Measuring Corruption Using Governmental Audits: Strategies and Pitfalls*

Michael Denly†

July 10, 2020

For the most recent version of this paper, please visit:
www.mikedenly.com/files/corruption-measurement.pdf

Abstract

For about 25 years, scholars have primarily relied on perceptions-based measures of governmental corruption, even though the drawbacks of these measures are ample and well-known. More recently, analyses centered on Brazil have showcased the utility of randomized audits as a more objective alternative to perception-based measures (e.g., Ferraz and Finan, 2008). However, Brazil is the only country in the world with randomized audits. In this paper, I show how scholars can objectively estimate governmental corruption with audits even when they are not randomized. Specifically, I show that it is acceptable to use non-randomized audits to measure corruption when: 1) the distribution of audits is not biased against opposition party politicians, especially following close elections; 2) scholars appropriately take into account the structure of government in each country; and 3) the legal framework underpinning supreme audit institutions' legal framework prevent them from being patronage bodies of the executive. I demonstrate the utility of my approach by analyzing subnational audits from India, Mexico, Honduras, and Guatemala, each of which figure (slightly) differently on the above criteria. The new data and approach proposed in this paper will help researchers undertake more objective analyses of governmental corruption around the world.

*For extraordinary research assistance, I thank 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, Amila Lulo, Ana Suarez, Ana Ramirez, Awab Ahmed, Blaine Finstein, Brayan Salgado, Caleb Rudow, Caroline Hastings, Caroline Woodman, Connor Hanna, David Tibaquira, Dorna Abdi, Elena Solimano, Emily Ibarra, Erin Eggleston, Ethan Tenison, Evelin Caro Gutierrez, Fabiola Ramirez, Felipa Mendez, Francesca D'Annunzio, Isabel Campos, Iulia Tothezan, Ishani Pandya, Ivana Jelensky, Jenny Rodriguez, Joelean Hall, Jonah Isaac, Josh Meuth Alldredge, Jude Rios, Judy Lane, Kais Bhandari, Karan Kanatala, Kate Hopp, Katherine Alonzo, Lissette Almanza, Luca Venegoni, Luisa Venegoni, Magdalena Ibarra, Marifer Guerrero, Mary White, Mishan Kara, Mohamed Abulfalgha, Mobin Piracha, Nicole Domingo, Nicky Pownall, Paige Johnson, Rachel Boles, Rachel Rosenburg, Ryan Jabra, Saamia Imtiaz, Samuel Gilette, Sara Lowe, Sterling Mosley, Tara Jackson, Temi Ajibola, Vanessa Gonzales, and Vanessa Lizcano. For advice or feedback, I thank Akshat Gautam, Arun Manuja, Jan Pierskalla, Mike Findley, John Gerring, Ken Greene, Xiaobo Lu, and Serge Nkuindja. All errors are those of the author.

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

Perhaps the original sin of empirical corruption research is that the majority of what scholars "know" about corruption comes from perceptions-based survey data, not objective measures of corruption.[1] It is an "original sin" because the first empirical studies on corruption acknowledged the imperfect nature of the perceptions data (e.g., Mauro, 1995, 1998; Treisman, 2000). Later, scholars showed that halo effects and information leakage generally make perceptions data unsuitable for advancing new theories and knowledge (Kurtz and Schrank, 2007*a*,*b*; Hollyer, 2018). Nevertheless, the findings of some of the most prominent recent theoretical advances on corruption continue to be based mainly off perceptions data (e.g., Mungiu-Pippidi, 2015; Schwindt-Bayer and Tavits, 2016; Potrafke, 2019).

Over the years, improvements to the objective measurement of corruption have taken two forms: experimental and observational. Experimental studies have provided many objectively-measured findings that have reshaped knowledge about corruption (e.g., Olken, 2007; Gingerich, 2013*b*; Findley, Nielson and Sharman, 2014), but such studies tend to be expensive and difficult-to-replicate. On the observational side, scholars have contributed useful improvements to the perceptions indexes and to accurately measuring corruption in surveys more broadly (e.g., Standaert, 2015; Kraay and Murrell, 2016; Gingerich et al., 2016). Ostensibly, though, studies using objective procurement data (e.g., Fazekas, Cingolani and Tóth, 2018), asset declarations (e.g., Klašnja, 2015), taxes (e.g., Khan, Khwaja and Olken, 2016), and audits (e.g., Ferraz and Finan, 2008) have produced the most notable innovations.

While the objective study of corruption through procurement data, taxes, and asset declarations has made significant progress, corruption studies using audit data have mostly not ventured beyond Brazil. To be sure, the numerous studies using Brazilian audit data showcase useful causal findings,[2] notably because the country has randomized municipal audits. However, the external validity of the Brazil studies remains a concern due to the

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

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

country's unique institutional setup and the fact that Brazil is the only country in the world with randomized audits. Fortunately, many other countries disseminate subnational audit data, and these data can help with knowledge development at the right level of aggregation (Gingerich, 2013*a*). However, scholars need a framework for discerning the validity of these data.

