Corruption and Citizens' Compliance with the Law: An Empirical Analysis*

Nicolas Guida-Johnson†

This Version: August 2022

First Version: September 2019

Abstract

In this paper, I empirically analyze the effect of the disclosure of corruption cases on citizens' compliance with the law. In order to do so, I use data on corruption cases generated by the Brazilian anti-corruption plan, "Programa de Fiscalização em Entes Federativos por Sorteios Públicos", which randomly audits municipalities for their use of federal funds. I measure non-compliance with the law by citizens using data on traffic offenses at the municipality level. The random selection of municipalities to be audited provides me with a straightforward empirical strategy. My main results indicate that the disclosure of corruption cases at the municipality level increases per capita traffic offenses by 1.2%, and an additional case of corruption disclosed increases traffic offenses per capita by 0.4%. These estimates are small and not statistically different from zero. Therefore, I cannot conclude that the disclosure of corruption impacts compliance with the law by citizens.

Keywords: Corruption, Compliance with the Law, Traffic Offenses, Brazil

JEL Codes: D73, H70, K42

*I thank M. Daniele Paserman, Patricia Cortés, Martin Fiszbein, Nicolás Ajzenman, Christian Ruzzier, and seminar participants at Boston University for helpful feedback and comments. All remaining errors are my own

†Boston University, Economics Department. Contact Information: guidaj@bu.edu

1 Introduction

Corruption and lawbreaking behavior among citizens are likely influenced by the same causes: institutions and social norms. Societies with less legal enforcement and where unlawful behavior is socially acceptable are more likely to have high levels of corruption and low levels of citizens' compliance with the law. But can corrupt public officials induce lawbreaking behavior among citizens?

In this paper, I empirically analyze whether the disclosure of corruption cases affects citizens' compliance with the law. Using data from the anti-corruption plan "Programa de Fiscalização em Entes Federativos por Sorteios Públicos" (Monitoring Program with Public Lotteries) from Brazil, I analyze if the corruption cases disclosed at the municipality level by this program had an impact on the compliance with the law by citizens. To measure non-compliance with the law, I use traffic offenses detected and reported by the "Polícia Rodoviária Federal" (Federal Highway Patrol). Corruption could affect citizens' behavior other than through infractions to the traffic law. However, given that the data on traffic offenses used in this paper is collected by a law enforcement agency independent from the authorities that are audited by the anti-corruption plan, variations in traffic offenses can be attributed to variations in rule-breaking by citizens. In other unlawful activities, for instance, evasion of municipal taxes, one could think that the anti-corruption plan could encourage local public officials to reduce local tax evasion. As a result, variations in the evasion of local taxes would not directly imply higher or lower levels of compliance with the law by citizens.

The anti-corruption plan implemented in Brazil between 2003 and 2015 consisted of random audits of municipalities for their use of federal funds. All municipalities with up to 500,000 inhabitants were eligible for selection. Initially, the first two lotteries selected 5 and 26 municipalities, respectively. Afterward, 60 municipalities were chosen randomly in each lottery. Given the sample period of information about traffic offenses (January 2007- September 2014), I work with the data coming from lotteries conducted between July 2006 and March 2013. In all of those municipalities audited during this period, at least one act of corruption associated with the use of federal funds was found in each municipality.

The identification of the causal effects of the disclosure of corruption on citizens' compliance with the law is challenging. Both corruption and unlawful behavior among citizens are likely to be determined by the same observable and unobservable characteristics of a society. A simple comparison of societies with different levels of corruption and lawbreaking behavior is bound to be biased. To address these empirical challenges, I exploit the exogenous variation in citizens' exposure to corrupt politicians generated by the anti-corruption plan. Thanks to the randomization built into this plan and the fact that in all municipalities audited between July 2006 and March 2013, there is at least one case of corruption, the basic empirical strategy is straightforward.

My main results do not allow me to conclude that the disclosure of corruption impacts compliance with the law by citizens. I find a small and not statistically different from zero effect of the disclosure of corruption on traffic offenses per capita: disclosing corruption cases is associated with an increase in the number of violations to the traffic law per capita of 1.2%. An extra corruption case disclosed leads to an increase of 0.4% in the number of infractions per capita. These estimates are larger when I focus on the disclosure of high levels of corruption or on municipalities with local media sources. However, all of these estimates are not statistically different from zero. Consequently, based on my empirical results, I cannot conclude that corruption harms compliance with the law by citizens.

From a theoretical standpoint, Acemoglu and Jackson (2015) show that a society's leaders or prominent agents can impact the evolution of social norms. These authors take as an example the figure of a prominent police officer who can choose a highly visible honest action to break the social norm of corruption. Social norms shape beliefs and behavior and can change over time in response to individual behavior by prominent agents. Additionally, Shleifer and Vishny (1993) and Olken and Pande (2012) argue that the incentive structure (determined by government institutions) faced by public officials as well as the bureaucratic organization might be important factors determining the level of corruption. Wages in the public sector, the level of monitoring, the probability and severity of punishment, and other incentives help explain the individual's decision to engage in corrupt behavior. Consequently, the disclosure of corruption cases could be perceived by citizens as a signal of an institutional setting where the benefits of engaging in corrupt behavior exceed the costs of it. Then, given that corruption is mostly a lawbreaking activity, citizens might internalize these signals and reconsider the benefits and costs associated with breaking the law.

Empirically, most studies find strong correlations between individual behavior and aggregate levels of corruption. For instance, Barr and Serra (2010) find among a population of undergraduate students a correlation between the individual propensity to act corruptly and the level of corruption in their home country. Abbink et al. (2018), using a sequential bribery game experimental design, show that participants who knew that they were interacting with a partner from a group with a majority of corrupt (as opposed to honest) partners offered twice as many bribes. In other words, the norms that arise from the typical behavior of a group affect individual decisions. Moreover, in this case, the social norms associated with corruption impact the individual propensity to engage in corrupt behavior.

Moreover, Gächter and Schulz (2016) using an index of "prevalence of rule violation" (PRV), which is based on country-level data on corruption, tax evasion, and fraudulent politics, find that individual intrinsic honesty (measured by a die-rolling experiment) is stronger in the subject pools of low PRV countries than those of high PRV. In other words, individuals from countries where corruption, tax evasion, and fraudulent politics are pervasive problems are more likely to be intrinsically dishonest than those from countries where those measures of rule violation are low. Even though this last paper focuses on intrinsic honesty, it is related to my research study since measures of honesty have been shown to be correlated with fraudulent behavior. In particular, Hanna and Wang (2017) show that among public sector employees (government nurses), those who are likely to be more dishonest tend to be fraudulently absent from work. Therefore, dishonest individuals are more likely to engage in lawbreaking activities. Combining the results from Hanna and Wang (2017) and Gächter and Schulz (2016), individuals from more corrupt countries are more likely to be willing to break the law.

Additional evidence of the correlation between individual dishonest behavior and corruption at the country level is given by Orosz et al. (2018). Using data from 40 countries, the authors find a strong relationship between self-reported academic cheating on exams and the country level of the corruption perception index (based on the index constructed by International Transparency).

On the other hand, Cameron et al. (2009) find mixed evidence about corruption and individual propensity to engage in corrupt behavior and to punish it. In particular, their results from India and Australia suggest that greater exposure to corruption in daily life may build a greater tolerance of corruption, with the Indian subjects showing a greater propensity to engage in and a lower propensity to punish corrupt behavior. However, the results from Singapore and Indonesia do not support this argument. Indonesia is consistently ranked as having high levels of corruption, yet its subjects displayed a relatively low tolerance for corruption. In contrast, Singapore is a relatively low corruption country, but the Singaporean subjects in their experiments showed a relatively high willingness to engage in corruption and reluctance to punish it. The authors conjecture that the findings in Singapore and Indonesia could reflect the recent institutional changes that these countries overtook.

There is strong evidence in support of a correlation between the level of corruption of a country and the propensity to break the law by its citizens. However, it is not possible to derive a causal relationship between these two variables from these studies. For instance, it could be the case that both are jointly determined by the institutional characteristics of the country, and then an exogenous increase in corruption would not have any impact on the propensity to break the law by citizens. In this sense, by exploiting the random nature of the disclosure of corruption embedded in the anti-corruption plan implemented in Brazil, this paper would help to shed some light on this relation.

My paper is closely related to Ajzenman (2021). In this paper, the author finds that in Mexico, following the revelations of corruption by local officials, cheating in cognitive tests by secondary school students increases significantly. Given that the disclosure of corruption cases is not random, the author uses a difference-in-difference approach which can suffer from internal validity issues if the typical assumptions in this framework are not reasonable. Here, I take advantage of the random selection of municipalities to be audited in Brazil under the "Programa de Fiscalização em Entes Federativos" and therefore, the empirical strategy will not be subject to the potential issues in Ajzenman (2021).

