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

Michael Denly†

January 30, 2024

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

www.mikedenly.com/research/audit-measurement

Abstract

Audits provide one of the most effective ways to measure as well as mitigate corruption. However, audits can be biased along many different dimensions, so corruption researchers need a framework to assess audit data quality. I respond to this need by developing a three-pillar framework emphasizing: 1) the audit institution's political independence; 2) the absence of distributional, group-level biases, such as toward particular ethnicities or opposition-party politicians; and 3) the lack of implementation or dosage biases across similar types of audits. I demonstrate the utility of the framework by collecting and analyzing a massive new dataset of subnational audits without random assignment from India, Mexico, Honduras, and Guatemala. Using a regression discontinuity design based on party alignment and various other regression methods, I find that the audit data are mostly unbiased along the aforementioned dimensions. Especially given the lack of governments conducting randomly assigned audits of their subnational units, the new audit data and framework proposed in this paper are crucial for advancing knowledge about corruption and how to combat it.

JEL codes: D73, H8, M42

*For funding, I thank Texas A&M University and 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, Alexander Ivey, Amila Lulo, Ana Ramirez, Ana Suarez, Ansh Samdaria, Anoushka Ramesh, Augustine D'Eramo, Awab Ahmed, Benjamin Clardy, Blaine Finstein, Braden Holt, Briana Gonzalez, Brayan Salgado, Briana Rocha, Caleb Rudow, Calista LaMotte, Camila Aguilar, Carlo Moran, Carmen Magana, Caroline Hastings, Caroline Woodman, Connor Hanna, Dalia Dabbagh, David Tibaquira, David Zulli, Debasmita Bhakta, Diana Geng, Dorna Abdi, Dulce Alonso, Elena Solimano, Emily Ibarra, Erin Eggleston, Ethan Tenison, Evan Shimek, Evelin Caro Gutierrez, Fabiola Ramirez, Felipa Mendez, Francesca D'Annunzio, Geraldyn Campos, Grace Thomas, Isabel Campos, Isabel Ayala, Ishani Pandya, Iulia Tothezan, Ivana Jelensky, Jenny Rodriguez, Joelean Hall, Jonah Isaac, Josh Meuth Alldredge, Jude Rios, Judy Lane, Kaetana Mendoza, Kais Bhandari, Karan Kanatala, Kate Hopp, Katya Bandouil, Katherine Alonzo, Kelsia Adil, Leah Lewis, Lissette Almanza, Luca Venegoni, Luisa Venegoni, Magdalena Ibarra, Marco Mejia, Maria Fernanda Guerrero, Marianne Nader, Mary White, Michael Fernandez, Mishan Kara, Mobin Piracha, Mohamed Abulfalgha, Monserrat Ramon, Nathan Stern, Nicholas Galindo, Nick Hoffman, Nicole Domingo, Nicky Pownall, Paige Johnson, Paul Orszag, Payton Bellman, Preston Saunders, Rachel Boles, Rachel Rosenburg, Regina Cruz, Ryan Jabra, Saamia Imtiaz, Samuel Gilette, Sara Lowe, Sofia Gonzalez, Sterling Mosley, Tara Jackson, Temi Ajibola, Thomas Boswell, Ximena Mercado Garcia, Vanessa Crespo, Vanessa Gonzales, Vanessa Lizcano, and Veronica Lulu Stracher. For advice or feedback, I thank Akshat Gautam, Arun Manuja, Dan Nielson, Horacio Larreguy, Ilona Wysmulek, Jan Pierskalla, John Gerring, John Ogallo, Ken Greene, Lorenzo Crippa, Mateo Montenegro, Mike Findley, Serge Nkuindja, and Xiaobo Lü. All errors are those of the author. Interviews referenced in the paper were approved by the UT Austin IRB under protocols 2016-04-0099, 2018-07-0062, and 2018070062-MODCR01.

† Assistant Professor, Texas A&M University, Bush School of Government & Public Service, 4220 TAMU, College Station, Texas, USA, +1 (202) 539-0452, ✉ mdenly@tamu.edu

Since the publication of Mauro's (1995) seminal study around 30 years ago, empirical scholarship on governmental corruption has primarily relied on perceptions data.[1] That is a grave problem for knowledge and theory creation, as corruption perceptions often differ markedly from actual corruption levels (Olken, 2009). The halo effects, construct validity challenges, and questionable error structures inherent to perceptions data constitute only a few reasons why (see Appendix A).

More recently, analysts have used many more objective measures of corruption, among which audits stand out as perhaps the most promising of the perceptions data alternatives. Corruption is a multifaceted, complex, and clandestine phenomenon that is hard to define,[2] but audits capture that complexity. Not only do audits capture many of the objective corruption measures in the literature (e.g., procurement flags, abnormal asset growth, and political connections),[3] but they also capture fraud, theft, nepotism, obstruction, and more. Consequently, audits also outperform experience-based measures in terms of scope and content validity, including Transparency International's (2021) Global Corruption Barometer, websites like ipaidabribe.com, and the World Bank's (2021) Enterprise Surveys. Given that audits are specific to individuals or institutions, audit also overcome the level-of-analysis problems that plague most research using perceptions data (see Gingerich, 2013*a*).

To date, scholars using audit data have focused mostly on the randomly assigned Brazilian municipal audits of federal transfer fund spending.[4] On the one hand, these data from

[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] "The misuse of public office for private gain" is perhaps the most common definition of corruption, but Rose-Ackerman and Palifka (2016) provide more context and explain the various forms of corruption.

[3] Procurement: See, for example, Bandiera, Prat and Valletti (2009), Mironov and Zhuravskaya (2016), and Fazekas, Cingolani and Tóth (2018). Asset declarations: See, for example, Eggers and Hainmueller (2009) and Fisman, Schulz and Vig (2014). Political connections: See, for example, Fisman (2001), Faccio (2006), and Fisman and Wang (2015).

[4] Bobonis, Cámara Fuertes and Schwabe (2016), Buntaine et al. (2018), Chong et al. (2015), Arias, Balán, Larreguy, Marshall and Querubín (2019), Arias, Larreguy, Marshall and Querubín (2019), Arias et al. (2022), Larreguy, Marshall and Snyder (2020), Ajzenman (2021), Chu et al. (2021), and Berliner and Wehner (2022) are, to my knowledge, the only studies using audits outside of Brazil—aside from those on one-off programs, such as Di Tella and Schargrodsky (2003), Olken (2007, 2009), Nikolova and Marinov (2017), and De La O, González and Weitz-Shapiro (2023).

Brazil are useful because they overcome endogeneity problems and allow for causal estimation when used as an independent variable. In turn, scholars have used these Brazil data to produce many useful findings.[5] On the other hand, the Brazil data are not as useful as a dependent variable: each of the audit lotteries/waves has different fiscal focuses (e.g., health, education, agriculture),[6] so the data are not fully comparable across waves.

Fortunately, some other countries disseminate subnational audit data relevant for capturing corruption, but the data are often not immediately useful for researchers. Notably, the data are mostly not available in machine-readable format, and researchers need a framework to assess the validity of these data. That is especially accurate when the audits are not randomly assigned, which is the case for all countries outside of Brazil for the 2002-2015 period.[7] On that score, an in-person interview with one of the heads of the supreme audit agencies whose data I use for the empirical analysis in this paper reveals that there is a significant lack of policy interest for conducting randomly assigned audits.[8]

In this paper, I address both of the above challenges by not only proffering new, open-source, audit-derived corruption data but also a new framework to validate their use for causal inference. The framework's first pillar concerns whether the audit agency has *de facto* political independence and a legal (*de jure*) basis for it to be credible. Second, given that most countries' audit agencies have discretion to perform risk-based audits based on relevant vulnerabilities, the frequency, selection, or distribution of audits must not exhibit bias toward any 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,

[5]See, for example, Ferraz and Finan (2008), Ferraz, Finan and Moreira (2012), Melo, Pereira and Figueiredo (2009), Pereira and Melo (2015), Timmons and Garfias (2015), Bologna and Ross (2015), Bologna (2016, 2017), Cavalcanti, Daniele and Galletta (2018), Funk and Owen (2020), and Colonnelli and Prem (2022).

[6]For example, lottery/wave 21 focuses on "housing, sanitation and urban planning", whereas lottery/wave 22 focuses on "agricultural organizations, energy, and environmental management". See link here.

[7]See Ferraz and Finan (2018) and Odilla and Rodriguez-Olivari (2021) for more on Brazil's randomized audit program and its termination in 2015.

[8]Interview held December 5, 2023.

opposition-party politicians must not be subject to more stringent audits than those sharing the same party as the executive.

I demonstrate the utility of my framework by collecting and analyzing a massive, new, micro-level dataset of subnational audits without random assignment from India, Mexico, Honduras, and Guatemala. Given that Pillar 1 on political independence involves using data from the World Bank's ranking audit institutions as well as my own supplemental qualitative analysis, I begin empirical testing with Pillar 2 on the audit distributions. To test the unbiasedness of the latter, I use a regression discontinuity design (RDD) based on party alignment as well as count regression models. For Pillar 3 on audit dosage, I complements RDDs and count models with survival analysis. With the potential exception of ethnic biases analysis,[9] I find that the audits mostly do not exhibit biases according to the above framework. Especially because audits are one of the most effective tools to tackle corruption (Gans-Morse et al., 2018; Lagunes, 2021), the new data and framework proposed in this paper provide an essential toolkit for researchers, governments, NGOs, and development institutions.

1. New Subnational Audit-Based Corruption Data

To address the dearth of political corruption studies employing audit data outside of Brazil (see Table 1), I collected new subnational audit data from Mexico, India, Honduras, and Guatemala. I selected these cases not only due to data availability and language familiarity but also because they are 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, almost no INTOSAI member countries make their subnational audits easily accessible to citizens or researchers.[10]

[9] See Section 2.2.2

[10] Obtaining and cleaning the data involved considerable work (see Appendices for codebooks).