In this article, I demonstrate how scholars can profitably use subnational audit data to objectively measure corruption and proffer causal findings even when the audits are not randomized. To do so, scholars must examine three criteria. First, because different government structures lead to diverse power configurations, scholars must take the relevant subnational government structure into account. Second, the distribution of audits at the appropriate level of analysis must not be biased against opposition party politicians, especially following close elections. Third, supreme audit institutions' (or regional audit bodies') reporting requirements and hiring procedures must protect them from being patronage bodies of the executive.

I demonstrate the utility of my approach by analyzing subnational audit data from India, Mexico, Honduras, and Guatemala. Each of which figure (slightly) differently on the above criteria. The new approach proposed in this paper will help researchers undertake more objective analyses of governmental corruption around the world.

The paper proceeds as follows. In Section X, I lay out the challenges to measuring corruption and explain why the benefits and drawbacks of perceptions-based indicators. Section XX describes why objective

1. Measuring Corruption with Observational, Perception-Based Data

Corruption entails "the misuse of public office for private gain",[3] and many analysts define the phenomenon to be much broader (e.g., Gingerich, 2013*b*; Mungiu-Pippidi, 2015). Irrespective of how one defines corruption, though, one thing is certain: by its very nature, corruption is a clandestine activity, so it is very difficult to measure. Another certainty is that corruption is one of the most endogenous phenomena in all of social science. For example, numerous studies suggest that corruption deteriorates the quality of democracy,[4] but others suggest that having political regimes that are less democratic fuel more corruption.[5] In sum, corruption is a difficult-to-measure phenomenon for which it is methodologically challenging to make causal claims.

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,[6] which is especially useful because corruption is multidimensional concept with experience-near and experience-distant meanings.[7]

Although development of the perceptions indexes represented major steps forward in

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

[4] See, for example, Lederman, Loayza and Soares (2005), Treisman (2000, 2007), and Lagunes (2012).

[5] See, for example, Jetter, Montoya Agudelo and Ramírez Hassan (2015), Kolstad and Wiig (2016), and Musila (2013).

[6] 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).

[7] "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)."

terms of measuring corruption, they never were suitable 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.[8] Among its many limitations, Transparency International CPI data are not suitable for over time comparisons, which severely scholars ability to use them for theory (Andersson and Heywood, 2009, 758). The World Bank's Control of Corruption score is perhaps the the most sophisticated of the three measures and is suitable for over time comparisons. Still, the measure has limited construct validity and discriminant validity,[9] as well as suffers from information leakage, halo effects, and content opacity (Kurtz and Schrank, 2007*a*,*b*; Langbein and Knack, 2010; Thomas, 2010; Bersch and Botero, 2014; Gisselquist, 2014; Hollyer, 2018).[10]

Information leakage, halo effects, content opacity, and low construct validity are very problematic from the perspective of theory and knowledge-building (Hollyer, 2018). All of these deficits are related, too. 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). Additionally, as Kurtz and Schrank (2007*a*,*b*) show, countries' growth trajectories cloud survey respondents' answers on governance-related questions. Such problems compound as perceptions indexes include more sources, too, which makes the precise scope of perception-based measures necessarily opaque. With opaque measures, it is impossible to precisely verify the existence of construct validity: that is, that concepts measure what they are supposed to measure (Trochim, 2006; Thomas, 2010). Which begs the question: how can a measure without construct validity be useful for theorizing? In short, its ability to do so is very limited, especially when investigating proximate and

[8] "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).

[9] Construct validity concerns whether the concepts measures what it is supposed to measure (Trochim, 2006). Discriminant validity concerns whether measures are not associated with measures that they are not supposed to be associated with. Discriminant validity is generally considered to be part of construct validity. For more, see Thomas (2010).

[10] 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.

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).[11] The most sophisticated of these measures is the Varieties of Democracy (V-Dem) Project's hybrid approach (Marquardt and Pemstein, 2018; Coppedge et al., 2020). V-Dem uses both fact-based and perceptions data, and complements them with targeted measures to improve reliability and validity. Nevertheless, any index relying on perceptions-based measures cannot fully overcome the aforementioned mentioned trade-offs, because they still exist—albeit to a smaller degree (Hollyer, 2018, 118, 128).

2. Audit Findings as Objective Measures of Corruption, and the External Validity of the Brazil Studies

Although objective measures of corruption do not capture as much conceptual intention as the perception-based ones, objective measures generally do not exhibit the same challenges as perception-based indicators. Consequently, objective measures generally have greater construct validity, and are more appropriate for theorizing and knowledge creation. Perhaps equally as significant, objective measures can solve the level-of-analysis problem that plagues most country-level perceptions data: that is, the fact that within-country variation often overwhelms national-level variation (Gingerich, 2013*a*, 505, 538). Aside from equivalence considerations,[12] the use of subnational data also tends to facilitate natural experiments and causal inferences—as opposed to correlational analyses with regression (Stanig, 2018, 150-151).

On the subject of innovative corruption measurement techniques that use natural ex-

[11] 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 Coppedge et al. (2020) allow for less missing data, incorporation of prior data, and are more attuned to measuring intractable, unobservable concepts like corruption (Fariss, Kenwick and Reuning, 2020).

