, Sporting Infrastructure, Public Lighting, and Rehabilitation of Educational Infrastructure for the Boroughs of Mexico City	54	1.14
9	Municipal Institutional Development Trust Fund	35	0.74
10	Compliance with Law of Financial Discipline	31	0.66
11	Fund for Sporting Infrastructure	18	0.38
12	Regional Development Projects	16	0.34
13	Fund for Culture	14	0.30
14	Fund for Paving and Municipal Development	6	0.13
15	Funds from Branch 33 and Their Unspent Balances	1	0.02
16	Habitat Fund	1	0.02
17	Youth Power Program	1	0.02

Table D2 provides a breakdown of the ASF audits by fund type. The ASF also undertakes a great diversity of audits and takes many different types of actions after these audits (see Tables D3 and D4).

Table D3: Types of ASF Audits (2007-2018)

	Audit Type	Count	Percent
1	Financial and Compliance	3667	77.74
2	Compliance	354	7.50
3	Financial Compliance with Performance Focus	252	5.34
4	Financial Compliance	202	4.28
5	Compliance and Performance	134	2.84
6	Performance	97	2.06
7	Physical Investment	6	0.13
8	Forensic	5	0.11

Table D4: Types of ASF Audit Actions (2007-2018)

	Actions	Count	Percent
1	Recommendation	12321	45.85
2	Statement of Financial Irregularities with Presumed Intent	5345	19.89
3	Punitive Noncompliance with Regulations	4854	18.06
4	Financial Irregularity with Required Compensatory Action	1689	6.29
5		1481	5.51
6	Request for Clarification	418	1.56
7	Performance Recommendation	391	1.46
8	Tax Evasion or Financial Regulatory Noncompliance with Presumed Intent	222	0.83
9	Report of Crime	150	0.56

D.2 Variables in the Dataset

Table D5: Variables in The Dataset

Variable	Definition
<code>state</code>	State where the audit took place.
<code>municipality</code>	Municipality where the audit took place.
<code>year</code>	Year that the audit finalized.
<code>audit_dummy</code>	municipality audited dummy variable.
<code>audit_count</code>	times municipality audited (all funds).
<code>infractions</code>	total number of infractions (all funds).
<code>log_money_missing</code>	Log amount of stolen/misappropriated money (previously deflated to constant 2013 Mexican pesos.)
<code>money_missing</code>	Amount of stolen/misappropriated money in constant 2013 Mexican pesos.

Continued on next page

Table D5: Variables in the Dataset – *continued*

Variable	Definition
<code>money_sample</code>	Amount of money audited in constant 2013 Mexican pesos.
<code>money_sample_share</code>	Share of audited money as a percent of the total.
<code>money_universe</code>	Amount of potentially auditable money in constant 2013 Mexican pesos.
<code>audit_compliance</code>	Number of compliance audits.
<code>audit_compliance_and_perf</code>	Number of compliance and performance audits.
<code>audit_financial_and_compli</code>	Number of financial and compliance audits.
<code>audit_financial_compliance</code>	Number of financial compliance audits.
<code>audit_fin_compli_with_perf</code>	Number of financial compliance with performance focus audits.
<code>finding_fin_irreg_payback</code>	Number of definitive financial irregularities with required compensatory actions.
<code>audit_forensic</code>	Number of forensic audits.
<code>audit_performance</code>	Number of performance audits.
<code>audit_physical_investment</code>	Number of physical investment audits.
<code>finding_perf_rec</code>	Number of definitive financial irregularities with required compensatory action.
<code>finding_noncompliance</code>	Number of instances of punitive noncompliance with regulations.
<code>finding_recommendation</code>	Number of audit recommendations.
<code>finding_crime</code>	Number crime reports.
<code>finding_clarification</code>	Number of requests for clarifications.
<code>finding_fin_irreg</code>	Number of statements of financial irregularities with presumed intent.
<code>finding_none</code>	Number of audits without any infractions.
<code>finding_tax_evasion_fin_irreg</code>	Number of instances of tax evasion or financial regulatory noncompliance with presumed intent.
<code>fund_branch33_unspent</code>	Number of audits regarding unspent funds from branch 33.
<code>fund_culture</code>	Number of Culture Fund audits.
<code>fund_eval_perf_fed_spend</code>	Number of audits concerning the evaluation system for the performance of federal spending.
<code>fund_fed_inv_muni</code>	Number of audits concerning federal investments in municipalities.
<code>fund_fism</code>	Number of audits concerning the Fund for Municipal Social Infrastructure.
<code>fund_habitat</code>	Number of audits concerning the Habitat Fund.
<code>fund_law_finan_discipline</code>	Number of audits concerning the law for financial discipline.