Table 1: New Data and Existing, Publicly-Available Subnational Audit Data

Panel A: Audit Data

	New Data (My Contribution)				**Others**				
	Honduras	**Guatemala**	**Mexico**	**India**	**Brazil**	**China**	**Puerto Rico**	**South Africa**	**Uganda**
Administrative unit	Municipal	Municipal	Municipal	State	Municipal	Prefecture	Municipal	Municipal	District
Funds covered	All	All	Federal only	All	Federal only	All	All	All	All
Years covered	2002-2018	2004-2019	2007-2018	2004-2021	2003-2015	2006-2016	1987-2005	2007-2015	2013-2014
Number of audits	900	3,350	3,211	1,496	2,200	2,940	326	X	111
Audited annually %	21%	88%	11%	100%	3%	79%	23%	X	100%
Sector/Fund details	No	No	Yes	Yes	Yes	No	No	No	No
Sub-sector details	No	No	No	Yes	No	No	No	No	No
Department details	No	No	No	Yes	No	No	No	No	No
Sub-sector observations	n/a	n/a	n/a	9,248	n/a	n/a	n/a	No	No
Department observations	n/a	n/a	n/a	15,499	n/a	n/a	n/a	No	No
Infractions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Details of infractions	Yes	No	Yes	Partial	Yes	Yes	Yes	No	Yes
Money stolen/missing	No	Yes	Yes	Yes	Yes	Yes	No	No	No
Follow-up data	No	No	No	Yes	No	No	No	No	No
Whistleblower complaints	No	Yes	No	No	No	No	No	No	No
Charges filed	No	Yes	Yes	No	No	No	No	No	No
Money audited	No	No	Yes	Partial	Yes	No	No	No	No
Money outside audit	No	No	Yes	No	Yes	No	No	No	No
Matching election data	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes

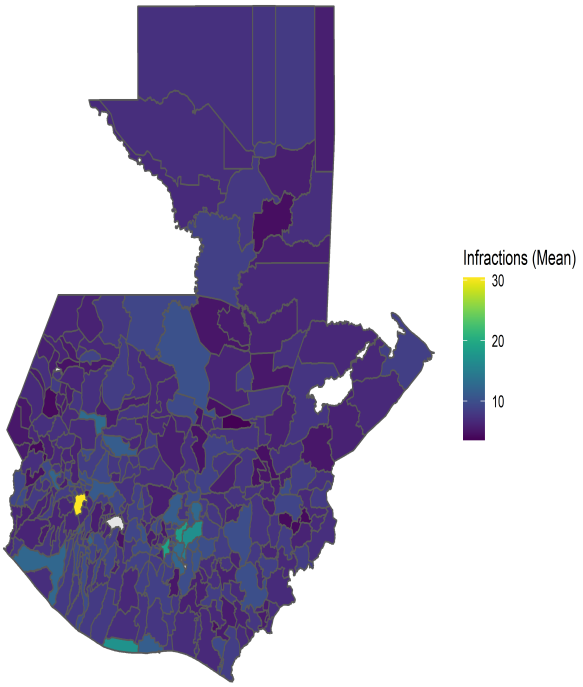
Panel B: World Bank Supreme Audit Institution (SAI) Independence Scores

Country	Score	Effective Rank	Constitution	SAI Head Appointment Transparency	Financial Autonomy	Audit Types	Operational Autonomy	Staffing	Audit Mandate	Audit Scope Autonomy	Information Transparency	Audit Report Obligation
South Africa	10/10	1/16	1	1	1	1	1	1	1	1	1	1
Brazil	9/10	3/16	1	0.5	1	1	1	1	1	1	1	0.5
China	9/10	3/16	1	1	0.5	1	1	0.5	1	1	1	1
Mexico	9/10	3/16	1	0.5	1	1	1	0.5	1	1	1	1
Uganda	9/10	3/16	0.5	1	0.5	1	1	1	1	1	1	1
Guatemala	8.5/10	4/16	1	0.5	0.5	1	1	0.5	1	1	1	1
India	8.5/10	4/16	1	0.5	1	1	1	0.5	1	1	0.5	1
Honduras	7/10	7/16	0.5	0.5	1	1	0.5	0.5	1	0.5	0.5	1

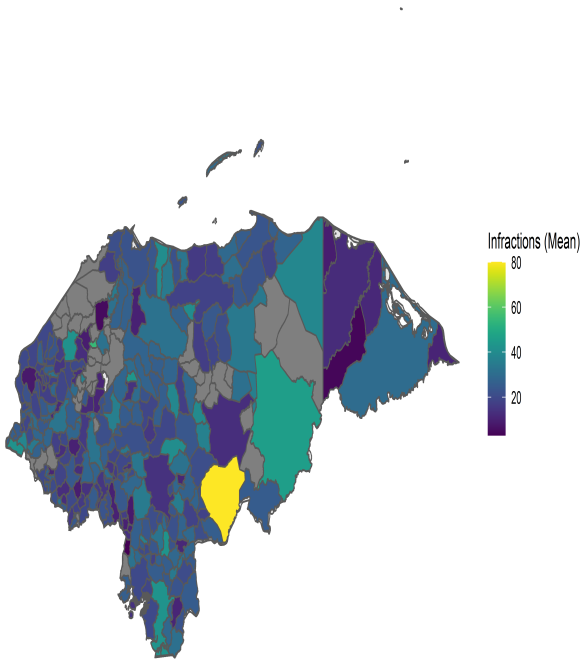
Note: See Gurazada et al. (2021) for more on the World Bank ranking of audit agencies, which receive scores of 0, 0.5, and 1.0 for each indicator. See Appendix B for the Honduras codebook; Appendix C for the India codebook; Appendix D for the Guatemala codebook; Ferraz and Finan (2018) for more on the Brazil data; Chu et al. (2021) for more on the China data; Bobonis, Cámara Fuertes and Schwabe (2016) for more on the Puerto Rico data; Berliner and Wehner (2022) for more on the South Africa data; and Buntaine et al. (2018) for more on the Uganda data. Chong et al. (2015), Arias, Balán, Larreguy, Marshall and Querubín (2019), Arias, Larreguy, Marshall and Querubín (2019), Arias et al. (2022), Larreguy, Marshall and Snyder (2020), and Ajzenman (2021) also use audit data from Mexico, but they focus exclusively on audits from one specific fund related to infrastructure. The Mexico codebook in Appendix E provides more details on the 17 different Mexican funds included in the present analysis.

Figure 1: Maps of Subnational Corruption Data by Administrative Unit

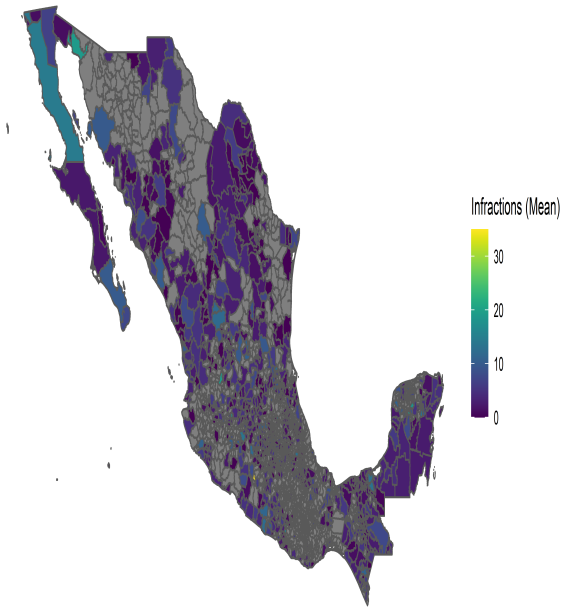
(a) Mean Infractions in Guatemala, 2004-2019



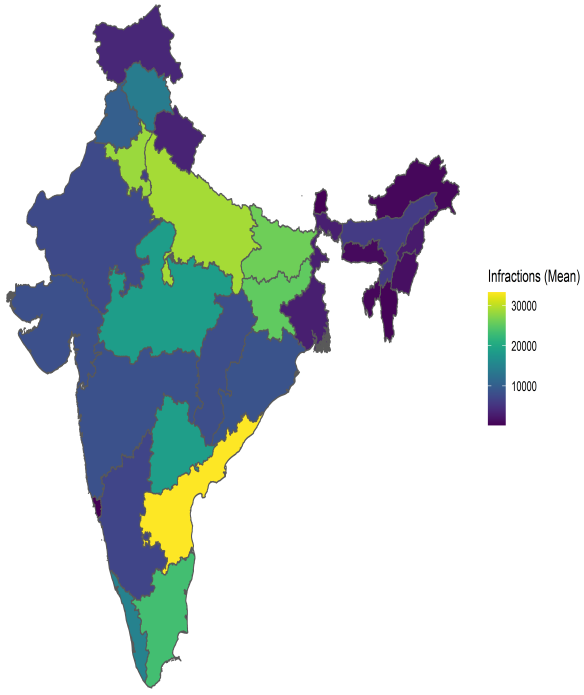
(b) Mean Infractions in Honduras, 2002-2018



(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.

Panel A of Table 1 summarizes the new audit data advanced in this paper. 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 municipals expenses. By contrast, the Mexican municipal audits examined in this study follow those from the highly-studied ones in Brazil and are more limited in scope, pertaining only to federal funding. Unlike existing papers using the Mexican audits,[11] which only focus on audits of one infrastructure fund, the present study examines audits from 17 different funds.

With respect to the geographical coverage of the audits, all of the countries in the new dataset provide higher annual shares of audited subnational units than Brazil (3%).[12] Similarly, none of the sample countries restrict audits to smaller municipalities above certain population thresholds like in Brazil. However, Honduras (21%) and Mexico (11%) audited lower shares of their municipalities each year than Puerto Rico (22%),[13] China (79%),[14] South Africa (100%), and Uganda (100%). All of the countries provide long time-series of data except Uganda, for which Buntaine et al. (2018) only provide one fiscal year of data.

In terms of the specific types of corruption data included in the new audit dataset, three out of the four countries provide details on either the count of infractions and/or the amount of misappropriated/stolen money detailed in the audits. Both Mexico and Guatemala have data on the charges filed. India is the only country that provides the audit data by department and sub-sector details, and it is also the only country that provides panel data on audit follow-up. These follow-up data are particularly crucial: if researchers

[11] See Chong et al. (2015), Arias, Balán, Larreguy, Marshall and Querubín (2019), Arias, Larreguy, Marshall and Querubín (2019), Arias et al. (2022), Larreguy, Marshall and Snyder (2020), and Ajzenman (2021).

[12] Calculation: $1,881/(5,568*12) = 2.8\%$. Note that there are 5,568 municipalities in Brazil; the randomized audit program lasted effectively for 12 years (2003-2014), as the final year (2015) only had very few audits; and 1,881 unique municipalities had been selected for audit through 2014 (Colonnelli and Prem, 2022, 699).

[13] Calculation: $326/(78*19) = 22\%$. Note: Bobonis, Cámara Fuertes and Schwabe (2016) provide full results for 326 unique municipalities, spanning 1987-2005 (5 elections), and there are 78 municipalities in Puerto Rico.

[14] Calculation: $2,940/(339*11) = 79\%$. Note that Chu et al. (2021) provide data for 2,940 reports across 11 years, and there are 339 prefectures in China.

are to learn what reduces corruption, then it is necessary to know the conditions under which politicians and bureaucrats take action to reduce corruption after they are alerted to it.

2. 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 political 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.

2.1. Pillar 1: Political Independence

For audit agencies to produce credible corruption data, they must enjoy political independence to pursue unbiased auditing and have a legal basis for it. Political independence ostensibly has a *de facto* self-enforcing, norms component to it (see Weingast, 1997). However, it is the *de jure*, legal basis for independence that makes an agency's commitment to unbiased auditing credible (see North and Weingast, 1989; Acemoglu, Johnson and Robinson, 2005). In particular, an audit agency's independence is most credible when it enjoys protections from the executive branch and is its own institutional veto player,[15] which is most often necessary to specify in the constitution (INTOSAI, 2019).