[12] For more on equivalence, see Locke and Thelen (1995) and Stegmueller (2011).

periments at the right level of aggregation, Ferraz and Finan's (2008) study of randomized municipal audits in Brazil pioneered an entire research agenda.[13] From the perspective of measurement, these objective data on corrupt acts like procurement fraud or over-invoicing from Brazil measure true corruption, not preceptions of it (Ferraz and Finan, 2008, 710). Also, due to the fact that the audits are randomized, they obviate the many endogeneity concerns that corruption typically engenders (Escaleras, Anbarci and Register, 2007). In turn, from the perspective of knowledge creation, these data enabled Ferraz and Finan (2008) to conclusively show—for the first time—that information about politicians' corruption levels can hurt their re-election prospects. Other important results stemming from the Brazilian audit infractions data include that: re-election incentives condition politicians' corruption levels (Ferraz and Finan, 2011); getting audited nudges subnational government entities to reduce their future corruption levels—although effects are sectorally heterogeneous (Avis, Ferraz and Finan, 2018; Zamboni and Litschig, 2018); malfeasance revelations affect the types of candidates that parties put forth on party lists (Cavalcanti, Daniele and Galletta, 2018); and female politicians are less corrupt than their male counterparts (Brollo and Troiano, 2016).

The above studies are all very precisely estimated with objective data, but how much external validity do they have? In other words, how certain can scholars be that the results apply equally to other settings, time periods, units (e.g., countries), mechanisms, as well as different operationalizations of the treatment and outcomes (Findley, Kikuta and Denly, 2021)? For practical, substantive, and methodological considerations, the corruption literature needs to move beyond Brazil.

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

2.1. Substantive and Practical External Validity Concerns with the Brazil Studies

Brazil is the only country in the world with randomized subnational audits. In turn, that means Brazil is the only country where scholars can be certain that partisan considerations do not influence who gets audited and at what time. Unequivocally, that is very useful from the perspective of causal inference. By the same token, even countries with Weberian bureaucracies like Sweden are subject to partisan pressures (Dahlström and Holmgren, 2019). Accordingly, understanding the extent of impartiality in the distribution of subnational audits is part and parcel of answering the question of whether audits deter political corruption.[14]

Practically, most countries' supreme audit institutions (SAIs) need to use their own discretion to choose where to audit through risk-based audits. High-corruption areas, or ones for which SAIs receive whistleblower complaints, simply need more attention than low corruption areas. Similarly, it is practically unreasonable for large, populous cities with large budgets to always have the same probability of receiving an audit as smaller villages with small budgets. Perfect randomization is not always fair or useful without some sort of blocking strategy,[15] which is something that the Brazilian government has not implemented.

At least three additional substantive features make the external validity of the Brazil studies limited. First, the country's media facilitates collective action by exposing corrupt politicians.[16] Indeed, both Ferraz and Finan (2008) and Cavalcanti, Daniele and Galletta (2018) show that their results are stronger when there is greater local media presence, which is something that not all relatively corrupt countries enjoy. Second, the country has compulsory voting that contributes to relatively large voter turnout and more engagement in

[14] Olken (2007), Duflo, Hanna and Ryan (2012), and Alm, Martinez-Vazquez and McClellan (2016) show that audits work in other contexts, but Dizon-Ross, Dupas and Robinson (2017) and Dhaliwal and Hanna (2017) have also shown that audits demotivate qualified staff, resulting in many qualified staff to leave their positions.

[15] See Gerber and Green (2012) for more on blocking.

[16] For more on how the media plays a crucial role in exposing corruption and fostering electoral accountability, see Besley and Burgess (2002), Snyder and Strömberg (2010), and Larreguy, Marshall and Snyder (2020).

politics than in many countries. Third, at least since 1988, the country has a "web" of institutions (i.e., Comptroller General, Federal Audit Court, Regional Audit Courts, Public Ministry, Federal Police) that promote horizontal accountability in anti-corruption (Ferraz and Finan, 2018, 255).[17] Clearly, not all countries with challenges controlling corruption have such a robust institutionalal presence. That is especially significant because there is scholarly consensus that institutions are the fundamental cause of development, growth, and corruption (Acemoglu and Robinson, 2012; Fisman and Golden, 2017, 18).

2.2. Methodological Challenges with the External Validity of the Brazil Studies

Methodologically, there are concerns with over-relying on one country as well, particularly one with *sui generis* features. To understand why, consider the simple difference-in-means estimator, $\hat{\delta}_S$. Assuming a binary treatment, Findley, Kikuta and Denly (2021) show that it is possible to decompose $\hat{\delta}_S$ using the potential outcomes framework as follows:

$$\hat{\delta}_S = \delta_P + b_{S1} + b_{S2} + b_P + b_V. \tag{1}$$

Above, δ_P corresponds to the Population Average Treatment Effect (PATE), which is the ideal quantity of theoretical interest;[18] b_{S1} refers to the selection bias;[19] b_{S2} is the difference in the strength of the treatment effects between the treatment and control groups—i.e., within-sample effect heterogeneity; b_P is the bias due to non-random sample selection; and b_V refers to variable heterogeneity bias—i.e., in treatments, outcomes, and mechanisms.

The only bias that randomization in the assignment of Brazil municipal audits data solves is b_{S1}. The unique institutional features described in Section 2.1 enter into Equation

[17] Horizontal accountability refers to the ability of the bureaucracy to exert checks and balances on itself (O'Donnell, 1998).