There is a widespread consensus on the importance of the rule of law to development. In this manner, my paper contributes to the literature that shows the potential adverse effects of corruption on different aspects of development. Among others, Bardhan (1997) shows the effects of corruption on efficiency and growth, Ferraz et al. (2012) find a negative impact of corruption in the educational performance of primary school students, Del Monte and Papagni (2001) show the negative impact of corruption on the efficiency of public expenditures and Svensson (2003) show how corruption can affect the cost of doing business for firms.

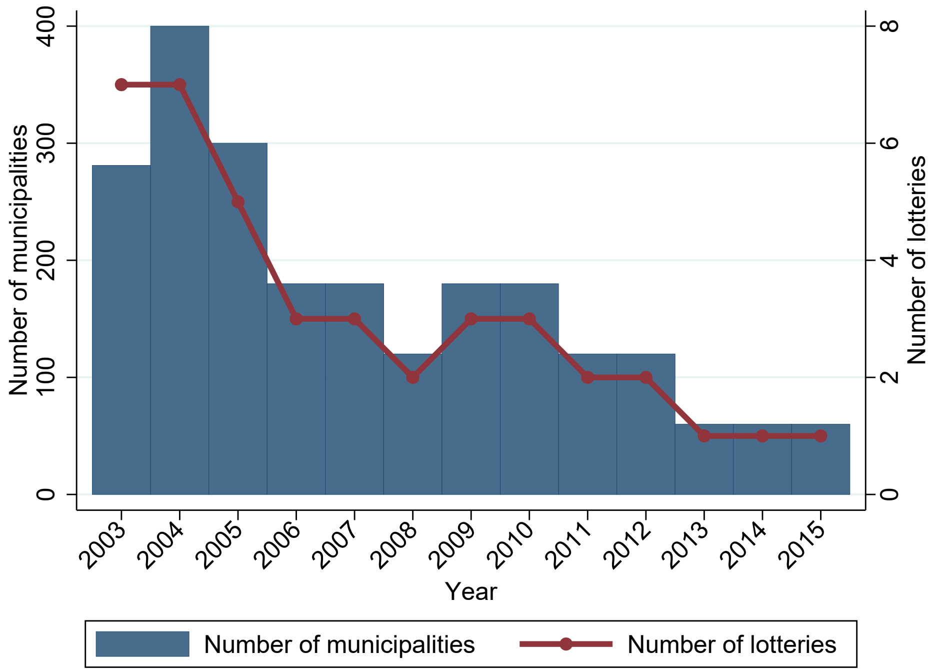
The remainder of this paper proceeds as follows: Section 2 describes the anti-corruption plan implemented in Brazil between 2003 and 2015 and presents its main characteristics. Section 3 discusses the data sources used in this paper to empirically analyze the relationship between the disclosure of corruption and citizen compliance with traffic law. Section 4 describes the empirical strategy, and Section 5 shows the results. Finally, concluding remarks follow in section 6.

2 Background

Between 2003 and 2015, the federal government of Brazil implemented an anticorruption program called "Programa de Fiscalização em Entes Federativos por Sorteios Públicos" (Monitoring Program with Public Lotteries), based on random audits of municipalities for their use of federal funds. The program was conducted by the "Controladoria Geral da Uniõa" (Office of Comptroller-General), an autonomous and ministerial rank agency that focuses on combating corruption at different administrative levels of Brazil. The selection of municipalities to be audited was random, and the lotteries were held publicly in conjunction with the national lottery in Brasilia.

All municipalities with up to 500,000 inhabitants were eligible for selection (from lottery 24 onward, capital municipalities were not eligible for selection). Initially, the first two lotteries selected 5 and 26 municipalities, respectively. Afterward, 60 municipalities were chosen randomly in each lottery. As of February 2015, 2,241 audits across 40 lotteries in 1,913 municipalities have been conducted. From this set of randomly selected municipalities, 1,619 were selected and audited only once, 261 two times, 32 three times, and only one municipality was selected and audited four times. Figure 1 shows the evolution over time of the program. The program reduced the number of lotteries per year over time and, thus, the number of municipalities audited. For the last three years in which the program was running, only one lottery was held per year, and 60 municipalities were audited each year.

Figure 1. Number of Lotteries Held and Municipalities Audited per Year.



Notes: this figure shows the number of municipalities randomly selected and the number of lotteries per year, for the full duration of the anticorruption plan in Brazil. Own elaboration based on data of the Office of Comptroller-General.

During each audit, the Office of Comptroller-General gathers information on all federal funds transferred to the municipality during the previous 3-4 years and issues a random selection of inspection orders on specific projects. The auditors from the Office of Comptroller-General are hired on the basis of a competitive public examination and earn highly competitive salaries; therefore, their incentives for corruption during the audits are lower than those of other federal-level bureaucrats (Avis et al., 2018).

After the audit is completed, a detailed report describing all the irregularities found is submitted to the central office of the Office of Comptroller-General in Brasilia. Moreover, a summary of the findings is posted online and disclosed to primary media sources. Each inspection took approximately ten days (Ferraz and Finan, 2008); this time could vary depending on the size of the municipality and the number of projects to be audited. Furthermore, from the moment a municipality was drawn from a lottery, it took approximately ten days until the audit began. Moreover, the results of the audits were publicly available between 6 to 12 months after the lottery. Once all reports from the municipalities audited in a particular lottery were finished, the Office of Comptroller-General announced through its webpage that reports were publicly available and also provided a summary of the main findings of the audits. Clearly, there are two critical dates in the process of this anticorruption plan that would be relevant to the central question of this paper. First, the date in which a municipality is randomly chosen to be audited and second, the date in which the results of the audit are released to the public. When these results show that there were cases of corruption at the municipality level in the use of federal funds, I take this event as a disclosure of corruption.

Although over time there were some changes to the program (such as limiting the number of randomly selected sectors in larger municipalities to be inspected and changing the time that had to elapse for a municipality to be audited again), these do not affect a municipality's audit probability conditional on being eligible to be audited. The randomization was performed at the state level; therefore, the probability of being audited is constant for municipalities within the same state. For small states, only one or two municipalities were typically drawn in a single lottery, while for large states, around ten municipalities were typically drawn.

Given the objective of this study, i.e., to analyze the impact of corruption on citizen compliance with the law, it is of great relevance whether the citizens learned about the information of use or misuse of federal funds at the municipality level that came out from the audits. Ferraz and Finan (2008) provide anecdotal evidence suggesting that the information from the audits reached citizens, and in particular, it was used during municipal campaigns. Moreover, the same authors find in their study that corrupt politicians were punished relatively more in places where local radio stations were present to divulge the findings of the audit reports. Additionally, anecdotal evidence from newspaper articles shows that information about the results of the audits reached citizens and also in some municipalities took them by surprise. Overall, there is evidence that the information about the audits reached citizens through local radios or newspapers. Furthermore, the evidence supports the idea that this information was not irrelevant to the population. Consequently, it could have affected citizen behavior regarding compliance with laws.

3 Data

In this paper, I analyze whether the disclosure of corruption cases at the municipality level impacts the number of traffic offenses or infractions to the traffic law. In order to do so, my two primary sources of information are data on traffic offenses provided by the "Polícia Rodoviária Federal" (Federal Highway Patrol) and data on corruption cases at the municipality level generated by the anticorruption plan implemented in Brazil and conducted by the Office of Comptroller-General.

The data on traffic offenses is publicly available on the website of the law enforcement agency and covers the period 2007-2014. The data contains 11,752,912 traffic offenses detected between January 2007 and September 2014. It contains information on all traffic offenses detected by the federal law enforcement agency on federal roads and highways. It provides information on the time and date of the infraction, the municipality, and state where the infraction took place, the state where the registration plate of the car was issued, the highway or road where the infraction took place, the type of infraction, the type of vehicle and the brand and model of the vehicle.

Given the quality of the data, I can construct a monthly panel data of municipalities with the total number of traffic offenses detected in each municipality per month. Furthermore, I can disaggregate the data by type of infraction. In this manner, I consider the most common types of traffic offenses and classified the data into nine categories of infractions: speeding, driving without a seat belt, illegal parking, driving under the influence of illegal substances, illegal equipment (defective equipment or modifications to the vehicle that are illegal), illegal driving (for example, making a U-turn when it is not allowed, overtaking another vehicle when it is illegal, etc.), illegal ID (not proper vehicle or individual identification), red-light crossing and other infractions.