To date, there is only one index from the World Bank that measures the independence of different countries' supreme audit institutions (Gurazada et al., 2021). Besides constitutional protection, the World Bank SAI Index defines independence on the basis of 9 other criteria: appointment process transparency for the SAI head; financial autonomy; audits type diversity; operational autonomy; staffing; mandate to decide on audit scope; access to records and information; and audit report rights and obligations. Each SAI then receives

[15] Tsebelis (2002, 2) defines a veto player as the actor(s) who must agree for a policy to change.

a final 0-10 score, ranging from 10 (only South Africa and Seychelles) to 2.5 (only Chad). Given that there are 16 different scores in the SAI index, the SAIs of Mexico, Guatemala, India, and Honduras all have effective ranks above the average (see Panel B of Table 1). Additionally, only Honduras' SAI has a rank that is not equal to or simply one half point below that of Brazil. Overall, the independence of the audit institutions whose data this paper is advancing is relatively strong. I support this conclusion with my own qualitative assessment of each country in the sample's supreme audit institution in Appendix F as well.

2.2. 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 selectorate. 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 imply that random assignment immediately solves this second pillar, some forms of random assignment are more credible than others,[16] and there is an even more significant problem with random assignment for audits: no SAIs are currently undertaking randomized audits of subnational audits since those in Brazil ended in 2015. Along those lines, given that SAIs are mostly undertaking risk-based audits at bureaucrats' discretion, quantitative analysis of the respective audit distributions 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.[17] When doing so, it is useful to take into account demographics such

[16] For example, complete randomization, block randomization, and stratified sampling are more credible than simple random assignment (see Gerber and Green, 2012).

[17] It is possible to estimate the conditional mean via numerous methods. I provide examples in Section 1 below.

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),[18] examining 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 the data are likely biased.

2.2.1. Are Guatemalan Municipal Audits Politically Biased?

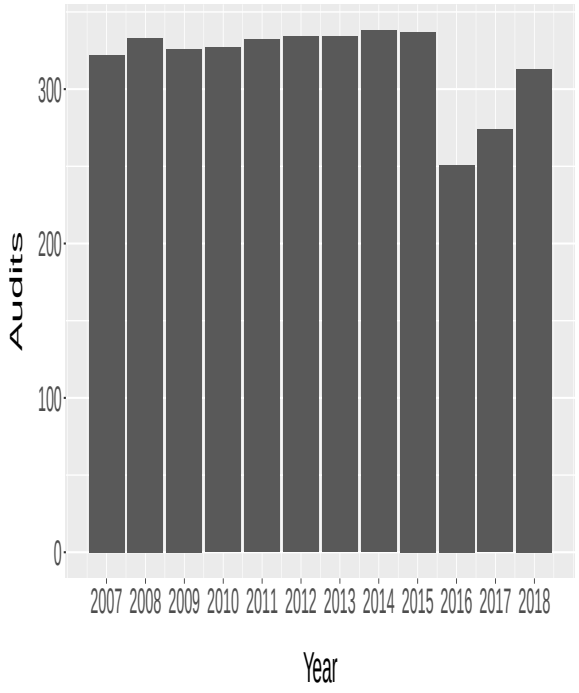
Below, I use the Guatemalan audit data to demonstrate how to examine their distribution for political bias, and Appendix G contains the relevant analyses for the other sample countries. Each year, the Guatemalan Comptroller General (CGC, *Contraloría General de Cuentas*) audits circa 310 of Guatemala's 340 municipalities (see Figure 2a).[19] Even though there is not much variation in which municipalities the CGC audits each year, there is a potential for audit distribution bias across the four-year electoral terms. I thus analyze that potential bias by electoral term and start with the RDD given potential endogeneity pressures as well as the vulnerability of electorally-competitive areas to audit distribution bias. Following Brollo and Nannicini (2012) and Marshall (2023), the parameter of interest for the

[18] Note: In Brollo and Nannicini's (2012) RDD setup, the treatment is alignment (or non-alignment for the opposition party), and the running variable is the margin of victory for the aligned or unaligned/opposition party.

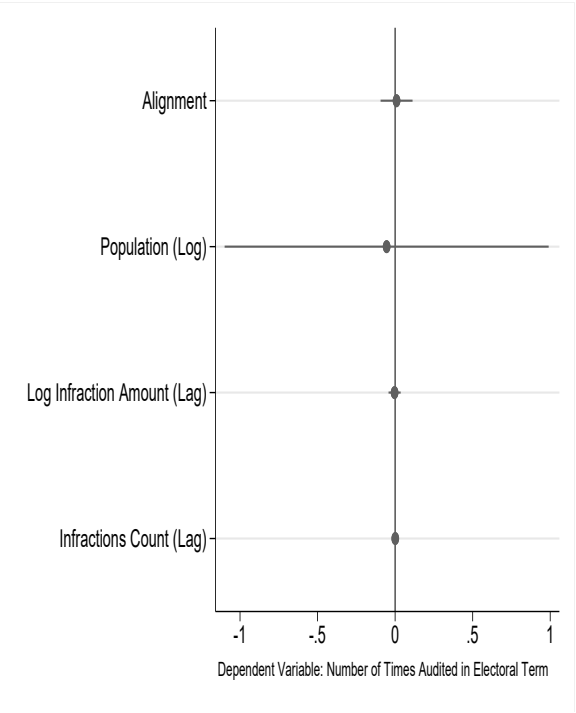
[19] This pattern starts in 2007. From 2004-2006, the Comptroller General undertook much fewer audits.

Figure 2: Main Results: Guatemala's Municipal Audit Distribution

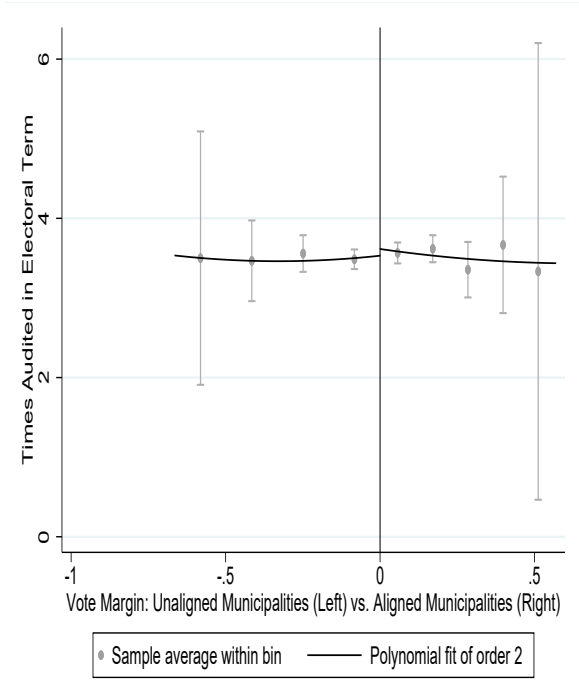
(a) Number of Audits Conducted by Year



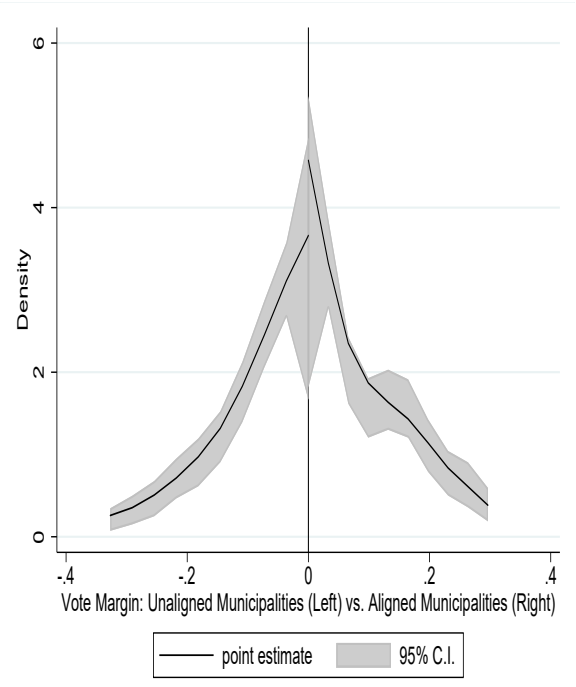
(b) Poisson Model with Municipality and Term FE



(c) RD Results for Audit Count



(d) McCrary Density Plot for RD



sharp RDD is a compound local average treatment effect, τ:

$$\begin{aligned}\tau &= \mathbf{E}[a_{(\text{aligned},it)} - a_{(\text{unaligned},it)} | MV_{it} = 0] \\ \tau &= lim_{MV \downarrow 0}\ \mathbf{E}[a_{it} | MV_{it} = h \geq \text{MV} > c] - lim_{MV \uparrow 0}\ \mathbf{E}[a_{it} | MV_{it} = -h \leq \text{MV} < c]\end{aligned} \tag{1}$$

where a_{it} reflects the number of audits conducted in the aligned/unaligned municipality i in the most recent electoral term t; the running variable, MV_{it}, is the margin of victory for the aligned/unaligned mayor i in the most recent electoral term t; c corresponds to the cutoff for MV_{it}, which is zero and defines the aligned and unaligned mayor treatment; and $\pm h$ correspond to the upper and lower limits of an automatically derived, optimal close-election bandwidth for MV, which I estimate using Calonico, Cattaneo and Titiunik (2014).

I use local polynomial regression to estimate τ as follows:

$$\begin{aligned}\tau &= \mu_{\text{aligned}} - \mu_{\text{unaligned}} \\ \mu_{\text{aligned}} &= \underset{h_0,\ldots,h_p}{\arg\min} \sum_{i=1}^{n} \mathbb{1}\left(h \geq MV_{it} > c\right) \left(a_{it} - h_0 - h_1\left(MV_{it} - c\right) - h_2\left(MV_{it} - c\right)^2\right) W\left(\frac{MV_{it} - c}{h}\right) \\ \mu_{\text{unaligned}} &= \underset{h_0,\ldots,h_p}{\arg\min} \sum_{i=1}^{n} \mathbb{1}\left(-h \leq MV_{it} < c\right) \left(a_{it} - h_0 - h_1\left(MV_{it} - c\right) - h_2\left(MV_{it} - c\right)^2\right) W\left(\frac{MV_{it} - c}{h}\right)\end{aligned} \tag{2}$$

where W represents a triangular weight;[20] and μ_{aligned} and $\mu_{\text{unaligned}}$ are the weighted least squares estimates for the respective aligned and unaligned intercepts. Following Gelman and Imbens (2019), I use second-order polynomial fits to allow for some potential non-linearity but avoid potential bias-variance trade-offs that can occur with higher-order polynomials. I also cluster the standard errors at the municipality level per Bartalotti and Brummet (2017).

Overall, the RDD results suggest that party alignment does not predict the number of audits that municipalities receive in each electoral term. More specifically, Figure 2c does not have a discontinuous jump in the number of audits received between aligned and unaligned opposition parties. The absence of sorting, as evidenced by the passing of the McCrary

[20] Triangular weights put more emphasis on observations near the as-if randomly assigned cutoff and thus are theoretically better for causal inference purposes. However, weight choice usually does not significantly alter estimates (Cattaneo, Idrobo and Titiunik, 2019, 36).

(2008) density test in Figure 2d, further supports these results.