[18] Not all (quantitative) studies will aim to estimate the PATE. However, given that the ultimate goal of social science is to make inferences beyond the data at hand (King, Keohane and Verba, 1994, 8,34), having an estimate correspond to the population average treatment effect (PATE) is an implicit goal in social science.

[19] Selection bias corresponds to the bias due to lack of random assignment (Angrist and Pischke, 2008).

(1) through b_P and b_V—i.e., what Findley, Kikuta and Denly (2021) "external validity bias". Under many circumstances, it is feasible that the external validity bias can overwhelm the internal validity bias, represented by b_{S1} and b_{S2}. In short, while the Brazil studies represent excellent contributions to the corruption literature, scholars still need more evidence from other countries, time periods, settings with potentially different treatment, outcomes, and mechanisms. Then, scholars will be able to better discern the external validity of the Brazil studies, as well as how politics conditions the results.

3. Data and a Framework for Assessing the Utility of Non-Randomized Audit Data

There is a dearth of political corruption studies employing audit data outside of Brazil.[20] One reason likely pertains to the lack of easily-accessible audit data. A second reason is that scholars do not have systematized ways of discerning the validity of non-randomized audit data.

In this article, I attempt to address both of the above shortcomings. First, I proffer a wealth of subnational audits data from Honduras (2002-2018), Guatemala (2004-2018), Mexico (2000-2017), and India (2006-2018) that I cleaned with help from research assistants. As Table 1 shows, the data vary by country but cover infractions, money stolen/missing, follow-up on audit recommendations, whistleblower complaints, and relevant information on the sector and sub-sector of the corruption activity. Relevant electoral data accompany those on corruption as well.

Second, this article provides a framework for assessing the validity of non-randomized audit data. In the absence of randomly assigned audits, it is feasible that presidents or prime ministers can use their control of the executive branch, which includes the bureaucracy, to

[20] Larreguy, Marshall and Snyder (2020) and Ajzenman (2020) are, to my knowledge, the only audit studies using administrative data—outside of one-off studies like Di Tella and Schargrodsky (2003) and Olken (2007).

Table 1: Summary of Original Data Collected

	Honduras	**Guatemala**	**Mexico**	**India**
Administrative Level	Municipal	Municipal	Municipal	State
Years Covered	2002-2018	2004-2018	2000-2018	2004-2018
Sector	No	No	Yes	Yes
Infractions	Yes	Yes	Yes	Yes
Precise Details of Infractions	Yes	No	Yes	Partial
Money Stolen/Missing	No	Yes	Yes	Yes
Follow-up on Audit Recommendations	No	No	No	In progress
Whistleblower Complaints	No	Yes	No	No
Bureaucrat Sanctions	No	Yes	No	No

target audits at opponents such as opposition mayors. The risk is particularly high particularly in poor developing democracies and countries with low horizontal accountability.[21]

The first tests that analysts need to run for partisanship pertains the frequency of audits received. If all units (states, municipalities, etc.) do not receive the same number of audits for the given time interval, then it is essential to regress partisanship as well as demographic characteristics (e.g., population) on whether the unit receives an audit. As mentioned above, auditors need discretion to conduct risk-based audits, so analysts should include past corruption as a control variable as well. Since the consequences of partisan tampering with the audit process are highest after close elections, the most convincing tests will complement traditional regression analysis with a close-election regression discontinuity design. Specifically, the latter needs to follow the setup of Brollo and Nannicini (2012) and examine whether party alignment between the executive and lower-level government units predicts audit frequency after close elections. The regression discontinuity design is perhaps the most powerful tests because the close elections can serve as a natural experiment that obviates (Lee, 2008; Eggers et al., 2015; Erikson and Rader, 2017). If the test consistently predicts that opposition/unaligned subnational units receive more audits, then it is clear that there are partisan concerns with the audit process.

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

Partisanship, however, is not the only consideration for the audit independence of audits. The International Organization of Supreme Audit Institutions (2019) also highlights, for example, the need for regulations or legislation—preferably in the Constitution—on tenure security as well as operational and budget autonomy for the heads of the audit agencies.[22] For example, if a country's President or Prime Minister can easily fire and replace the audit agency head at will, it will be difficult for the audit process to be apolitical. The same is true for budgets: if audit agencies receive smaller outlays after undertaking unpopular audits, then independence becomes more elusive. That is why it is generally better if the audit agencies report to Congress or a parliament, such that there is less possibility for a single individual to unduly influence the audit process.

4. Analysis of the Country Cases

In this Section, I apply the framework in Section 3 to analyze the extent to which researchers can use the non-randomized, subnational audits India, Mexico, Guatemala, and Honduras without fear of bias. Essentially, is there an independent process underpinning each country conducts its governmental audits?

4.1. India

Legal Basis for Audit Independence. The Comptroller and Auditor General (CAG) is the institution responsible for conducting governmental audits in India. The Auditor General leads the CAG and 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.

[22] The International Organization of Supreme Audit Institutions codified these recommendations into existence for in its Lima Declaration (1977) and Mexico Declaration (2007) (International Organization of Supreme Audit Institutions, 1977, 2007).