Additionally, corrections or modifications to the information provided by the Federal Highway Patrol were performed. First, some infractions were coded by this agency occurring on a highway or road that does not exist (around 0.06% of the total number of infractions recorded). These observations were deleted from the sample. Second, in some cases, the information about the municipality and the state where the infraction took place was incongruous; i.e., the infraction was detected in a municipality that do not belong to the state coded (around 0.2% of the total number of infractions recorded). For example, an infraction is coded taking place in the municipality of São João da Fronteira which belongs to the state of Piauí, but the state code for that infraction was Bahía. When these errors were identified, the state was corrected. In the example provided, the state where the infraction took place was changed from Bahía to Piauí. However, when the name of the municipality corresponds to more than one municipality and the state code is incorrect, then the observation was deleted (around 0.01% of the total number of infractions). For example, the name "Alto Paraíso" corresponds to two different municipalities: one in the state of Rondonia and one in the state of Paraná. If for an infraction that took place in "Alto Paraíso", the state appearing in the dataset from the law enforcement agency was not either Rondonia or Paraná, then I am not able to identify to which state the observation belongs. In this manner, the observation is deleted. Finally, after these adjustments to the data from the Federal Highway Patrol, I obtain a panel dataset that contains 1,372 municipalities and covers the period from January 2007 to September 2014 (93 months).

Table 1 presents summary statistics for each infraction type over the sample of 1,372 municipalities and 93 periods. The data from three municipalities (Río Branco, Florianópolis, and Aracajú; capitals of the state of Acre, Santa Catarina, and Sergipe, respectively) was included only until July 2007, the date of the lottery 24. From lottery 24 onward, municipalities that were the capital of their state were not longer eligible to be randomly selected to be audited. As can be seen, the average municipality-month observation contains a total of 38.6 infractions. Moreover, there is a large variability in the number of infractions by municipality per month. For some municipalities, no infractions were detected by the Federal Highway Patrol in some months. There is a municipality in which there were 3,464 infractions in a month. From the classification of traffic offenses, the most frequent are speeding (average municipality-month observation has 6.6 speeding infractions), illegal driving (average municipality-month observation has 12.3 illegal driving infractions) and illegal ID (average municipality-month observation has 7.5 illegal ID infractions).

Table 1: Summary Statistics on Traffic Offenses

Variable	Observations	Mean	Std. Dev.	Min	Max
All Infractions	127,333	38.57	93.64	0	3464
Speeding	127,333	6.616	60.65	0	3428
No Seat Belt	127,333	2.78	9.05	0	354
Illegal Parking	127,333	0.58	3.19	0	148
Illegal Substance Usage	127,333	0.52	1.75	0	59
Illegal Equipment	127,333	3.50	10.52	0	294
Illegal Driving	127,333	12.30	30.40	0	625
Illegal ID	127,333	7.77	17.17	0	363
Red-Light Crossing	127,333	0.03	0.50	0	34
Other Infractions	127,333	4.50	10.95	0	224

Notes: this table shows statistics on the total number of each type of infraction per month over the sample of 1,372 municipalities constructed based on the data available for the period 01/2007-09/2014 from the Federal Highway Patrol.

In terms of the relevance of this data as a proxy for citizen compliance with the law, I believe it has two main advantages. First, behavior in the street and, in particular, the decision to comply with traffic law is a day-to-day activity. Consequently, I can obtain relatively high-frequency (monthly) data on citizen compliance with the law. It is clear, however, that traffic offenses partially capture compliance with the law by citizens. Other potential measures for non-compliance with the law would be, for example, tax evasion. However, data on this measure at the municipality level is not available. In this context, the second and most important advantage of using traffic offenses as a proxy for compliance/non-compliance with the law is that the Federal Highway Patrol does not respond to the municipality authorities. It is a federal law enforcement agency that responds to the National Ministry of Justice. Therefore, municipality authorities are unlikely to be able to manipulate the behavior of the law enforcement agency. One concern would be that once a municipality is chosen to be audited or when the information about the audits is available (and therefore, potential cases of corruption are disclosed), municipality authorities could increase the number of law enforcement agents as a response to the arrival of auditors to the municipality or as a response to the information disclosed by the audits. Therefore, there would be a spurious relation between the audits and traffic offenses. However, this is an unlikely event given that municipality authorities do not have control over this law enforcement agency.

The information about corruption cases comes from the information generated by the anticorruption plan implemented in Brazil. Given the sample period of information about traffic offenses (January 2007-September 2014), I work with the data coming from lotteries 22 to 38 (lotteries conducted between July 2006 and March 2013). Table 2 presents information for this set of lotteries on the number of municipalities audited in each lottery, the date on which the lottery took place, and the date on which the information gathered during the audits was made publicly available. As it was mentioned previously, there is a gap of around 6 to 12 months between the moment in which a municipality is chosen to be audited and the moment in which the information of the audit is released to the public. The dates on which information of the lotteries 21 and 39 were disclosed are out of the sample of the data on traffic offenses. Consequently, I work with the data coming from audits associated with lotteries 22 to 38.

Table 2: Lotteries, Number of Audits per Lottery and Dates of Lottery and Release of Information

Lottery	Number of Audited Municipalities	Date of Lottery	Date of Disclosure of Data from Audits
22	60	July 2006	July 2007
23	60	May 2007	January 2008
24	60	July 2007	April 2008
25	60	October 2007	June 2008
26	60	April 2008	December 2008
27	60	October 2008	April 2009
28	60	May 2009	January 2010
29	60	August 2009	April 2010
30	60	October 2009	July 2010
31	60	March 2010	January 2011
32	60	May 2010	January 2011
33	60	July 2010	March 2011
34	60	August 2011	April 2012
35	60	October 2011	June 2012
36	60	July 2012	January 2013
37	60	October 2012	June 2013
38	60	March 2013	November 2013

Notes: this table shows for lotteries 22 to 38, the number of audited municipalities, the data of each lottery and the date of the disclosure of the information on audited municipalities (i.e. mainly, the characteristics of the corruption cases found by the auditors, if any).

The measure of corruption disclosure at the municipality level is built based on a database generated by the Office of Comptroller-General; this is the same database used in Avis et al. (2018). Based on the findings of each inspection order, the auditors describe the irregularity found and classify it as (1) an act of mismanagement (for example, documents were not properly filled out or there was improper storage of food supplies and medical equipment), (2) an act of moderate corruption, or (3) an act of severe corruption.

As in Avis et al. (2018), the distinction between mismanagement and corruption is clear. However, the distinction between moderate and severe corruption made by the auditors does not seem to be clear or obvious. It is possible to find among the audits acts of corruption with a similar degree of severity that were classified as moderate corruption in some municipalities and severe corruption in others. Moreover, no clarification was made by the auditors to justify these classifications. For instance, consider these two cases of corruption: (1) a contract of transportation services was awarded to a firm that did not match the original proposal, and the value of the contract was an amount different from what was offered; (2) school lunches had not been delivered for an entire year in one school and had disappeared for a month in the other two schools, even though the municipality had received the money to pay for a school lunch program. Case (1) was classified as a severe act of corruption while case (2) as moderate act of corruption. One can argue that both cases have a similar degree of severity and, therefore should be classified as severe cases of corruption. In any case, the distinction made by auditors is not clear or obvious.

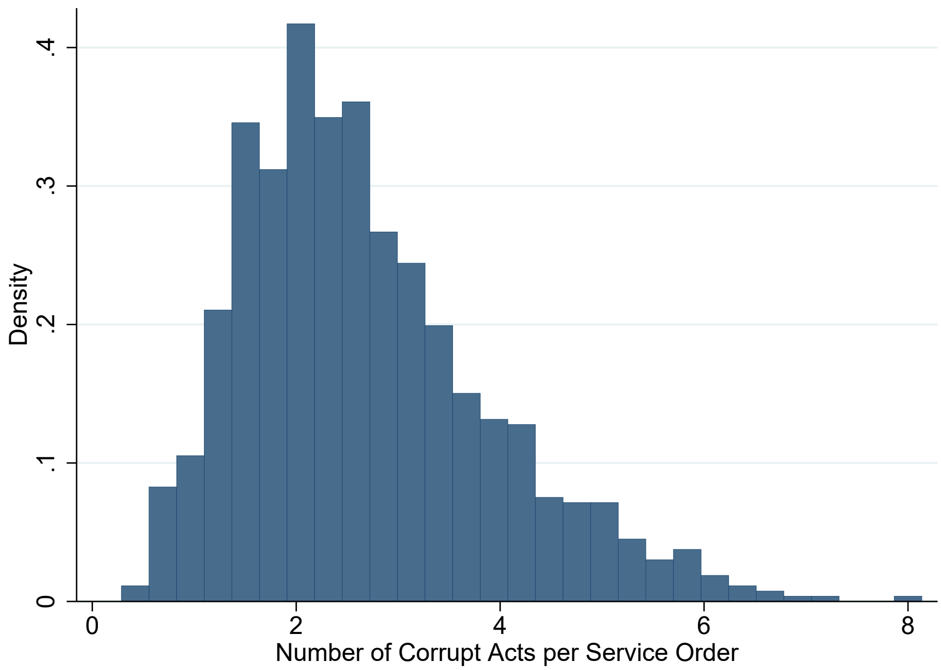
One possible explanation for these unclear classifications of corruption cases is that the distinction between moderate and severe corruption was made solely based on the auditors' views. Furthermore, there was not a clear criterion elaborated by the Office of Comptroller-General that the auditors could follow. Therefore, following the procedure of Avis et al. (2018), I choose to drop this distinction and consider both types of irregularities as corruption.