The above RDD results are useful both from causal and substantive perspectives, but they are not sufficient. The bandwidth restrictions for the RDD entails a focus on subsample of the observations, so I supplement the RDD results with a Poisson count model of the form below, with standard errors clustered by municipality:

$$\log(\lambda_{it}) = \beta_0 + \beta_1 \text{alignment}_{it} + \beta_2 \text{controls}_{it} \tag{3}$$

The controls are now necessary given the lack of random assignment in alignment. As shown in Figure 2b, alignment, past corruption (lagged infractions and log amounts attached to those infractions), and demographics (population) are all similarly poor predictors of how many audits a municipality receives in a given electoral term. Like Honduras, Guatemala also does not have political variation at the department (state-equivalent) level, so it is possible to analyze mayor-president party alignments without including department-level political controls.

2.2.2. Are Honduran Municipal Audits Ethnically Biased?

Another potential bias is ethnic bias, which I examine using data from Honduras, where there is significant ethnic conflict. To obtain data on ethnicity, I compiled data on the percent of indigenous peoples living in each municipality from the 2013 Honduran census. The groups included in my count were the Garifuna, Lenca, MayaChorti, 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 Poisson regression similar to Equation (3), using the indigenous peoples' share of each municipality's total population in 2013 as the main independent variable and the total number of 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

Table 2: Does Ethnicity Predict Audits in Honduras?

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

Standard errors clustered by municipality in parentheses.
Note: Cross-sectional Poisson regression.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(given that the audits took place across multiple terms). In all specifications in Table 2, I find that municipalities with higher shares of indigenous peoples are more likely to receive an audit, indicating potential ethnic bias. 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. It is thus feasible that these municipalities with higher shares of indigenous populations have higher corruption risks, making the supreme audit institution's risk-based auditing decisions justifiable—especially in the context of a limited budget.

2.3. 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 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 audited as percent of the total, etc. For this reason, some very skilled qualitative researchers may be able to uncover relevant biases through ethnographies or qualitative interviews, though it may be difficult to overcome Hawthorne effects or social desirability bias concerns.

2.3.1. Do Opposition Mayors Receive More Stringent Audits in Honduras?

Table 3: 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	0.587	0.497	0.470	-2.632	-3.170
	(0.316)	(0.381)	(0.347)	(0.332)	(1.957)	(2.260)
Population (log)	0.505***	0.607***	0.546***	0.522***	-3.869***	-3.959***
	(0.0961)	(0.131)	(0.111)	(0.103)	(0.647)	(0.743)
Constant		-19.22***	-11.63***	-10.24***	79.08***	79.60***
		(1.662)	(1.437)	(1.334)	(6.620)	(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***	-11.63***	-10.24***	79.08***	79.60***
		(1.662)	(1.437)	(1.334)	(6.620)	(8.312)

Clustered, robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Note: All specifications contain term fixed effects.

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.

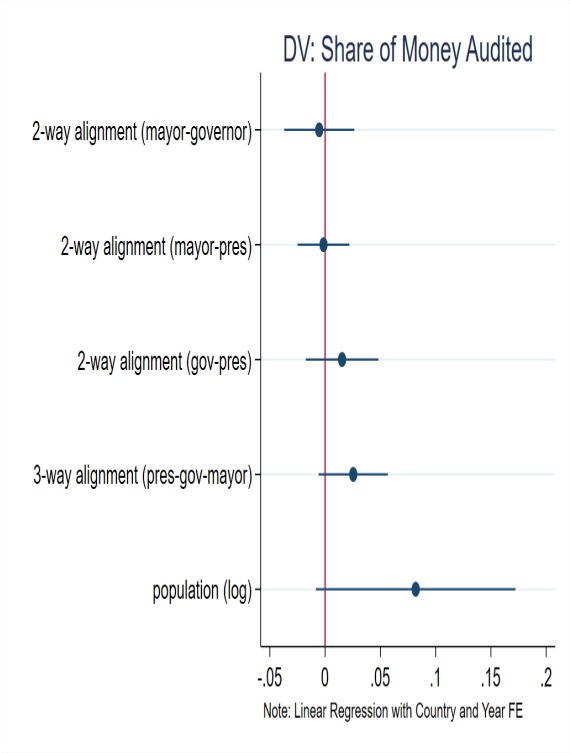
Irrespective of these potential challenges, Table 3 examines potential partisanship-related duration biases in audit intensity through Cox, Exponential, Weibull, Gompertz, Log-Normal, and Log-Logistic regression. I use these various functional forms because it is too difficult to theorize about the correct 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 H.1 provides further technical specifics of the survival analyses.

2.3.2. Do Opposition Mayors Receive More Intense Audits in Mexico?

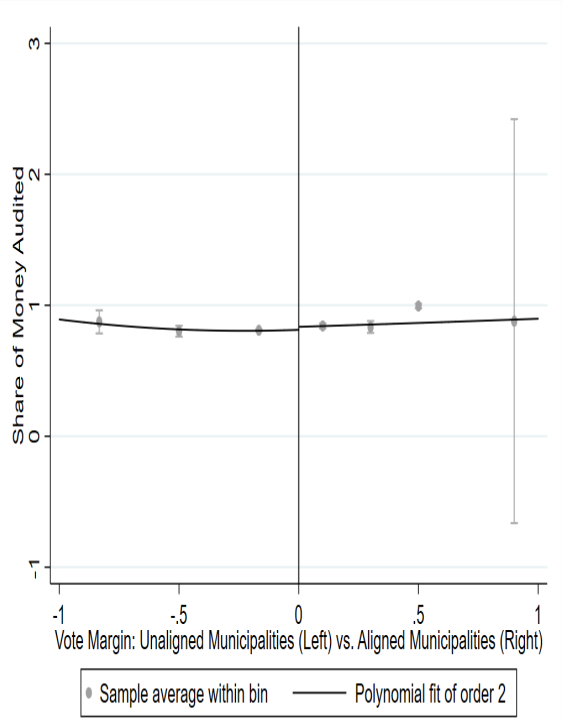
Mexico's data provide provides another method of assessing potential implementation/dosage biases. For each audit, Mexico's ASF provides the amount of money in Mexican pesos in the audit sample as well as the overall amount of money that it could have audited. If membership in an opposition party predicts a higher share of funds audited, it suggests that opposition parties are receiving more stringent audits. As results from Figure 3 suggest, that does not appear to be the case for Mexico. Both the linear regression model and regression discontinuity approach reach the same conclusion.

Figure 3: Mexico's Audit Implementation/Dosage

(a) Panel Linear Regression Model



(b) RDD Results (3-Way Alignment)



3. Discussion and Broader Applicability of the Framework

Table 4: Overview of Results

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

Table 4 provides an overview of the results discussed in the previous section as well as the supplemental analyses in the Appendix. 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. Empirical analyses suggest no political biases, but there do appear to be potential ethnic biases in Honduras—even though interview evidence suggest that the

patterns are likely justifiable given capacity challenges. The framework's third pillar on audit implementation/dosage is perhaps the most difficult to analyze. Nevertheless, the relevant quantitative measures on audit duration in Honduras and share of money audited in Mexico 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 hometown biases (see Chu et al., 2021),[21] 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 will prove particularly useful for governmental audit data from both more democratic and authoritarian countries. I mention authoritarian countries, because single-party authoritarian regimes such as China also conduct subnational audits (e.g., Chu et al., 2021), and the framework's pillars relating to independence and implementation/dosage remain particularly relevant for such countries. Going forward, future research may wish to examine the relevance of the present article's framework for state-owned enterprise (SOE) audits.

4. 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 on political connections,[22] procurement,[23] asset declarations,[24] taxes,[25] and customs duties.[26] However, audit data standout from the above measures for a simple reason: audits reveal a

[21] Even transparency requests to India and Honduras' SAIs did not yield any data on auditor sanctions.
[22] See, for example, Fisman (2001), Faccio (2006), and Fisman and Wang (2015)
[23] See, for example, Bandiera, Prat and Valletti (2009), Mironov and Zhuravskaya (2016), Fazekas (2017), Fazekas, Cingolani and Tóth (2018), Baltrunaite (2020), Broms, Dahlström and Fazekas (2019), and Fazekas and Kocsis (2020).
[24] See, for example, Eggers and Hainmueller (2009) and Fisman, Schulz and Vig (2014).
[25] See, for example, Khan, Khwaja and Olken (2016) and Naritomi (2019).
[26] See, for example, Fisman and Wei (2004, 2009), Sequeira and Djankov (2014), and Rijkers, Baghdadi and Raballand (2017).

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.

The challenge with audit-related corruption measures to date is that, with very few exceptions,[27] scholars have focused almost exclusively on the municipal audit data from Brazil. The numerous Brazil studies using these data as an independent variable 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 question due to the unique nature of Brazil's anti-corruption program, the program's termination in 2015, and the fact that no other country randomly audits all of its subnational units. With respect to the latter, interviews with the head of one of the supreme audit agencies whose data I used in this paper suggests that randomly assigned audits are not of great interest to policymakers. In particular, supreme audit agencies assign audits based on risk-based criteria, which preclude random assignment.

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, rely more on other countries besides Brazil, account for biases in corruption data and, in turn, better understand the diverse causes and consequences of corruption.

[27] See footnote 4.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2005. Institutions as a Fundamental Cause of Long-Run Growth. In *Handbook of Economic Growth*, ed. Philippe Aghion and Steven N Durlauf. Vol. 1A Elsevier B.V. pp. 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.

Arias, Eric, Horacio Larreguy, John Marshall and Pablo Querubín. 2019. When Does Information Increase Electoral Accountability? Lessons from a Field Experiment in Mexico. In *Information, Accountability, and Culmulative Learning: Lessons from Metaketa I*, ed. Thad Dunning, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Criag McIntosh and Gareth Nellis. Cambridge: Cambridge University Press chapter 5, pp. 118–155.

Arias, Eric, Horacio Larreguy, John Marshall and Pablo Querubín. 2022. "Priors Rule: When Do Malfeasance Revelations Help Or Hurt Incumbent Parties?" *Journal of the European Economic Association* 20(4):1433–1477.

Arias, Eric, Pablo Balán, Horacio Larreguy, John Marshall and Pablo Querubín. 2019. "Information Provision, Voter Coordination, and Electoral Accountability: Evidence from Mexican Social Networks." *American Political Science Review* 113(2):475–498.

Auditoría Superior de la Federación de México. 2013. "Política Institucional de Integridad de la ASF.".

Baltrunaite, Audinga. 2020. "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.

Bartalotti, Otávio and Quentin Brummet. 2017. Regression Discontinuity Designs with Clustered Data. In *Advances in Econometrics*, ed. Matias D. Cattaneo and Juan Carlos Escanciano. Vol. 38 Bingley, United Kingdom: Emerald Publishing pp. 383–420.

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.

Berliner, Daniel and Joachim Wehner. 2022. "Audits for Accountability: Evidence from Municipal By-Elections in South Africa." *Journal of Politics* 84(3):1581–1594.

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.

Blossfeld, Hans-Peter, Katrin Golsch and Götz Rohwer. 2007. *Event History Analysis with Stata*. New York: Psychology Press: Taylor and Francis Group.