Although the president appoints the Auditor General, 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 independence to carry out unbiased auditing.

Frequency, Selection, and Distribution of Audits. Each year, India's CAG publishes audit reports for each of India's 29 states on its website. The audits reports are very comprehensive, often spanning hundreds of pages. Although the audits from 2004-2018 cover 70 different "sectors", there are audit reports every year for state finances, revenue, civil, commercial, economic, and social sectors, among others. Within these sectoral reports, the CAG generally undertakes in-depth analysis of at least one different sub-sector each year. Accordingly, analysts can use these state-level audits data from India without concern of distributional bias along party lines.

4.2. Honduras

Legal Basis for Audit Independence. The Supreme Tribunal of Accounts (TSC, *Tribunal Superior de Cuentas*) is the institution responsible for government audits in Honduras. Chapter III (Articles 222-227) of the Honduran Constitution and the Organic Law of the TSC (2002, revised 2011) provide the legal basis of the TSC. Overall, its legal basis is strong. Notably, the TSC reports to Congress, not the President, so the chance that any one individual can control the TSC is low. That is especially the case because the head of the TSC rotates annually over a three-year term, with the heads being selected uniquely by Congress; and the Organic Law of the TSC specifically excludes numerous politicians, such as the President, from being a member of the TSC. From an operational autonomy perspective, the TSC's is independence 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.

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. 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 at the TSC is no exception.

Frequency/Selection of Audits. The TSC performs numerous different types of audits of various institutions in Honduras but, as compared to the other countries in this paper, it performs relatively relatively fewer municipal audits. The country has 298 municipalities but only in 2002, when the TSC became a formal SAI, did the country undertake more than 200 audits in a year (see Figure 1a). Since then—and particularly over the course of Juan Orlando Hernandez's presidency (2014-present)—the number of audits undertaken by the TSC has dropped steadily. The numbers will increase slightly as the TSC releases more multi-year audits reports on its website, but the decline is significant.

Regardless, the preliminary logit model in Figure 1b shows that partisan motivations are not driving which municipalities receive audits. I confirm this pattern with a more credible regression discontinuity design in Figure 1c. 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 audit allocation decisions. Consistent with the legal framework, that is not a concern here. As shown in Figure 1, the regression discontinuity analysis passes the McCrary (2008) density test, too, so we can be confident that random variation in close elections is smoothly distributed, and that there are no signs of electoral fraud that could skew the results.

4.3. Guatemala

Legal Basis for Audit Independence. The National Audit Office (*Contraloría General de Cuentas*) is the body in charge of undertaking audits within 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. 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."

Frequency/Selection of Audits. Each year, the National Audit Office audits circa 320 of Guatemala's 340 municipalities as well as numerous other

4.4. Mexico

Legal Basis for Audit Independence. The Supreme Audit Institution of the Federation (ASF, *Auditoría Superior de la Federación*) is the institution responsible for government audits in Mexico. The ASF receives its mandate directly from Articles 74, 79, and 113 of the Mexican Constitution and reports to the Chamber of Deputies, not the President. The Chamber of Deputies also elects the head of ASF (*Auditor Superior*), who serves a term of 8 years that may be renewed once.

Frequency/Selection of Audits. Honduras has 298 municipalities, but budgetary restrictions prevent the TSC from auditing every municipality each year. The table presents

Table 2: Corrective Actions/Findings by Mexican ASF in Municipal Audits (2007-2018)

	Corrective Actions/Findings	Frequency
1	Definitive financial irregularity with required compensatory action	925
2	Performance recommendation	386
3	Punitive noncompliance with regulations	3879
4	Recommendation	11519
5	Report of crime	147
6	Request for clarification	359
7	Statement of financial irregularity/ies with presumed intent	3465
8	Tax evasion or financial regulatory noncompliance with presumed intent	215

Table 3: Types of Mexican Municipal Audits (2007-2018)

	Type of Audit	Frequency
1	Compliance	49
2	Compliance and performance	164
3	Financial and compliance	17442
4	Financial compliance	511
5	Financial compliance with performance focus	799
6	Forensic	56
7	Performance	1831
8	Physical investment	43

5. Conclusion

Audit data have emerged as a suitable alternative to objectively measuring corruption and advancing knowledge than the perceptions data that have dominated the literature since Mauro (1995). However, aside from new studies of Mexico by Larreguy, Marshall and Snyder (2020) and Ajzenman (2020), scholars have focused too much on Brazil. The numerous Brazil studies have produced some useful causal findings, notably due to the randomized nature of the audit data, but the external validity of these studies remains a concern.

In this article, I put forth new subnational audit data from Mexico, Guatemala, Honduras, and India as well as a new framework for assessing whether researchers can profitably use these non-randomized data to analyze corruption. With these data and framework, analysts will be able to undertake sophisticated analyses of corruption across the world. In the process, researchers will move beyond Brazil, stop using problematic perceptions-based

data for knowledge creation,[23] and better understand the true causes and consequences of corruption.