Consequently, the data generated by the anticorruption plan contains measures of corruption and mismanagement at the municipality-lottery level. The measure of corruption is the number of irregularities classified as either moderate or severe. The measure of mismanagement is the number of irregularities associated with administrative and procedural issues.

In this paper, I will focus on acts of corruption. My thesis is that the disclosure of corruption could impact the behavior of citizens regarding the law; knowing that the authorities where the individual lives engage in acts of corruption could affect his or her incentives to break the law. Therefore, acts of mismanagement do not seem to potentially affect the incentives to comply or not with the law. These acts can be interpreted as a signal of a low quality of politicians. On the other hand, corruption can be interpreted as a signal of a low probability of getting caught breaking the law, a low level of monitoring, or a low degree of punishment, within the municipality. These signals could potentially have an impact on compliance with the law by citizens.

Figure 2 depicts the distribution of irregularities associated with corruption per service order for the municipalities audited between lottery 22 and lottery 38. Recall that each audit was composed of randomly selected inspection orders. For instance, one municipality selected to be audited could have been investigated about school construction and purchase of medicine, while another municipality selected to be audited could have been investigated about transportation services and the provision of lunch to schools. As can be seen from Figure 2, in every audit, some irregularity associated with corruption was found. This characteristic of the corruption disclosed on audits associated with lotteries 22 to 38 is an essential component of my empirical strategy (as it will be described in Section 4). The auditors discovered, on average, 2.7 acts of corruption, and the maximum number of corrupt acts per service order is 8.

Figure 2. Distribution of the Number of Corrupt Acts per Service Order.



Notes: This figure shows the distribution of the number of corruption acts per service order from the municipalities randomly selected in lotteries 22-38. Own elaboration based on information from the Office of Comptroller-General.

Then, I combine the information of the data set of corruption acts and the data set of traffic offenses. In order to empirically answer the central question of this paper, I restrict the analysis to municipalities never selected between the lotteries 22 to 38 and municipalities selected only once and drawn from one of the lotteries 22 to 38. By doing so, out of the 1,372 municipalities with information on traffic offenses, 218 municipalities were randomly selected in lotteries 22 to 38. The rest of the municipalities were never selected to be audited or were selected after the period under analysis. Therefore, in approximately 16% of the municipalities in my data set, corruption acts were disclosed by the random audits. Recall from Figure 2 that in all the municipalities audited during lotteries 22-38, at least one act of corruption was found.

I also use different sources of information in order to obtain certain relevant characteristics of the municipalities. From the 2000 Census, conducted by the "Instituto Brasileiro de Geografia e Estatística" (IBGE; Brazilian Institute of Geography and Statistics), I obtain information for each municipality on the population, population density, the share of women in the population, the share of the population living in urban areas, the share of youths (individuals aged 18 to 24 years), income per capita, Gini index, IDHM index, the proportion of poor individuals, the share of individuals aged 18+ with secondary education, the share of individuals aged 25+ with a college degree and the share of illiterates.

From the 2005 "Perfil dos Municípios Brasileiros" (a survey on municipalities characteristics), also conducted by the IBGE, I obtain information on whether or not each municipality has a university, an AM radio station, an FM radio station, and a TV station. Finally, from the "Instituto Nacional de Meteorologia" (National Institute of Meteorology), I obtain information at the state level for the period under analysis on weather characteristics. These variables include average cloudiness by month, total rain per month, mean maximum temperature, and mean minimum temperature.

Table 3 presents summary statistics for the municipalities in my sample, by whether they were randomly selected to be audited or not. For each characteristic, I also present the difference between these characteristics. The randomization built into the anticorruption plan guarantees, in principle, that on average the municipalities selected to be audited are similar in terms of observable and unobservable characteristics. However, it is important to notice that the randomization was performed over the entire set of municipalities in Brazil, not over those municipalities in my sample, which is a subset of municipalities with information on traffic offenses detected by the Federal Highway Patrol. In this manner, it is important to analyze whether the selection and audit of a municipality is still an exogenous variable in my sample of municipalities. Table 3 provides evidence that in this sample, the randomized nature of the selection of municipalities to be audited is preserved. As can be seen, out of 18 characteristics, only one (share of youth in total population) is statistically significant at 5% level. Therefore, from this evidence, I am able to state that in the sample of municipalities with information on traffic offenses, those selected to be audited are similar in observable characteristics to those not selected. Moreover, given this evidence, it is reasonable enough to assume that they are also similar in unobservable characteristics.

In terms of the characteristics of the municipalities in my sample, Table 3 shows that the majority of the population is urban (around 60%), there is a large proportion of poor individuals (approximately 40%), a very small fraction of the population has college education (2%) and, also a relatively small proportion has

Table 3: Mean Comparisons between Non-Audited and Audited Municipalities

	(1) Non-Audited Municipalities	(2) Audited Municipalities	(3) Difference (2) - (1)
Population 2000	24072.5	25587.9	-1515.4
	(34813.4)	(44670.8)	[2747.7]
Population Density	86.54	78.19	8.346
	(443.5)	(289.1)	[25.26]
Share Female	0.494	0.494	-0.000660
	(0.0141)	(0.0134)	[0.0014]
Share Urban	0.612	0.602	0.0105
	(0.226)	(0.216)	[0.0157]
Share Youth (18-24)	0.131	0.133	-0.00214*
	(0.0128)	(0.0129)	[0.0010]
Income per Capita (log)	5.699	5.611	0.0875
	(0.546)	(0.517)	[0.0483]
Gini Index	0.552	0.553	-0.000546
	(0.0654)	(0.0586)	[0.0051]
Human Development Index	0.527	0.509	0.0178
	(0.0971)	(0.0968)	[0.0098]
Share Poor	0.403	0.438	-0.0351
	(0.211)	(0.214)	[0.0210]
Number of Highways	2.205	2.206	-0.00105
	(1.297)	(1.298)	[0.0517]
Share College Education	0.0238	0.0208	0.00298
	(0.0221)	(0.0191)	[0.0017]
Share Secondary Education	0.133	0.125	0.00713
	(0.0657)	(0.0637)	[0.0058]
Has an Institution of Higher Education	0.398	0.349	0.0491
	(0.490)	(0.478)	[0.0369]
Has AM Radio Station	0.270	0.248	0.0227
	(0.444)	(0.433)	[0.0319]
Has FM Radio Station	0.568	0.564	0.00424
	(0.496)	(0.497)	[0.0356]
Has TV Station	0.137	0.115	0.0222
	(0.344)	(0.319)	[0.0251]
Average Cloudiness	5.489	5.559	-0.0702
	(0.522)	(0.548)	[0.0541]
Average Total Rain	115.6	113.8	1.799
	(31.57)	(37.35)	[4.9931]
Average Max. Temperature	29.36	29.73	-0.373
	(3.160)	(2.741)	[0.3046]
Average Min. Temperature	18.36	18.98	-0.621
	(3.014)	(2.942)	[0.3304]
Observations	1,154	218	

Notes: this table shows means and standard deviations (in parenthesis) of several characteristics by municipalities audited and municipalities non-audited in the sample of municipalities for which there is information on traffic offenses. The difference in means is shown in column (3), with standard error in brackets. As described in the text, these characteristics were obtain from either the 2000 Census or a 2005 survey on municipalities characteristics. * $p < 0.05$

secondary education (around 13%). Additionally, in approximately 40% of the municipalities in my sample, there is at least one institution of higher education. In around 25% there is an AM radio station, in 56% there is an FM radio station, and in around 12% there is a TV station.

4 Empirical Strategy

Thanks to the randomization built into the anticorruption plan implemented in Brazil and the fact that in all municipalities audited between lotteries 22 to 38, there is at least one case of corruption per service order, the basic empirical strategy is straightforward. In this manner, the econometric specification chosen has the following form,

$$Y_{it} = \beta DisclosureCorruption_{it} + X_{it}\alpha + \mu_i + f_t + \epsilon_{it} \tag{1}$$

where Y_{it} is the log of the number of traffic offenses detected by the Federal Highway Patrol per inhabitant in municipality i at period t, $DisclosureCorruption_{it}$ is an indicator variable that takes the value 1 when there was a disclosure of a corruption case at the municipality level and zero otherwise, X_{it} is a vector of control variables (in particular, weather characteristics potentially correlated with traffic offenses), μ_i is a municipality fixed effect, f_t is a period fixed effect and ϵ_{it} an error term.