Bobonis, Gustavo J., Luis R. Cámara Fuertes and Rainer Schwabe. 2016. "Monitoring Corruptible Politicians." *American Economic Review* 106(8):2371–2405.

Bologna, Jamie. 2016. "The Effect of Informal Employment and Corruption on Income Levels in Brazil." *Journal of Comparative Economics* 44(3):657–695.

Bologna, Jamie. 2017. "Contagious Corruption, Informal Employment, and Income: Evidence from Brazilian Municipalities." *Annals of Regional Science* 58(1):67–118.

Bologna, Jamie and Amanda Ross. 2015. "Corruption and Entrepreneurship: Evidence from Brazilian Municipalities." *Public Choice* 165(1-2):59–77.

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.

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.

Cattaneo, Matias D., Nicolas Idrobo and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Part I*. Cambridge: Cambridge University Press.

Cavalcanti, Francisco, Gianmarco Daniele and Sergio Galletta. 2018. "Popularity Shocks and Political Selection." *Journal of Public Economics* 165:201–216.

Chong, Alberto, Ana L. De La O, Dean Karlan and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout." *Journal of Politics* 29(1):55–71.

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.

Colonnelli, Emanuele and Mounu Prem. 2022. "Corruption and Firms." *Review of Economic Studies* 89:695–732.

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.".

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 La O, Ana L., Lucas I. González and Rebecca Weitz-Shapiro. 2023. "Voluntary Audits: Experimental Evidence on a New Approach to Monitoring Front-line Bureaucrats." *World Development* 162(106131).

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.

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.

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. 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.

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.

Funk, Kendall D. and Erica Owen. 2020. "Consequences of an Anti-Corruption Experiment for Local Government Performance in Brazil." *Journal of Policy Analysis and Management* 39(2):444–468.

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.

Gelman, Andrew and Guido W. Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business and Economic Statistics* 37(3):447–456.

Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. W.W. Norton & Company.

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. 2013*a*. "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. 2013*b*. *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.

Gurazada, Srinivas, Zohra Farooq Noorryar, Arun Manuja, Moses Wasika, Mona El-Chami, Michael Jacobs, Zac Mills, Timothy Williamson, Seongjun Kim, Kolajo Joshua Adegbenro, Lara Saade, Jose Carlos Ferreyra, Barbara Joan Rice, Marlee Jenean Beers, Maria Lopez, Luidmila Uvarova and Bertin Lopez. 2021. *Supreme Audit Institutions Independence Index: 2021 Global Synthesis Report*. Washington, DC: World Bank.

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_Indep*

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.

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.

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. 2007*a*. "Growth and Governance: A Defense." *Journal of Politics* 69(2):563–569.

Kurtz, Marcus J. and Andrew Schrank. 2007*b*. "Growth and Governance: Models, Measures, and Mechanisms." *Journal of Politics* 69(2):538–554.

Lagunes, Paul. 2021. *The Eye and Whip: Corruption Control in the Americas*. New York: Oxford University Press.

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.

Magnusson, Leandro M. and Yashar Tarverdi. 2020. "Measuring Governance: Why Do Errors Matter?" *World Development* 136:105061.

Marshall, John. 2023. "Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?" *American Journal of Political Science* .

Mauro, Paolo. 1995. "Corruption and Growth." *Quarterly Journal of Economics* 110(3):681–712.

McCrary, 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. 2022. "Assessing Data Quality: An Approach and An Application." *Political Analysis* 30(3):426–449.

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 Seventeeth-Century England." *Journal of Economic History* 49(4):803–832.

Odilla, Fernanda and Denisse Rodriguez-Olivari. 2021. Corruption Control under Fire: A Brief History of Brazil's Office of the Comptroller General. In *The Politics of Anti-Corruption Agencies in Latin America*, ed. Joseph Pozsgai-Alvarez. London: Routledge chapter 7, pp. 108–133.

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.

Olken, Benjamin A. 2009. "Corruption Perceptions vs. Corruption Reality." *Journal of Public Economics* 93(12-13):950–964.

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.

Thomas, Melissa A. 2010. "What Do the Worldwide Governance Indicators Measure?" *European Journal of Development Research* 22(1):31–54.

Timmons, Jeffrey F. and Francisco Garfias. 2015. "Revealed Corruption, Taxation, and Fiscal Accountability: Evidence from Brazil." *World Development* 70:13–27.

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.

Appendices

Appendix A Measuring Corruption with Observational, Perception-Based Data

Corruption entails "the misuse of public office for private gain",[28] 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,[29] which is especially useful because corruption is multidimensional concept with experience-near and experience-distant meanings.[30]

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.[31] Among its many limitations, Transparency International CPI data are not suitable for over time comparisons, which severely limit scholars

[28] 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).

[29] 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).

[30] "Experience-distant concepts are ones that specialists of one sort or another... employ to forward their scientific, philosophical, or and 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)."

[31] "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).

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,[32] has limited construct validity,[33] and suffers from information leakage,[34] halo effects,[35] and content opacity (Kurtz and Schrank, 2007*a*,*b*; Langbein and Knack, 2010; Thomas, 2010; Bersch and Botero, 2014; Gisselquist, 2014; Hollyer, 2018; Magnusson and Tarverdi, 2020).[36]

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 lines, Kurtz and Schrank (2007*a*,*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

[32] 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.

[33] 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).

[34] Information leakage refers to when a survey taker's perspective on one phenomenon is influenced by something else.

[35] Halo effects specifically refer to when positive effects about something color or influence a person's opinion in a positive way about something else.

[36] Kurtz and Schrank (2007*a*,*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.

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).[37] 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., 2022). 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).

[37]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. (2022) allow for less missing data, incorporation of priors, and are more attuned to measuring intractable, unobservable concepts like corruption (Fariss, Kenwick and Reuning, 2020).

Appendix B Honduran Audits Codebook

B.1 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 B2). 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 B2).

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.[38] 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 B2. Appendix B provides further details about these data and their coding procedure.

B.2 Variables in the Honduran Audits Dataset

[38] 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.

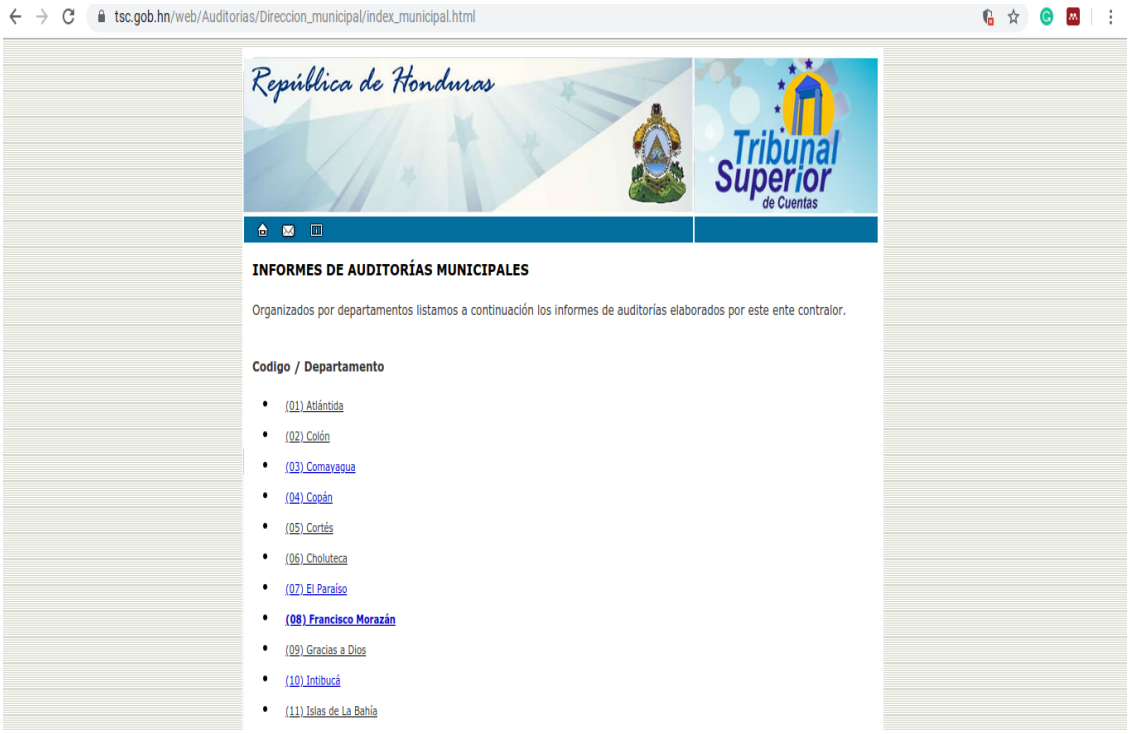
Table B1: Variables in the Dataset

Variable	**Definition**
`unique_number`	Unique identifier for the municipal audit. On occassions 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 palce in specific municipalities.
`municipio`	Municipality where the audit took place.
`departamento`	Department (state/province equivalent) where the audit took place.
`audit_start_date`	Date when the audit started.
`audit_end_date`	Date when the audit ended.
`audit_duration`	Duration in days of the audit.
`total_infractions`	Total number of infractions in the audit.
`corrupt_infractions`	Total number of infractions in the audit.
`severity_infractions`	Severity of infractions in the audit. These qualitative assessments take the form of an ordinal score from 1 (not severe) to 7 (high severity).

B.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".

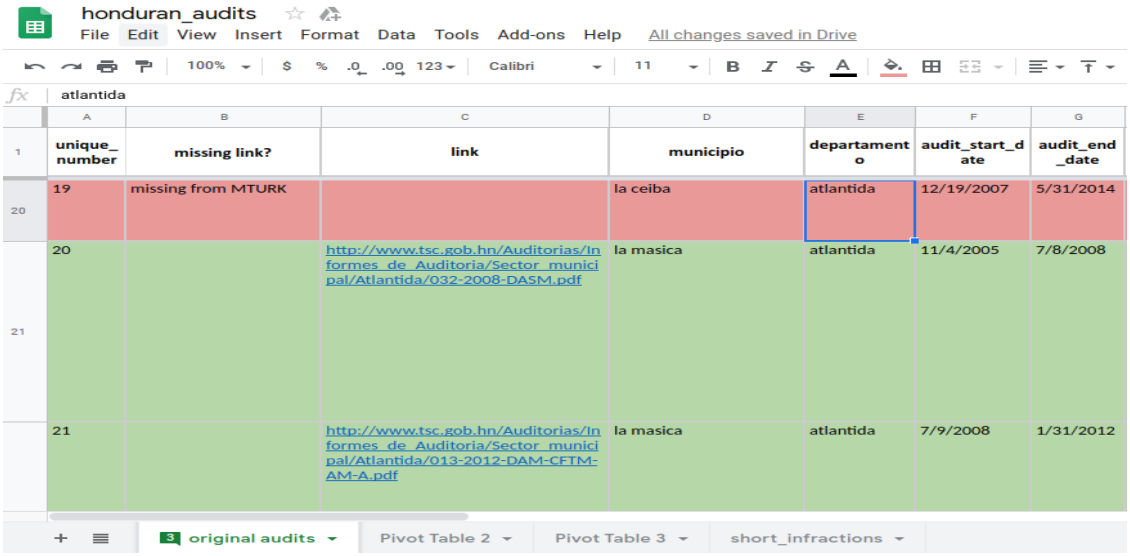
Next, click on "Informes de Auditorías", and select the respective *departamento* (department).