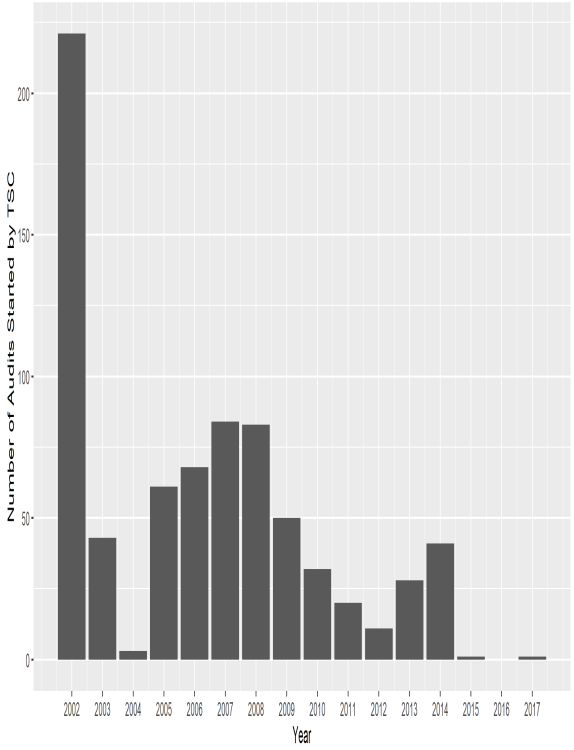
Of course, the sample of countries examined in this paper is not random. In my search for audit data across Latin American and African countries, the above countries are the ones that published their audit reports online for more than one electoral period.[24] Not coincidentally, transparency is also one of the criteria that the International Organization of Supreme Audit Institutions (2019) suggests is essential for audit independence. Nevertheless, the data and framework in this article help open up new areas of knowledge for the study of corruption.

[23] Hollyer (2018) argues that the perceptions are still very useful for ranking countries. I have a different perspective. More recently, as more data on money laundering has come to the fore, it has become even more difficult to reconcile the fact that the corruption perception indexes rank the same countries favorably that fuel grand corruption at the highest scale through shell corporations and the banking system (e.g., Findley, Nielson and Sharman, 2014).

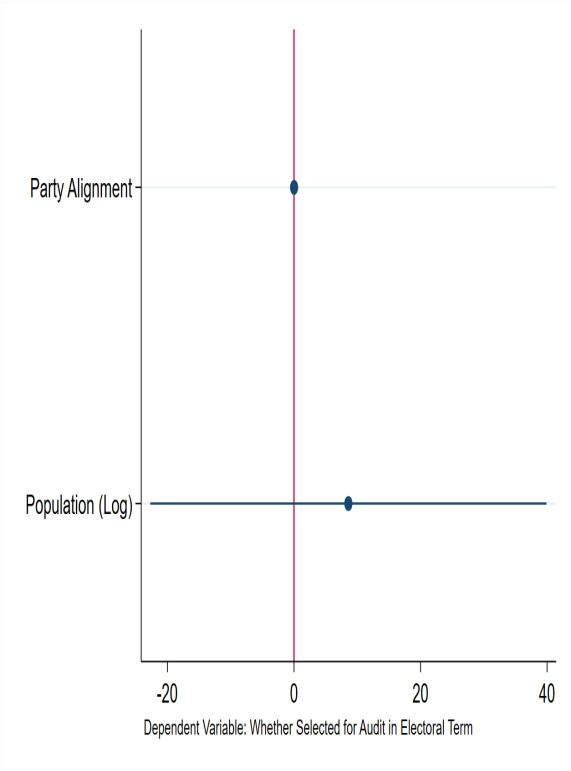
[24] Even then, it required significant effort to clean and convert to the data a usable format. In Africa, Kenya's was the only supreme audit institution to publish multiple years of subnational audit reports online, but the data did not encompass more than one electoral period.