Continued on next page

Table D5: Variables in the Dataset – *continued*

Variable	Definition
<code>fund_muni_institution_dev</code>	Number of municipal institutional development trust fund audits.
<code>fund_muni_state_infras</code>	Number of fund for municipal and state infrastructure audits.
<code>fund_paving_muni_dev</code>	Number of fund for paving and municipal development audits.
<code>fund_pub_goods_cdmx</code>	Number of fund for paving, sporting infrastructure, public lighting, and rehabilitation of educational infrastructure for the boroughs of Mexico City audits.
<code>fund_region_dev</code>	Number of regional development project audits.
<code>fund_security_cdmx</code>	Number of fund for subsidizing public security for the boroughs of Mexico City audits.
<code>fund_strength_cdmx</code>	Number of fund for strengthening the boroughs of Mexico City audits.
<code>fund_social_infras_cdmx</code>	Number of fund for strengthening the boroughs of Mexico City's social infrastructure audits.
<code>fund_sport_infras</code>	Number of fund for sporting infrastructure audits.
<code>fund_youth_power</code>	Number of fund for youth power audits.

Appendix E Technical Appendix: Survival Analysis of the Honduran Audit Distribution

This section provides an overview of the data and survival methods used to test whether partisan affiliation affected the intensity or dosage of the municipal audits undertaken by Honduras' Supreme audit agency (*Tribunal Superior de Cuentas*, TSC).

E.1 Data for Survival Analysis

Before conducting the actual survival analysis, it was necessary to transform data described in Section 3.2.4 and Appendix A into municipality-day format. Honduras has general elections for its president, congress, governors, and mayors every four years in November, and there is approximately a two-month delay between the elections and when the politicians take office in late January. As Table E6 shows, the presidential terms start around January 27-28, leaving most of the month of January under the mandate of the previous presidents. Although that is less than one month of the year, those 27-28 days are very significant. For example, all 247 audits under Flores' tenure started between January 25-27, 2002.

In total, expanding the 935 audits that the TSC started from January 25, 2002 - December 31, 2018 resulted in a final dataset of 1.8 million observations. These 1.8 million correspond to the fact that Honduras has 298 municipalities, there are 17 years of available data, and there are 365 days per year in which municipalities could receive a new audit.

E.2 Survival/Event History Methods

In this paper, I employ a diverse array of event-history models, including the semi-parametric Cox regression as well as parametric Weibull, Gompertz, Exponential, Log-Normal, and Log-Logistic models. Of these methods, my preferred event-history specification

Table E6: Electoral Term Start Dates and Audits by Term

President	Electoral Term Dates	Audits Started
Flores	1/27/1998-1/27/2002	247
Maduro	1/28/2002-1/27/2006	132
Zelaya/Micheletti	1/28/2006-1/27/2010	372
Lobo	1/28/2010-1/27/2014	150
Hernandez	1/28/2014-present	150

Note: Zelaya/Michelletti corresponds to the fact that Honduras had a promisory coup in 2009, during which Micheletti took over for Zelaya and offered to have new elections and not run in them ([Bermeo, 2016](#)). Because Micheletti was from the same party as Zelaya, and Micheletti only served for 9 months, these can be considered as part of the same term.

is the multi-episode semi-parametric Cox proportional hazards model:

$$L^p = \prod_{k \in D} \prod_{i \in \epsilon_k} \frac{\exp(A^{(k)}(t_i)\alpha^{(k)})}{\sum_{l \in R(t_i)} \exp(A^{(k)}(t_l)\alpha^{(k)})} \quad (1)$$

“where $A^{(k)}(t)$ is the vector of covariates, specified for the transition to destination state k , $\alpha^{(k)}$ is the vector of associated covariates, and R is the risk set” ([Blossfeld, Golsch and Rohwer, 2007](#), 225). In my case, the risk set corresponds to not being under audit (the origin) and the destination state (being under audit), which can happen more than once. Since there is no intercept in the Cox model, the baseline hazard absorbs the constant. I adjudicate ties with the Breslow method.

I prefer the semi-parametric Cox model to the other event-history ones because I have no specific predictions concerning the time dependence, I have no particular theory concerning how time duration impacts the model, and I am more interested the effects and their direction than time. Given these limitations, it is difficult to hypothesize about the utility of any parametric model for my data. As [Blossfeld, Golsch and Rohwer \(2007\)](#) emphasize, theoretical conjectures about the particular functional form of event-history models are very challenging.

Many analysts would argue that it is necessary to test the proportionality assumption of the Cox model, and choose a parametric model in the case of nonproportionality. In my

case, the proportionality assumption is most certainly violated. That is largely a function of having to disaggregate the data to the municipality-day level, which enlarged my dataset to more than 1.8 million observations with unequal groups sizes.