Notice that $DisclosureCorruption_{it}$ will take value 1 for a randomly selected municipality between period t^* (when the information about the audit was released) and the end of the sample period. It will take value zero for the periods before the disclosure of information of the municipality's audit, and for municipalities never audited, it will take value zero for the entire sample period. In this manner, regression model (1) can be interpreted as a difference-in-difference model. For the period analyzed and set of audits analyzed, $DisclosureCorruption_{it}$ is exogenous in model (1) thanks to the randomization of municipalities and the fact that in all municipalities audited during this period, at least one case of corruption was found. Therefore, β in the specification (1) can be interpreted as the causal effect of the disclosure of corruption on traffic offenses.

I am also interested in analyzing the effect of not only the disclosure of corruption but also the number of corruption cases disclosed on traffic offenses. For that matter, I propose to estimate the following regression model,

$$Y_{it} = \beta CorruptionCases_{it} + X_{it}\alpha + \mu_i + f_t + \epsilon_{it} \tag{2}$$

where the only difference relative to model (1) is that $CorruptionCases_{it}$ is the number of corruption cases disclosed. In this case, it is clear that the variable $CorruptionCases_{it}$ will not be exogenous in the model (2), even when the municipalities were randomly selected. In other words, among those municipalities audited, the number of corruption cases disclosed is not exogenous. It is reasonable to assume that differences in the number of corrupt cases found by the auditors can be associated with unobserved variables correlated with the level of corruption in the municipality and the compliance with laws by citizens. For instance, in municipalities where institutions have been established to increase law obedience or where citizens have strong support for the rule of law, it would be likely to find low levels of corruption along with low levels of traffic offenses. These unobserved characteristics can potentially vary across municipalities and over time. Therefore, to solve this endogeneity issue, I use the random selection of municipalities as an instrument for the number of corrupt cases disclosed.

Notice that $CorruptionCases_{it}$ will take value 0 for all municipalities never selected to be audited and for all municipalities selected to be audited before the release of the information contained in the audits. On the other hand, $CorruptionCases_{it}$ will be equal to the number of corruption cases disclosed by the audit for all municipalities selected to be audited after the release of the information of the audits.

A potential concern with either of these models is associated with citizen changing their behavior regarding the law in response to the municipality being selected to be audited. In other words, it could be the case that after knowing that the municipality where the individual lives was selected to be audited and before the report about the use of federal funds at the municipality level is disclosed, the individual changes his or her behavior because of a potential sensation of being monitored more than usual, and in this case by the federal government.

If this sensation of being monitored leads citizens to comply more with laws, but the disclosure of corruption cases leads citizens to comply less with laws, then the parameter of interest in model (1) would not identify. Basically, $DisclosureCorruption_{it}$ would be endogenous: it would be correlated with unobserved factors (sensation of being monitored) relevant to the number of infractions to the traffic law. Similarly, for the case of the model (2), this concern becomes a problem for the validity of the instrument. In other words, it could be the case that the exclusion restriction is violated in the model (2); in other words, the instrumental variable (selected to be audited) could have an effect on the dependent variable other than through the endogenous variable (number of corruption cases disclosed). Moreover, therefore, the instrumental variable strategy would lead me to a biased estimate of the effect of the number of corruption cases disclosed on citizen behavior regarding the law.

As it was mentioned above, there is a gap of time between the moment a municipality is selected and the moment in which the report about the use of federal funds is released (see Table 2). This gap of time is, on average, eight months. Therefore, we can use this time gap to test whether citizens respond to the fact that the municipality has been chosen to be audited even before the audit results are available. In order to do so, I propose performing an event study analysis. In different contexts, this type of regression model has been used by Autor (2003), Simon (2016), Clemens et al. (2018), Lutz (2011), Dobkin et al. (2018), among others.

Through this exercise, I would be able to know whether the results of estimating model (1) come from the response of the citizens to the municipality being audited or the corruption disclosed by the audits. On the other hand, in the case of the model (2), it is clear that this exercise is not a test of the validity of the instrumental variable strategy. If I found that citizens do not respond to the fact that the municipality has been chosen to be audited, then it would be reasonable to assume that the instrumental variable strategy is valid.

The event study analysis will not only allow me to analyze the behavior of traffic offenses in that window period of 8 months prior to the disclosure of corruption information but also allow me to analyze whether or not there are differential trends between treatment and control groups (audited and non-audited municipalities). Additionally, it is useful to test if there is any anticipation effect and to test if there is the persistence of the main effect over time (more generally, they can be used to analyze the dynamic behavior of the effect after the treatment under analysis took place). Formally, the event study regression model takes the following form,

$$Y_{it} = \sum_{j=-25}^{j=25} \beta_j DisclosureCorruption_{it}^{j} + X_{it}\alpha + \mu_i + f_t + \epsilon_{it} \tag{3}$$

where, Y_{it}, X_{it} μ_i, f_t and ϵ_{it} has the same interpretation as in model (1). Importantly, $DisclosureCorruption_{it}^{j}$ is an indicator variable for an event happening j periods away from t, defined as,

$$DisclosureCorruption_{it}^{j} = \begin{cases} 1[t \leq e_i + j] & \text{if} \quad j = -25 \\ 1[t = e_i + j] & \text{if} \quad -25 < j < 25 \\ 1[t \geq e_i + j] & \text{if} \quad j = 25 \end{cases}$$

where e_i is the time period in which the disclosure of corruption cases for municipality i took place. For instance, suppose that for municipality i the disclosure of corruption cases took place in January 2010. Then $DisclosureCorruption_{it}^{0}$ is an indicator variable that takes value one only in January 2010, $DisclosureCorruption_{it}^{1}$ takes value one only in February 2010, $DisclosureCorruption_{it}^{2}$ takes value one only in March 2010, and so on. Moreover, $DisclosureCorruption_{it}^{-1}$ takes the value one only for December 2009, $DisclosureCorruption_{it}^{-2}$ takes value one only for November 2009, and so on. Additionally, $DisclosureCorruption_{it}^{j}$ are binned at the endpoints, i.e. at 25 and -25. Therefore, $DisclosureCorruption_{it}^{25}$ takes value one 25 months or more after the month in which the disclosure of corruption took place. Similarly, $DisclosureCorruption_{it}^{-25}$

takes value one 25 months or more before the disclosure of corruption.

5 Results

5.1 The Effect of the Disclosure of Corruption

Table 4 shows the results of estimating regression model (1). In all columns, the dependent variable is the log of the number of infractions to the traffic law per capita. In order to avoid losing observations with zero infractions, I replace the log of the number of infractions for these observations with a zero and include an indicator variable that equals one when the number of infractions is zero. The results are not driven by this choice on how to deal with zeros in the dependent variable. In the Appendix (see Tables A1 and A2), I present the results of estimating specification (1) and (2) with the dependent variable as the number of infraction per capita. The results are very similar.

Moreover, all regressions include municipality fixed-effects and time-fixed effects. Also, standard errors are clustered at the state level.

According to Table 4, for all infractions to traffic law and not controlling for weather covariates, the estimated effect of the disclosure of corruption is a 1.2% increase in the number of traffic offenses per capita (column (1), Panel A). This is a small and not statistically different from zero estimate. When I disaggregate the information on traffic offenses into different types of infractions (columns (2) to (10), Panel A), the results are relatively similar. For three (illegal equipment, red-light crossing, and other infractions) out of nine types of infractions, the point estimate is negative but small and not significant. For the rest of the infraction types, the point estimate is positive and similar to the one found for all infractions; the point estimates suggest a small increase between 0.8% and 1.6% in traffic offenses per capita. All estimates are small in magnitude and not statistically different from zero.

Table 4: Effect of the Disclosure of Corruption on Traffic Offenses

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. No Weather Covariates.										
Disclosure of Corruption	0.0119	0.0158	0.0160	0.00814	0.00879	-0.0147	0.0115	0.0148	-0.00409	-0.00267
	(0.0390)	(0.0174)	(0.0264)	(0.0187)	(0.0165)	(0.0218)	(0.0291)	(0.0307)	(0.00239)	(0.0255)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964
Panel B. Weather Covariates Included.										
Disclosure of Corruption	0.0119	0.0158	0.0160	0.00807	0.00876	-0.0146	0.0115	0.0148	-0.00410	-0.00283
	(0.0390)	(0.0174)	(0.0264)	(0.0187)	(0.0165)	(0.0218)	(0.0291)	(0.0307)	(0.00239)	(0.0256)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964

Notes: this table shows the results of estimating regression model (1). In all cases, the dependent variable is the log of the number of traffic offenses per capita. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects and time fixed effects. Standard errors (in parenthesis) are clustered at the state level. Panel A shows the effect of interest without controlling for weather characteristics, while regressions in Panel B include weather characteristics as covariates (these are average cloudiness, average total rain, average maximum temperature and average minimum temperature).