Figure 1: Results: Honduras

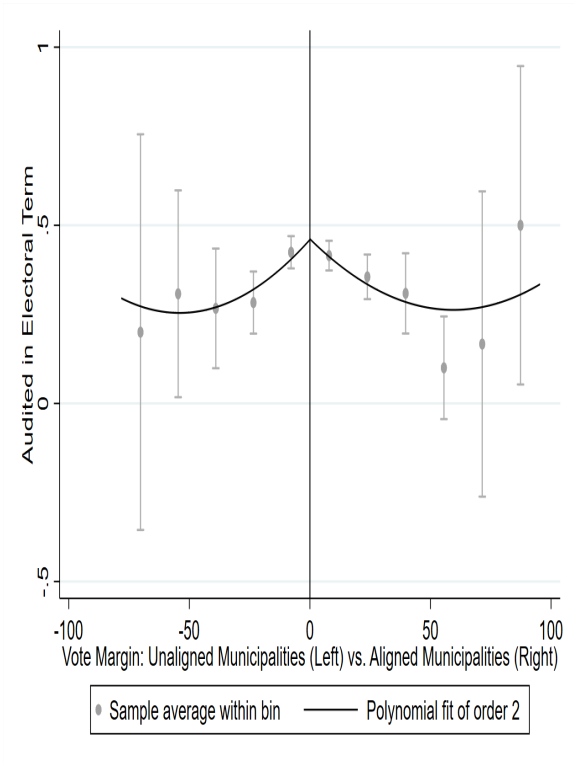
(a) Audits Started by the TSC Over Time



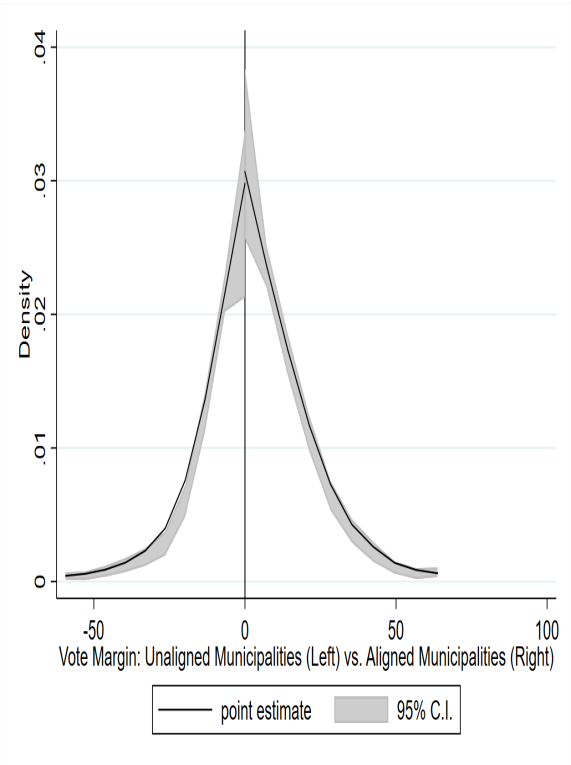
(b) Logit Model with Municipality and Term FE



(c) Regressions Discontinuity Results (Term)



(d) McCrary Density Plot for RDD



References

Acemoglu, Daron and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* New York: Crown Business.

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. 2020. "The Power of Example: Corruption Spurs Corruption." *American Economic Journal: Applied Economics* (Forthcoming).

Alm, James, Jorge Martinez-Vazquez and Chandler McClellan. 2016. "Corruption and Firm Tax Evasion." *Journal of Economic Behavior and Organization* 124:146–163.

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.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, New Jersey: Princeton University Press.

Avis, Eric, Claudio Ferraz and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126(5):1912–1964.

Bersch, Katherine and Sandra Botero. 2014. "Measuring Governance: Implications of Conceptual Choices." *European Journal of Development Research* 26(1):124–141.

Besley, Timothy and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117(4):1415–1451.

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):1–20.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti and Guido Tabellini. 2013. "The Political Resource Curse." *American Economic Review* 103(5):1759–1796.

Brollo, Fernanda and Ugo Troiano. 2016. "What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil." *Journal of Development Economics* 122:28–45.

Broms, Rasmus, Carl Dahlström and Mihály Fazekas. 2019. "Political Competition and Public Procurement Outcomes." *Comparative Political Studies* 52(9):1259–1292.

Calonico, Sebastian, Matias D. Cattaneo and Rocío Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

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

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.

Dahlström, Carl and Mikael Holmgren. 2019. "The Political Dynamics of Bureaucratic Turnover." *British Journal of Political Science* 49(3):823–836.

Dhaliwal, Iqbal and Rema Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 124:1–21.

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.

Dizon-Ross, Rebecca, Pascaline Dupas and Jonathan Robinson. 2017. "Governance and the Effectiveness of Public Health Subsidies: Evidence from Ghana, Kenya and Uganda." *Journal of Public Economics* 156:150–169.

Duflo, Esther, Rema Hanna and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4):1241–1278.

Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall and James M. Snyder. 2015. "On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races." *American Journal of Political Science* 59(1):259–274.

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.

Erikson, Robert S. and Kelly Rader. 2017. "Much Ado About Nothing: RDD and the Incumbency Advantage." *Political Analysis* 25(02):269–275.

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

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 and Gábor Kocsis. 2020. "Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data." *British Journal of Political Science* 50(1):155–164.

Fazekas, Mihály, Luciana Cingolani and Bence Tóth. 2018. Innovations in Objectively Measuring Corruption in Public Procurement. In *Governance Indicators: Approaches, Progress, Promise*, ed. Helmut Anheier, Matthias Haber and Mark Andreas Kayser. Oxford: Oxford University Press chapter 7, pp. 154–185.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of

Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2):703–745.

Ferraz, Claudio and Frederico Finan. 2011. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." *American Economic Review* 101(4):1274–1311.

Ferraz, Claudio and Frederico Finan. 2018. Fighting Political Corruption: Evidence from Brazil. In *Institutions, Governance, and the Control of Corruption*, ed. Kaushik Basu and Tito Cordella. Cham, Switzerland: Palgrave Macmillan chapter 9, pp. 253–284.

Ferraz, Claudio, Frederico Finan and Diana B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96(9-10):712–726.

Findley, Michael G., Daniel L. Nielson and Jason C. Sharman. 2014. *Global Shell Games: Experiments in Transnational Relations, Crime, and Terrorism*. New York: Cambridge University Press.

Findley, Michael G., Kyosuke Kikuta and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* .

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 Miriam A. Golden. 2017. *Corruption: What Everyone Needs to Know*. Oxford: Oxford University Press.

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.

Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho and Mauricio Ruiz-Vega. 2016. "When to Protect? Using the Crosswise Model to Integrate Protected and Direct Responses in Surveys of Sensitive Behavior." *Political Analysis* 24:132–156.

Gisselquist, Rachel M. 2014. "Developing and Evaluating Governance Indexes: 10 Questions." *Policy Studies* 35(5):513–531.

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.

International Organization of Supreme Audit Institutions. 1977. "The Lima Declaration.".

International Organization of Supreme Audit Institutions. 2007. "Mexico Declaration on Supreme Audit Institutions (SAI) Independence.".

International Organization of Supreme Audit Institutions. 2019. "Independence of Supreme Audit Institutions.".
URL: *https://www.intosai.org/fileadmin/downloads/focus_areas/independence/EN_0219_Handout_Indep*

Jetter, Michael, Alejandra Montoya Agudelo and Andrés Ramírez Hassan. 2015. "The Effect of Democracy on Corruption: Income is Key." *World Development* 74:286–304.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal on the Rule of Law* 3(2):220–246.

Keefer, Philip. 2007*a*. "Clientelism, Credibility, and the Policy Choices of Young Democracies." *American Journal of Political Science* 51(4):804–821.

Keefer, Philip. 2007*b*. The Poor Performance of Poor Democracies. In *Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan C. Stokes. Oxford: chapter 36, pp. 886–909.

Khan, Adnan Q., Asim Ijaz Khwaja and Benjamin A. Olken. 2016. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics* pp. 219–271.

King, Gary, Robert Keohane and Sydney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, New Jersey: Princeton University Press.

Klašnja, Marko. 2015. "Corruption and the Incumbency Disadvantage: Theory and Evidence." *Journal of Politics* 77(4):928–942.

Kolstad, Ivar and Arne Wiig. 2016. "Does Democracy Reduce Corruption?" *Democratization* 23(4):1198–1225.

Kraay, Aart and Peter Murrell. 2016. "Misunderestimating Corruption." *Review of Economics and Statistics* 98(3):455–466.

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. 2012. "Corruption's Challenge to Democracy: A Review of the Issues." *Politics and Policy* 40(5):802–826.

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* .

Lederman, Daniel, Norman V. Loayza and Rodrigo R. Soares. 2005. "Accountability and Corruption: Political Institutions Matter." *Economics and Politics* 17(1):1–35.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics* 142(2):675–697.

Locke, Richard M. and Kathleen Thelen. 1995. "Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics." *Politics and Society* 23(3):337–367.

Lührmann, Anna, Kyle L. Marquardt and Valeriya Mechkova. 2020. "Constraining Governments: New Indices of Vertical, Horizontal, and Diagonal Accountability." *American Political Science Review* 114(3):811–820.

Marquardt, Kyle L. and Daniel Pemstein. 2018. "IRT Models for Expert-Coded Panel Data." *Political Analysis* 10(4):431–456.

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

Mauro, Paolo. 1998. "Corruption and the Composition of Government Expenditure." *Journal of Public Economics* 69(2):263–279.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2):698–714.

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.

Mungiu-Pippidi, Alina. 2015. *The Quest for Good Governance: How Societies Develop Control of Corruption*. New York: Cambridge University Press.

Musila, Jacob W. 2013. "Does Democracy Have a Different Impact on Corruption in Africa?" *Journal of African Business* 14(3):162–170.

Naritomi, Joana. 2019. "Consumers as Tax Auditors." *American Economic Review* 109(9):3031–3072.

O'Donnell, Guillermo. 1998. "Horizontal Accountability in New Democracies." *Journal of Democracy* 9(3):1112–1126.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(21):200–249.

Pereira, Carlos and Marcus André Melo. 2015. "Reelecting Corrupt Incumbents in Exchange for Public Goods: Rouba mas faz in Brazil." *Latin American Research Review* 50(4):88–115.

Potrafke, Niklas. 2019. "Electoral Cycles in Perceived Corruption: International Empirical Evidence." *Journal of Comparative Economics* 47(1):215–224.

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.

Schwindt-Bayer, Leslie A. and Margit Tavits. 2016. *Clarity of Responsibility, Accountability, and Corruption*. New York: Cambridge University Press.

Snyder, James M. and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118(2):355–408.

Standaert, Samuel. 2015. "Divining the Level of Corruption: A Bayesian State-Space Approach." *Journal of Comparative Economics* 43(3):782–803.

Stanig, Piero. 2018. Considerations on the Method of Constructing Governance Indices. In *Governance Indicators: Approaches, Progress, Promise*, ed. Helmut Anheier, Matthias Haber and Mark Andreas Kayser. Oxford: Oxford University Press chapter 6, pp. 134–153.

Stegmueller, Daniel. 2011. "Apples and Oranges? The Problem of Equivalence in Comparative Research." *Political Analysis* 19(4):471–487.

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

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.

Trochim, William. 2006. "The Research Methods Knowledge Base, 2nd Edition.".
URL: *http://www.socialresearchmethods.net/kb/*

Zamboni, Yves and Stephan Litschig. 2018. "Audit Risk and Rent Extraction: Evidence from a Randomized Evaluation in Brazil." *Journal of Development Economics* 134(April):133–149.