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.

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.



honduran_audits

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

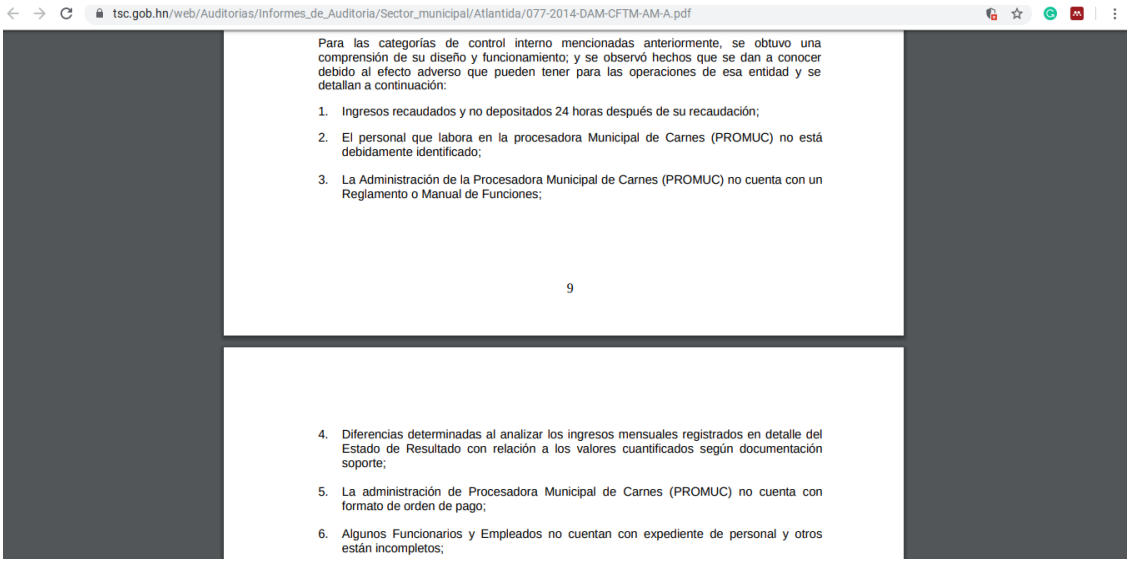
atlantida

	A	B	C	D	E	F	G
1	unique_ number	missing link?	link	municipio	departamento	audit_start_date	audit_end_date
20	19	missing from MTURK		la ceiba	atlantida	12/19/2007	5/31/2014
21	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

original audits | Pivot Table 2 | Pivot Table 3 | short_infractions

B.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.



tsc.gob.hn/web/Auditorias/Informes_de_Auditoria/Sector_municipal/Atlantida/077-2014-DAM-CFTM-AM-A.pdf

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:

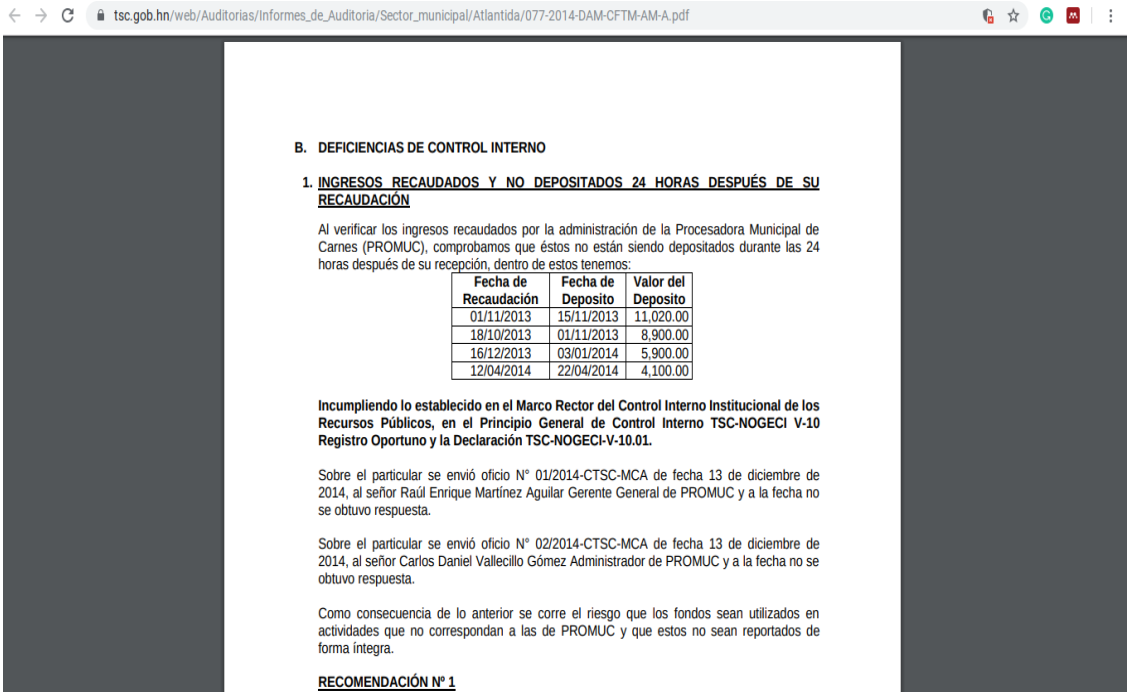
1. Ingresos recaudados y no depositados 24 horas después de su recaudación;
2. El personal que labora en la procesadora Municipal de Carnes (PROMUC) no está debidamente identificado;
3. La Administración de la Procesadora Municipal de Carnes (PROMUC) no cuenta con un Reglamento o Manual de Funciones;

9

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:



tsc.gob.hn/web/Auditorias/Informes_de_Auditoria/Sector_municipal/Atlantida/077-2014-DAM-CFTM-AM-A.pdf

B. DEFICIENCIAS DE CONTROL INTERNO

1. **INGRESOS RECAUDADOS Y NO DEPOSITADOS 24 HORAS DESPUÉS DE SU RECAUDACIÓN**

Al verificar los ingresos recaudados por la administración de la Procesadora Municipal de Carnes (PROMUC), comprobamos que éstos no están siendo depositados durante las 24 horas después de su recepción, dentro de estos tenemos:

Fecha de Recaudación	Fecha de Deposito	Valor del Deposito
01/11/2013	15/11/2013	11,020.00
18/10/2013	01/11/2013	8,900.00
16/12/2013	03/01/2014	5,900.00
12/04/2014	22/04/2014	4,100.00

Incumpliendo lo establecido en el Marco Rector del Control Interno Institucional de los Recursos Públicos, en el Principio General de Control Interno TSC-NOGECI V-10 Registro Oportuno y la Declaración TSC-NOGECI-V-10.01.

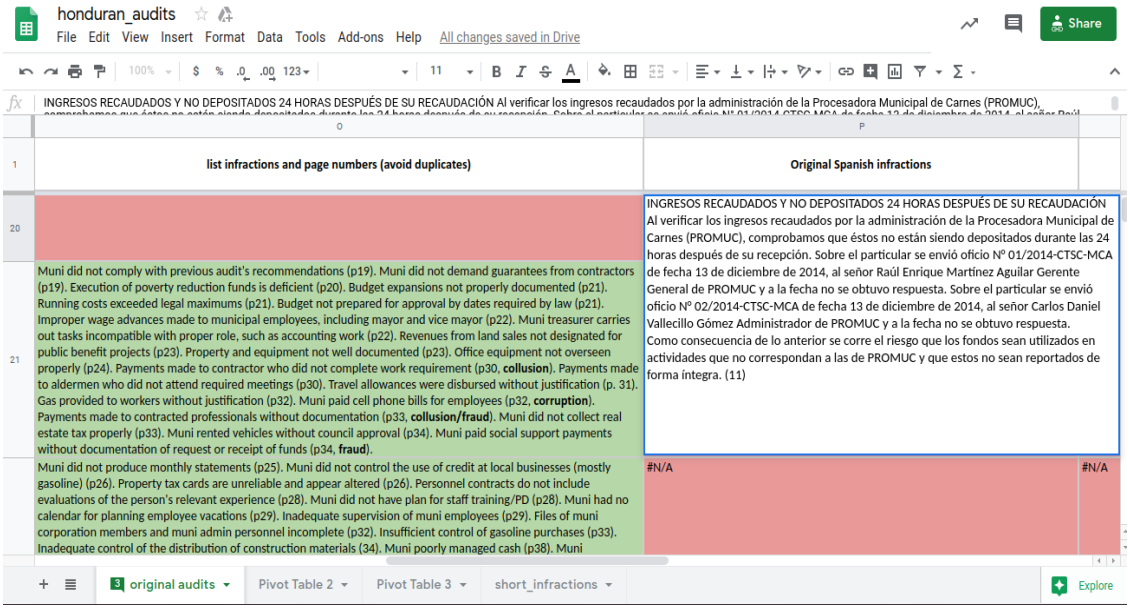
Sobre el particular se envió oficio N° 01/2014-CTSC-MCA de fecha 13 de diciembre de 2014, al señor Raúl Enrique Martínez Aguilar Gerente General de PROMUC y a la fecha no se obtuvo respuesta.

Sobre el particular se envió oficio N° 02/2014-CTSC-MCA de fecha 13 de diciembre de 2014, al señor Carlos Daniel Vallecillo Gómez Administrador de PROMUC y a la fecha no se obtuvo respuesta.

Como consecuencia de lo anterior se corre el riesgo que los fondos sean utilizados en actividades que no correspondan a las de PROMUC y que estos no sean reportados de forma íntegra.