In Panel B of Table 4, weather covariates are included mainly to reduce the unexplained variability in the dependent variable and, therefore, increase the precision in the estimation of the effect of interest. Notice, however, that in general the inclusion of this set of covariates does not modify the estimates in Panel A. In general, all estimates for the effect of the disclosure of corruption and their corresponding standard error are very similar or the same.

From the results presented in Table 4, I am not able to state that the disclosure of corruption cases has an impact at the municipality level on the number of traffic offenses. All estimates are small, close to zero and not statistically significant.

5.2 The Effect of the Number of Corruption Cases Disclosed

As mentioned previously, I am also interested in analyzing whether the number of corruption cases disclosed by the auditors of the Office of Comptroller-General has an impact on traffic offenses. To this end, I proceed to present the results of estimating regression model (2), in which the regressor of interest is the number of corruption cases disclosed.

In order to guarantee the validity of the instrumental variable strategy in terms of the relevance of the instrument, in Table 5 I present the results of the first stage estimation. As expected, the random selection of municipalities to be audited is highly correlated with the number of corruption cases disclosed. The F-statistics are 89.97 and 102.75, with no covariates and with covariates, respectively. I can clearly rule out any concern associated with having a weak instrument.

Notice that, as mentioned above, the number of corruption cases disclosed for municipalities never selected to be audited is zero. Moreover, for all municipalities selected to be audited, there was at least one case of corruption disclosed. In this manner, the estimated coefficient of the random selection of municipalities to be audited in Table 5 is basically the average number of corruption cases disclosed in audited municipalities.

First, Table 6 shows the results of estimating specification (2) by Ordinary Least Squares (OLS). As expected, even when the estimates are not statistically different from zero, there is in general, a positive correlation between the number of corruption cases disclosed and infractions to the traffic law. Based on the point estimation, for all types of infractions, one extra case of corruption disclosed leads to an increase of 1.9% in the number of infractions per capita. As explained previously, these results cannot be considered unbiased estimates of the effect of interest given the potential endogeneity issues associated with these regressions. As before, Panel A shows the results without covariates, and Panel B shows the results, including weather covariates in the regressions. There are no significant differences between the results in Panel A and Panel B. All point estimates and corresponding standard errors are very similar or the same.

Table 5: First Stage - Random Selection and Number of Corruption Cases

	(1)	(2)
Random Selection To be Audited	2.7303***	2.7303***
	(0.2883)	(0.2883)
Controls	No	Yes
Observations	127,333	127,333
R-squared	0.675	0.675
F-statistic	89.97	102.75

Notes: this table shows the first stage regression associated with estimating specification (2) by 2SLS. In both columns, the dependent variable is the number of corruption cases disclosed by the auditors of the Office of Comptroller-General. Column (1) no covariates are included, while in Column (2) weather variables are included as covariates. Standard errors (in parenthesis) are clustered at the state level. *** $p < 0.001$

Table 7 shows the results of estimating equation model (2) by 2SLS using the random selection of municipalities as an instrument for the number of corruption cases disclosed. As it can be seen from column (1) Panel A, the effect of an additional corruption case disclosed is an increase of 0.4% in traffic offenses per capita. This is a small and not statistically different from zero estimate. As in the regression model with disclosure of corruption as the key independent variable, the effect of the number of corruption cases is negative, small, and not statistically significant for illegal equipment, red-light crossing, and other infractions. For the rest of the types of infractions, the estimated effect of an additional corruption case disclosed is between 0.2% and 0.6%. All these estimates are small, close to zero, and not statistically significant. Additionally, results when controlling for weather characteristics are similar or the same (Table 7, Panel B).

Again, the results from Table 7 do not allow me to state that the number of corruption cases has an impact at the municipality level on the number of traffic offenses.

To further explore the potential role of the number of corruption cases on traffic offenses, I estimate a similar regression model as specification (2), but instead of using the number of corruption cases disclosed as the key independent variable, I use an indicator variable that equals one when the number of corruption cases disclosed is greater or equal than the median number of corruption cases found by the Office of the Comptroller-General in the audits performed associated with lotteries 22 to 38 (the median number is 2.46 corruption cases per service order). I define this indicator as the disclosure of high levels of corruption. As in model (2), this indicator variable is endogenous and has to be instrumented with the exogenous variable associated with the random selection of audits.

In Table 8, I show the results of the first stage associated with this alternative regression model. As expected, the random selection of municipalities is highly correlated with disclosing high levels of corruption. Additionally, given the structure of my data and the way in which the indicator for high levels of corruption was constructed, the point estimate is 0.5. Moreover, the F-statistics are large enough to conclude that the instrument has a strong first stage.

In Table 9, the results of estimating the effect of disclosing high levels of corruption on traffic offenses are presented. In Panel A the coefficients were estimated using OLS, while in Panel B using 2SLS, instrumenting the disclosure of high levels of corruption with the random selection of municipalities to be audited. Both panels show the results including weather covariates. As can be seen, the OLS estimates are in general larger than the 2SLS estimates. Clearly, this makes sense given the endogeneity issues discussed above (in particular, unobserved variables, such as the general respect for law in a given municipality, that are positively correlated with corruption and traffic offenses would tend to bias the OLS estimates upward). The 2SLS results indicate that disclosing high levels of corruption increases by 2.4% the number of all types of traffic offenses. As in the previous estimations, this result is small and not statistically different from zero. Also, the results for illegal equipment, red-light crossing, and other infractions indicate a negative effect of the disclosure of high levels of corruption. However, neither estimate is statistically different from zero. For the rest of the types of infractions, for which the effect is positive, the estimated impact of disclosing high levels of corruption range from 1.6% to 3.3%. Comparing these results with those displayed in Table 4 (in which I show the effect of disclosing any corruption case), it is clear that when I focus on the disclosure of high levels of corruption, the point estimates are larger suggesting a larger effect on traffic offenses. However, the point estimates from Table 4 are not statistically different from those in Table 9.

Overall, the evidence presented so far does not allow me to state or conclude that the disclosure of corruption impacts citizen compliance with traffic law. In all cases, the effects estimated are small, close to zero, and not statistically different from zero.

Table 6: The Effect of the Number of Corruption Cases Disclosed on Traffic Offenses (OLS Regressions)

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. No Weather Covariates.										
Corruption Cases	0.0190	0.0103	0.0155	0.0147	0.0116	-0.0256	0.0332	-0.00880	-0.00252	-0.000738
	(0.0421)	(0.0180)	(0.0281)	(0.0214)	(0.0153)	(0.0301)	(0.0351)	(0.0311)	(0.00211)	(0.0236)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964
Panel B. Weather Covariates Included.										
Corruption Cases	0.0192	0.0103	0.0155	0.0147	0.0116	-0.0255	0.0333	-0.00878	-0.00253	-0.000800
	(0.0421)	(0.0180)	(0.0281)	(0.0214)	(0.0153)	(0.0300)	(0.0351)	(0.0310)	(0.00211)	(0.0237)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964

Notes: this table shows the results of estimating regression model (2) by OLS. In all cases, the dependent variable is the log of the number of traffic offenses per capita. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects and time fixed effects. Standard errors (in parenthesis) are clustered at the state level. Panel A shows the effect of interest without controlling for weather characteristics, while regressions in Panel B include weather characteristics as covariates (these are average cloudiness, average total rain, average maximum temperature and average minimum temperature).

Table 7: The Effect of the Number of Corruption Cases Disclosed on Traffic Offenses (2SLS Regressions)

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. No Weather Covariates.										
Corruption Cases	0.00435	0.00580	0.00587	0.00298	0.00322	-0.00539	0.00420	0.00542	-0.00150	-0.000978
	(0.0143)	(0.00634)	(0.00951)	(0.00665)	(0.00591)	(0.00777)	(0.0107)	(0.0112)	(0.000954)	(0.00918)
Observations	127333	127333	127333	127333	127333	127333	127333	127333	127333	127333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964
Panel B. Weather Covariates Included.										
Corruption Cases	0.00437	0.00577	0.00588	0.00296	0.00321	-0.00535	0.00422	0.00542	-0.00150	-0.00103
	(0.0142)	(0.00635)	(0.00951)	(0.00665)	(0.00592)	(0.00778)	(0.0107)	(0.0112)	(0.000954)	(0.00920)
Observations	127333	127333	127333	127333	127333	127333	127333	127333	127333	127333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964

Notes: this table shows the results of estimating regression model (2) by 2SLS using the random selection of municipalities as an instrument for the number of corruption cases disclosed. In all cases, the dependent variable is the log of the number of traffic offenses per capita. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects and time fixed effects. Standard errors (in parenthesis) are clustered at the state level. Panel A shows the effect of interest without controlling for weather characteristics, while regressions in Panel B include weather characteristics as covariates (these are average cloudiness, average total rain, average maximum temperature and average minimum temperature).