Appendix F Additional Regressions

F.1 Honduras Ethnicity

To test the hypothesis that ethnicity predicts which municipalities receive audits in Honduras, I compiled data on the percent of indigenous peoples living in each municipality from the 2013 census. The groups included in my count were the Garifuna, Lenca, Maya-Chorti, Miskito, Nahua, Tawahka, and Tolupan peoples. I also included the census category of “other” indigenous peoples.

Unfortunately, yearly panel data on ethnic compositions do not exist in Honduras. Accordingly, I ran a cross-sectional regression, using the indigenous peoples’ share of each municipality’s total population in 2013 as the main independent variable and the total number of ASF audits received for all available periods (2002-2018) as the dependent variable. As control variables, I added the (log) population in 2013, the poverty rate in 2013 (measured via unmet basic needs), and the mean of party alignment. In all specifications in Table F1, I find that municipalities with higher shares of indigenous peoples are more likely to receive an audit.

F.2 Mexico Regressions

Table F1: Does Ethnicity Predict Audits in Honduras?

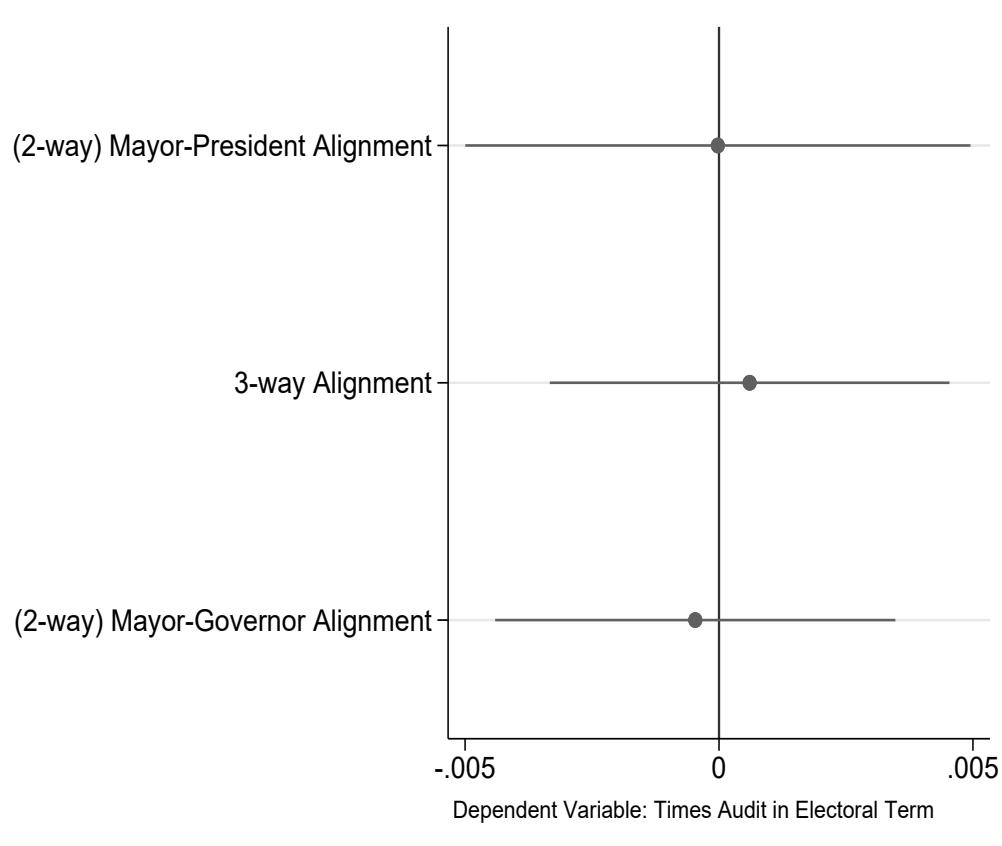
	(1) No. of Audits	(2) No. of Audits	(3) No. of Audits	(4) No. of Audits
No. of Audits				
Indigenous peoples (%)	0.272*** (0.081)	0.289*** (0.084)	0.315*** (0.093)	0.316*** (0.095)
Population (log)		0.031 (0.047)	0.019 (0.045)	0.020 (0.045)
Poverty rate			-0.324 (0.218)	-0.319 (0.218)
Party Alignment (mean)				-0.019 (0.136)
Constant	1.095*** (0.034)	0.801* (0.447)	1.108*** (0.429)	1.112*** (0.430)
Observations	298	298	296	296

Standard errors clustered by municipality in parentheses.

Note: Cross-sectional Poisson regression. Standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure F.1: Does Alignment Predict More Audits at the State Level?



Note: Estimations contain state and year fixed effects. Standard errors are clustered at the state level.