RECOMENDACIÓN Nº 1

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



honduran_audits

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

Share

INGRESOS RECAUDADOS Y NO DEPOSITADOS 24 HORAS DESPUÉS DE SU RECAUDACIÓN Al verificar los ingresos recaudados por la administración de la Procesadora Municipal de Carnes (PROMUC),

	list infractions and page numbers (avoid duplicates)	Original Spanish infractions	
20		INGRESOS RECAUDADOS Y NO DEPOSITADOS 24 HORAS DESPUÉS DE SU RECAUDACIÓN Al verificar los ingresos recaudados por la administración de la Procesadora Municipal de Carnes (PROMUC), comprobamos que éstos no están siendo depositados durante las 24 horas después de su recepción. Sobre el particular se envió oficio Nº 01/2014-CTSC-MCA de fecha 13 de diciembre de 2014, al señor Raúl Enrique Martínez Aguilar Gerente General de PROMUC y a la fecha no se obtuvo respuesta. Sobre el particular se envió oficio Nº 02/2014-CTSC-MCA de fecha 13 de diciembre de 2014, al señor Carlos Daniel Vallecillo Gómez Administrador de PROMUC y a la fecha no se obtuvo respuesta. Como consecuencia de lo anterior se corre el riesgo que los fondos sean utilizados en actividades que no correspondan a las de PROMUC y que estos no sean reportados de forma íntegra. (11)	
21	Muni did not comply with previous audit's recommendations (p19). Muni did not demand guarantees from contractors (p19). Execution of poverty reduction funds is deficient (p20). Budget expansions not properly documented (p21). Running costs exceeded legal maximums (p21). Budget not prepared for approval by dates required by law (p21). Improper wage advances made to municipal employees, including mayor and vice mayor (p22). Muni treasurer carries out tasks incompatible with proper role, such as accounting work (p22). Revenues from land sales not designated for public benefit projects (p23). Property and equipment not well documented (p23). Office equipment not overseen properly (p24). Payments made to contractor who did not complete work requirement (p30, **collusion**). Payments made to aldermen who did not attend required meetings (p30). Travel allowances were disbursed without justification (p. 31). Gas provided to workers without justification (p32). Muni paid cell phone bills for employees (p32, **corruption**). Payments made to contracted professionals without documentation (p33, **collusion/fraud**). Muni did not collect real estate tax properly (p33). Muni rented vehicles without council approval (p34). Muni paid social support payments without documentation of request or receipt of funds (p34, **fraud**).		
	Muni did not produce monthly statements (p25). Muni did not control the use of credit at local businesses (mostly gasoline) (p26). Property tax cards are unreliable and appear altered (p26). Personnel contracts do not include evaluations of the person's relevant experience (p28). Muni did not have plan for staff training/PD (p28). Muni had no calendar for planning employee vacations (p29). Inadequate supervision of muni employees (p29). Files of muni corporation members and muni admin personnel incomplete (p32). Insufficient control of gasoline purchases (p33). Inadequate control of the distribution of construction materials (34). Muni poorly managed cash (p38). Muni	#N/A	#N/A

original audits | Pivot Table 2 | Pivot Table 3 | short_infractions | Explore

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 B2. 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	Montlhy 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 B2: 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 C Indian Audits Codebook

C.1 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 the Codebook below 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 C). 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.

C.2 Dataset Overview and Scope

This section provides an overview of the dataset. 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 2021.

The CAG conducts its annual state-level audits by sector. Accordingly, the primary the 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 C.3 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.

C.3 Variables in the Indian Audit Datasets

Table C1: 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 C2: 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 C3: Variables in the Receipts 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 C4: Variables in the Subsector Dataset

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

C.4 Coding Procedure

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

C.4.1 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 observatins	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.

(₹ in crore)

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
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**

C.4.2 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.

(₹ in crore)

Year of Inspection Report	No. of units audited	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**

C.4.3 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**

(Rupees in crore)

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			Position of IRs in respect of which first reply not received from March 1992 to March 2006		
	No. of IRs	No. of para-graphs	Money value	No. of IRs	No. of para-graphs	Money value	No. of IRs	No. of para-graphs	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

C.4.4 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 lahks 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**

C.4.5 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	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 subsector coding.

C.4.6 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		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

(₹ in lakh)

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
Mandya	-	-	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

C.4.7 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**

C.4.8 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 tablse below would be coded in the department dataset. They are relevant

because the shows paragraphs/infractions 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

(₹ in lakh)

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**

(₹ in crore)

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

C.4.9 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.	445	1,131	111.75
		Entertainment	10	15	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

C.4.10 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

(₹ in crore)

Sl. No.	Categories	Number of cases	Amount
1	Non-settlement of vested lands	16	0.10
2	Non-settlement of *sairats*	9	0.02
3	Other cases	153	3.77
Total		**178**	**3.89**

C.4.11 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.

f_the_Union_Territory_of_Puducherry.pdf 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)	Number of cases	Amount (₹ in lakh)
1	Adi-Dravidar Welfare	3	3.26	1	0.43	4	3.68
2	Agriculture	6	1.89	2	8.15	8	10.04
3	Animal Husbandry	3	0.06	1	41.57	3	41.63
4	Art and Culture	3	0.37	Nil	Nil	3	0.36
5	Education	18	1.70	3	43.81	21	45.51
6	Electricity	265	4,380.73	1	0.17	266	4,380.91
7	Health and Family Welfare	13	1.37	Nil	Nil	13	1.37
8	Information Technology	1	1.15	Nil	Nil	1	1.15
9	Labour	2	0.70	Nil	Nil	2	0.70
10	Local Administration	5	2.03	1	0.47	6	2.51
11	Police	Nil	Nil	2	0.50	2	0.50
12	Public Works	2	0.20	2	5.01	4	5.20
13	Revenue	2	0.84	Nil	Nil	1	0.84
14	Tourism	1	*	Nil	Nil	1	*
15	Women and Child Development	1	3.75	Nil	Nil	1	3.75
	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	4	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

(₹ in crore)

Year	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 D Guatemalan Audits Codebook

D.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.

D.1.1 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 D provides a full codebook of these data.

D.2 Variables in the Dataset

D.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.

Table D1: Variables in the Guatemalan Audits Dataset

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

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 (i.e., state equivalent).

DEPARTAMENTO DE IZABAL

		Denuncias		Sanciones		Formulaciones	
No.	Departamento	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 E Mexican Audits Codebook

E.1 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*).[39] 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 **??**). The ASF reports to the Mexican Chamber of Deputies.

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 E2 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. It thus adds significant corruption data for researchers to exploit.

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 E3 and E4). Appendix E provides a full codebook of the Mexican audits dataset.

Table E2 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

[39] EFSEs are also called *las Entidades de Fiscalización Superior Locales* (EFSLs).

Table E2: 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

(see Tables E3 and E4).

Table E3: 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 E4: 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

E.2 Variables in the Dataset

Table E5: Variables in The Dataset

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

Continued on next page

Table E5: Variables in the Dataset – *continued*

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

Continued on next page

Table E5: Variables in the Dataset – *continued*

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

Appendix F Qualitative Independence Analyses

F.1 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 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).[40] 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 TSC also 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 needs to answer to the TSC's Probity 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

[40] 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).

Cuentas de Honduras, 2018).

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 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.

F.2 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).

F.3 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). 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). The one more challenging area pertains to the budget, which prevents the ASF from conducting more audits (OECD, 2017).

F.4 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.

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.

Appendix G Additional Political Biases Analyses

G.1 Frequency, Selection, and Distribution of Audits (Mexico)

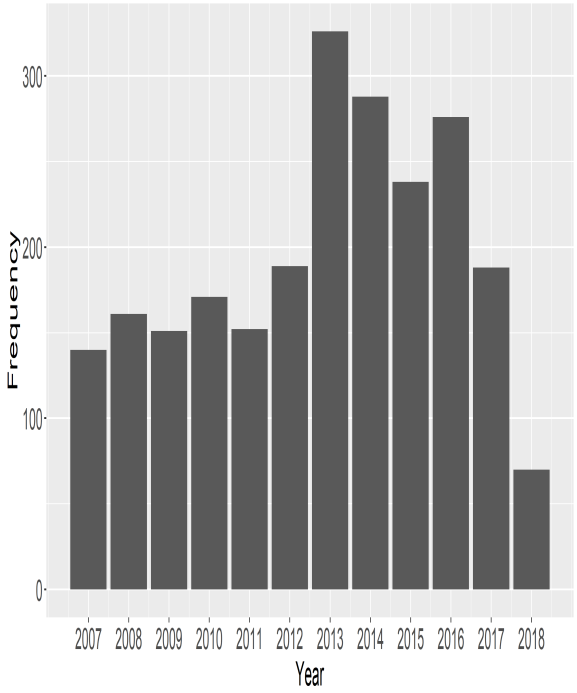
Mexico's ASF does not audit all of Mexico's 2,454 municipalities each year (see Figure 1c). Instead, the ASF conducts risk-based audits, which are decided upon by the ASF planning commission in consultation with the Chambers of Deputies. According to an analysis by the OECD (2017), the decision process 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 2016,[41] some of 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.

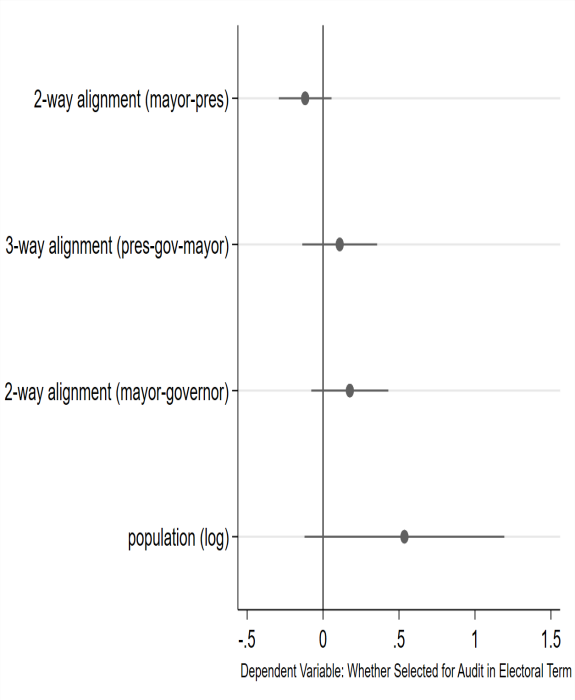
[41] For more on the relevant electoral reform, see Motolinia (2021).

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

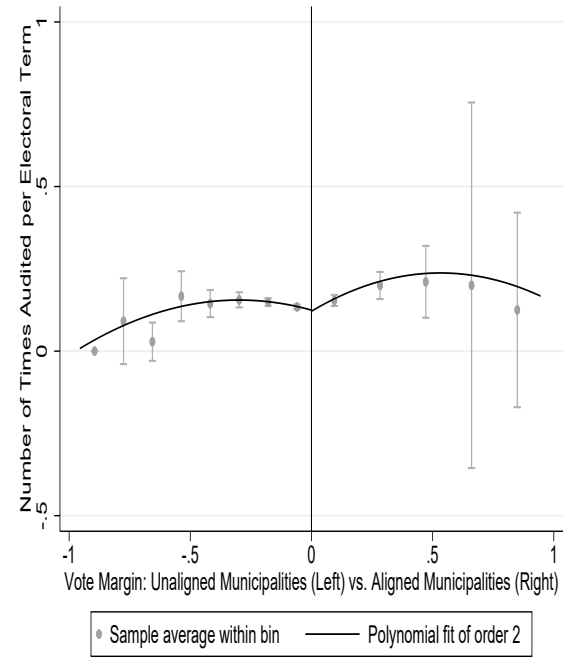
(a) Municipalities Receiving Audits by Year