Table 8: First Stage - Random Selection and High Levels of Corruption Cases

	(1)	(2)
Random Selection To be Audited	0.492***	0.492***
	(0.0390)	(0.0390)
Controls	No	Yes
Observations	127,333	127,333
R-squared	0.493	0.493
F-statistic	80.91	89.22

Notes: this table shows the first stage regression associated with estimating by 2SLS a regression model where the key independent variable is an indicator variable for disclosure of high corruption (number of corruption cases greater or equal to the median number of corruption cases disclosed in audits associated with lotteries 22 to 38). In both columns, the dependent variable is a dummy variable that equals one when the number of corruption cases is greater or equal to the median number of cases. In column (1) no covariates are included, while in Column (2) weather variables are included as covariates. Standard errors (in parenthesis) are clustered at the state level. *** $p < 0.001$

Table 9: Effect of the Disclosure of High Levels of Corruption on Traffic Offenses - OLS and 2SLS Estimation

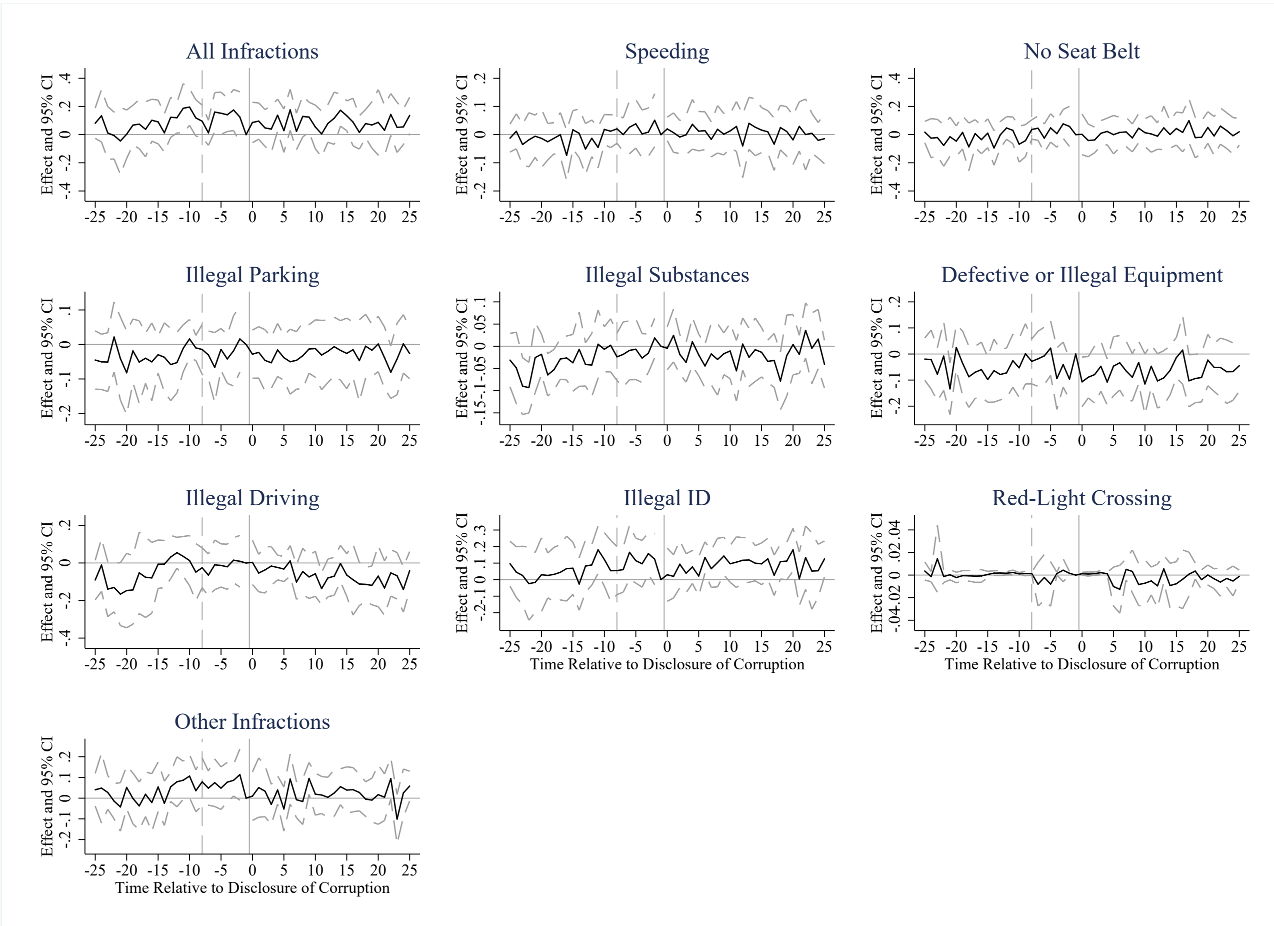
	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. OLS Estimates.										
Disclosure of High	0.0808	0.0270	0.0454	-0.00235	0.0254	-0.0211	0.0852	0.0293	-0.00298	0.0411
Corruption	(0.0683)	(0.0233)	(0.0455)	(0.0278)	(0.0304)	(0.0599)	(0.0612)	(0.0569)	(0.00412)	(0.0495)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964
Panel B. 2SLS Estimates.										
Disclosure of High	0.0242	0.0321	0.0326	0.0164	0.0178	-0.0297	0.0234	0.0301	-0.00833	-0.00574
Corruption	(0.0773)	(0.0352)	(0.0521)	(0.0366)	(0.0326)	(0.0446)	(0.0581)	(0.0606)	(0.00481)	(0.0511)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964

Notes: this table shows the results of estimating a regression model where the key independent variable is an indicator variable for disclosure of high corruption (number of corruption cases greater or equal to the median number of corruption cases disclosed in audits associated with lotteries 22 to 38). Panel A uses OLS and Panel B uses 2SLS, using as instrument the random selection of municipalities. In all cases, the dependent variable is the log of the number of traffic offenses per capita. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects, time fixed effects and weather controls. Standard errors (in parenthesis) are clustered at the state level.

5.3 Event Study Results

Figure 3 presents the results of estimating regression model (3). The solid black line represents the estimated β_j with a corresponding 95% confidence interval (grey dashed line). The horizontal axis represents time relative to the disclosure of corruption (the month in which the information about the audit was released). In the graphs, the disclosure of corruption takes place at point 0 of the horizontal axis.

Figure 3. Event Study.



Notes: This figure shows the results of estimating the event study regression model (3). The horizontal axis represents months relative to the disclosure of corruption, which takes place at month 0. The solid black line represents the coefficients β_j (for $j = -25, ..., 25$), and the grey dashed lines are 95% confidence intervals. The vertical dashed line represents the moment in which the municipalities were chosen to be audited. The vertical solid line divides the period between before and after the disclosure of corruption. As mentioned in the text, the endpoints (-25 and 25) are binned, meaning that β_{25} and β_{-25} represent the effect for 25 or more periods into the future and, 25 periods or more ago, relative to the disclosure of corruption.

First, consistent with the previous results, there is no sign of an impact of the disclosure of corruption in the months following the disclosure of corruption (i.e., to the right of the solid vertical line). The trend of the black lines remains flat after the disclosure, and all point estimates are not statistically different from zero. This is true for all infractions and when considering different types of traffic offenses.

Second, the pattern previous to the disclosure of corruption does not support the concern of anticipation effects or differences in trends between municipalities audited and municipalities non-audited. Before the disclosure of corruption, even when the graphs show some fluctuations, the trend of the black line for the different graphs is horizontal, meaning that there are no differences in trends between treated and control groups. This is an expected result given the randomization associated with the disclosure of corruption.

Finally, as discussed previously, a potential concern is that citizens changed their behavior regarding the law in response to the municipality being selected to be audited. The different graphs in Figure 3 rule out this possibility. As it can be seen, in the gap between a municipality was selected to be audited (vertical dashed line) and the moment the information of the audits was disclosed (vertical solid line) there is no change in trends relative to the period before the selection of municipalities to be audited. In all graphs, relative to the period that corresponds with more than eight months before the disclosure of corruption (in the graphs, this period corresponds to the months to the left of the vertical dashed line), the trend of the black line remains the same in the period ranging from 8 months prior to the disclosure of the information of the audits (the moment in which a municipality is randomly selected to be audited) and the month when the information is released. Therefore, we can rule out the possibility that citizens changed their behavior regarding laws when they knew that their municipality was chosen to be audited. Consequently, concerns about the validity of the empirical strategy based on changes in the behavior of citizens in municipalities chosen to be audited can be ruled out. Additionally, these results provide more support for the validity of the instrumental variable strategy. The random selection of municipalities only affects the behavior of citizens regarding traffic law through its effect on the corruption disclosed.