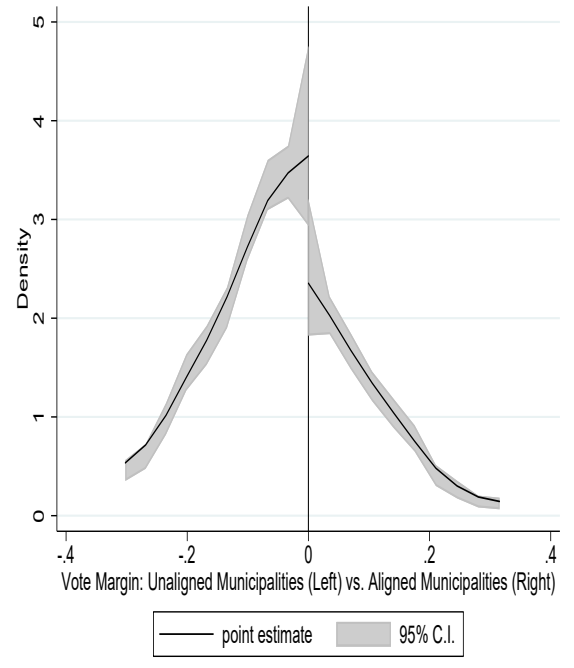
(b) Logit Model with Municipality and Year FE



(c) RD Results (Year)

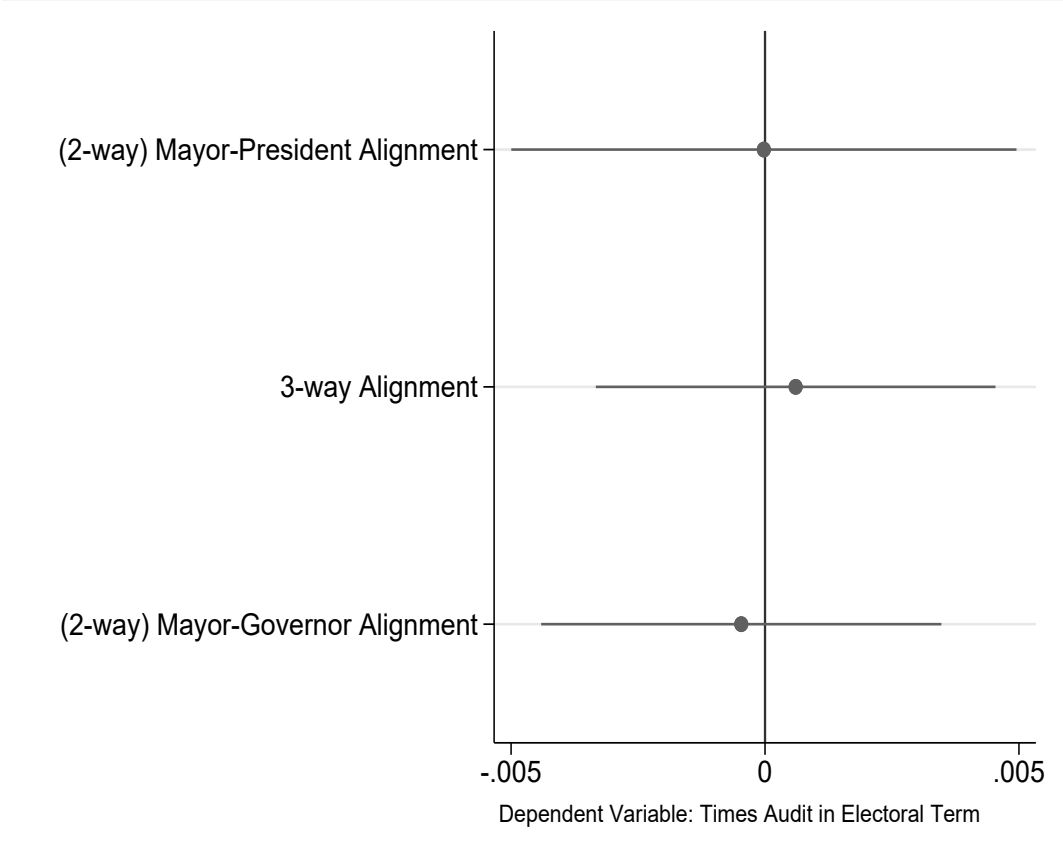


(d) McCrary Density Plot for RD



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.[42] 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 5, none of the alignment configurations predict greater or lower numbers of audits.

Figure 5: 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.

[42] 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.

In the corresponding regression discontinuity design model using three-way party alignment in Figure 4c,[43] 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.

G.2 Frequency, Selection, and Distribution of Audits (Honduras)

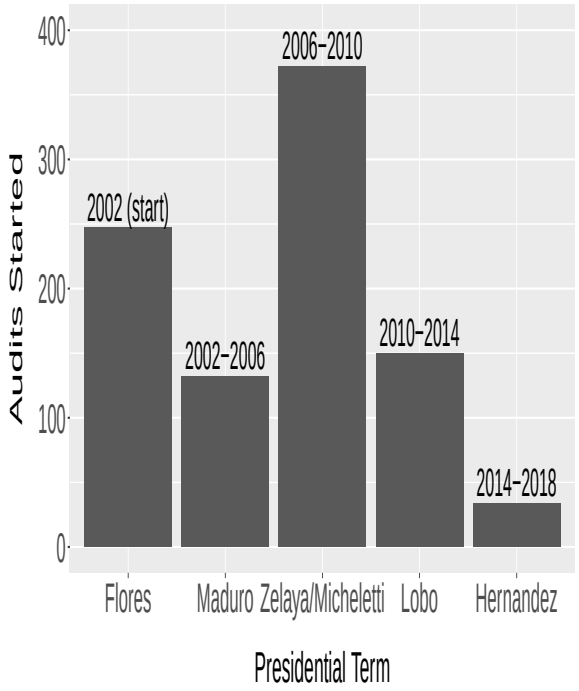
The TSC performs numerous audits of both municipalities and state-owned enterprises each year. However, as compared to the SAIs from other countries examined in this paper, the TSC undertakes relatively few municipal audits. Honduras has 298 municipalities but only in 2002—when the TSC became a formal SAI—did the country start more than 200 audits in a year (see Figure 6a). 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.

In terms of the relationship between political rivalry and auditing decisions, the logit model in Figure 6b 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.

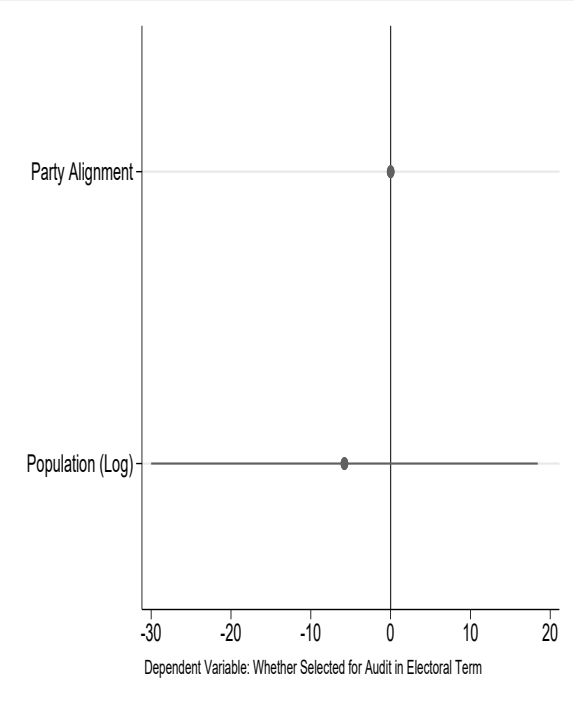
[43] Note: results are similar for the other alignment configurations.

Figure 6: Main Results: Honduras' Municipal Audit Distribution

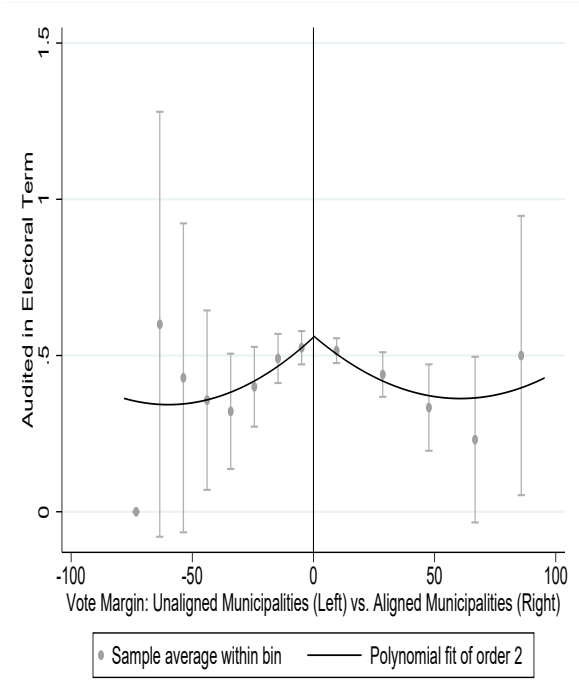
(a) Audits Started by the TSC Over Time



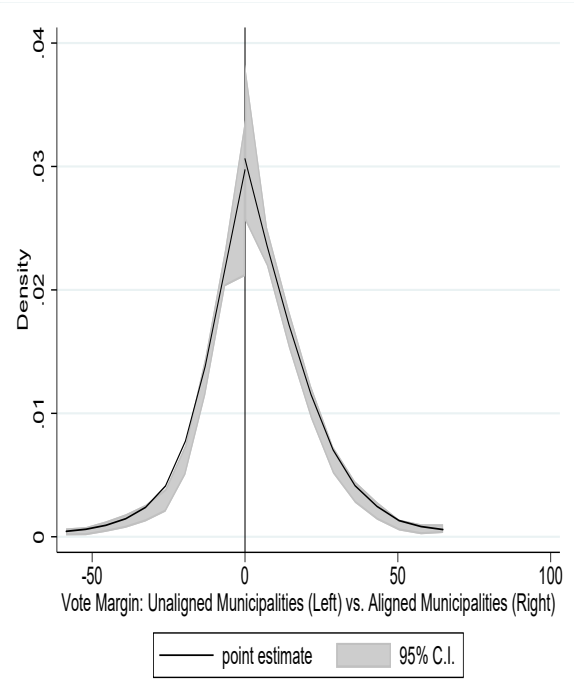
(b) Logit Model with Municipality and Term FE



(c) RD Results (Term)



(d) McCrary Density Plot for RDD



Consistent with the framework detailed in Section 2, I run a regression discontinuity design along the lines of Brollo and Nannicini (2012) as well. It uses random variation 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 6b and Figure 6c: party alignment does not affect auditing decisions. As shown in Figure 6d, 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.

G.3 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.

Appendix H Additional Implementation/Dosage Analyses

H.1 Honduras

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).

H.1.1 Data for Survival Analysis

Before conducting the actual survival analysis, it was necessary to transform data described in Section B.1 and Appendix B 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 H6 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.

Table H6: 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 Michelettti 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.

H.1.2 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 is the multi-episode semi-parametric Cox proportional hazards model:

$$L^p = \prod_{k \in D} \prod_{i \in \epsilon_k} \frac{exp\big(A^{(k)}(t_i)\alpha^{(k)}\big)}{\sum l \in R(t_i) exp\big(A^{(k)}(t_i)\alpha^{(k)}\big)} \tag{4}$$

"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.

H.2 Guatemala

Guatemala's CGC does not provide quantitative data for assessing audit implementation/dosage.

H.3 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.