5.4 The Role of Media

Clearly, all the estimations performed so far are relevant if citizens learned about the content of the audits. Finan and Ferraz (2008) provide anecdotal evidence suggesting that the information from the audits reached citizens, and in particular, it was used during municipal campaigns. However, I do not have data that confirms that citizens in audited municipalities learned about the corruption cases disclosed by the Office of Comptroller-General.

One possible approach to solve this lack of information is to use the information about the availability of local radio stations and TV stations in each municipality. It is much more likely that citizens learn about the content of the audits in municipalities that have either a local AM radio station, a local FM radio station, or a local TV station. Municipalities that do not have either of these media sources have only access to national media, which is less likely to divulge information about specific cases of corruption at the municipality level. On the other hand, in those municipalities with local media sources, citizens are more likely to learn about different events, such as the disclosure of corruption in the specific municipality.

In light of the previous results about the effect of the disclosure of corruption, it is possible that the impact is concentrated in municipalities with either a local AM radio station, a local FM radio station, or a local TV station, where citizens are more likely to learn about the contents of the audits performed at the municipality level. In order to explore this possibility, I perform a regression analysis following specification (1) but including an interaction term between the variable that indicates the disclosure of corruption and an indicator variable that equals one if the municipality has either a local AM radio station, a local FM radio station or a local TV station and, zero otherwise.

Table 10 presents the results of this exercise. For all infractions, the point estimate suggests that the effect of disclosing corruption is greater in municipalities with local media sources compared to municipalities with no local media sources. Following the point estimates, if a municipality has a local media source (AM Radio, FM radio or TV station) the disclosure of corruption implies an increase of 3.1% in the number of infractions to the traffic law; see column (1). While if a municipality does not have local media sources, the effect is 0.95%. These estimates are again small in magnitude and not statistically different from zero. When infractions are disaggregated by type of traffic offense, the results are qualitatively similar; municipalities with local media sources show a greater response to the disclosure of corruption in terms of traffic violations than municipalities with no local media sources (except for the case of illegal equipment, infraction for which the presence of local media have a negative effect). However, the point estimates are always small, close to zero, and not statistically significant.

Even when the sign of the estimated coefficients in Table 10 is consistent with the idea that citizens in municipalities with local media sources are more likely to be informed about the contents of the audits and respond with a greater increase in the number of infractions to the traffic law, the magnitude of these coefficients and their statistical significance do not allow me to state that there is an effect of the disclosure of corruption on traffic offenses which is concentrated in municipalities with local media sources.

Table 10: Effect of the Disclosure of Corruption on Traffic Offenses and the Role of Media

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Corruption Disclosure	0.00945	0.0110	0.0145	0.0090	0.0081	-0.00750	0.00745	0.0131	-0.00441	-0.0041
	(0.0381)	(0.0169)	(0.0259)	(0.0197)	(0.0197)	(0.0262)	(0.0273)	(0.0342)	(0.00264)	(0.0271)
Corruption Disclosure $\times$	0.0213	0.0425	0.0137	0.0006	0.0007	-0.0632	0.0351	0.0151	0.00279	0.0024
Media	(0.119)	(0.0898)	(0.0957)	(0.0659)	(0.0489)	(0.0996)	(0.0996)	(0.0623)	(0.00227)	(0.0797)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.885	0.896	0.963	0.975	0.983	0.967	0.931	0.953	0.984	0.964

Notes: this table shows the results of estimating a regression model like (1) but including an interaction term between the disclosure of corruption and an indicator variable that takes value one if the municipality has either a local AM radio station, a local FM radio station or a local TV station and, zero otherwise. In all cases, the dependent variable is the log of the number of traffic offenses per capita. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects, time fixed effects and controls for weather characteristics. Standard errors (in parenthesis) are clustered at the state level.

6 Conclusions

In this paper, I empirically analyzed whether or not the disclosure of corruption cases impacts compliance with the law by citizens. In particular, using the information from an anticorruption plan implemented in Brazil and combining this data with information on traffic offenses, I tested if the disclosure of corruption affects the number of infractions to the traffic law.

The results show that the disclosure of corruption cases has a small, positive, and not statistically significant effect on the number of traffic offenses detected by the law enforcement agency at the municipality level. Moreover, when analyzing the effect of the number of corruption cases disclosed, I find that as the number of corruption cases disclosed increases, the number of infractions to the traffic law increases. However, the estimated effect is also small and not statistically significant. When I focus the analysis on the disclosure of high levels of corruption, the estimated effect seems to be larger than the effect of disclosing any corruption level. However, again the estimates are small and not statistically different from zero. A similar result is obtained when I analyze the differential effect for municipalities with media sources (where individuals are more likely to learn about the information of the audits): a larger, but not significant, effect for municipalities with media sources.

Finally, I present evidence through an event study regression model consistent with the previous results. This event study also provides evidence to support the empirical strategy: no anticipation effects, parallel pre-treatment trends, and no change in traffic offenses associated with the fact that the municipality has been chosen to be audited.

The disclosure of corruption may affect the behavior of citizens regarding the law through other activities (such as tax evasion). However, the results of my empirical analysis do not allow me to conclude that the disclosure of corruption cases affects compliance with the law by citizens.

References

Abbink, Klaus, Esteban Freidin, Lata Gangadharan, and Rodrigo Moro, "The effect of social norms on bribe offers," *The Journal of Law, Economics, and Organization*, 2018, *34* (3), 457–474.

Acemoglu, Daron and Matthew O Jackson, "History, expectations, and leadership in the evolution of social norms," *The Review of Economic Studies*, 2015, *82* (2), 423–456.

Ajzenman, Nicolás, "The power of example: Corruption spurs corruption," *American Economic Journal: Applied Economics*, 2021, *13* (2), 230–57.

Autor, David H, "Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing," *Journal of labor economics*, 2003, *21* (1), 1–42.

Avis, Eric, Claudio Ferraz, and Frederico Finan, "Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians," *Journal of Political Economy*, 2018, *126* (5), 1912–1964.

Bardhan, Pranab, "Corruption and Development: A Review ofthe Issues," *Journal of Economic Perspectives*, 1997, *7*, 3–45.

Barr, Abigail and Danila Serra, "Corruption and culture: An experimental analysis," *Journal of Public economics*, 2010, *94* (11-12), 862–869.

Cameron, Lisa, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan, "Propensities to engage in and punish corrupt behavior: Experimental evidence from Australia, India, Indonesia and Singapore," *Journal of public economics*, 2009, *93* (7-8), 843–851.

Clemens, Michael A, Ethan G Lewis, and Hannah M Postel, "Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion," *American Economic Review*, 2018, *108* (6), 1468–87.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J Notowidigdo, "The economic consequences of hospital admissions," *American Economic Review*, 2018, *108* (2), 308–52.

Ferraz, Claudio and Frederico Finan, "Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes," *The Quarterly journal of economics*, 2008, *123* (2), 703–745.

_ , _ , and Diana B Moreira, "Corrupting learning: Evidence from missing federal education funds in Brazil," *Journal of Public Economics*, 2012, *96* (9-10), 712–726.

Gächter, Simon and Jonathan F Schulz, "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 2016, *531* (7595), 496–499.

Hanna, Rema and Shing-Yi Wang, "Dishonesty and selection into public service: Evidence from India," *American Economic Journal: Economic Policy*, 2017, *9* (3), 262–90.

Lutz, Byron, "The end of court-ordered desegregation," *American Economic Journal: Economic Policy*, 2011, *3* (2), 130–68.

Monte, Alfredo Del and Erasmo Papagni, "Public expenditure, corruption, and economic growth: the case of Italy," *European journal of political economy*, 2001, *17* (1), 1–16.

Olken, Benjamin A and Rohini Pande, "Corruption in developing countries," *Annu. Rev. Econ.*, 2012, *4* (1), 479–509.

Orosz, Gábor, István Tóth-Király, Beáta Bőthe, B Paskuj, Mihály Berkics, Márta Fülöp, and Christine Roland-Lévy, "Linking cheating in school and corruption," *European Review of Applied Psychology*, 2018, *68* (2), 89–97.

Shleifer, Andrei and Robert W Vishny, "Corruption," *The quarterly journal of economics*, 1993, *108* (3), 599–617.

Simon, David, "Does early life exposure to cigarette smoke permanently harm childhood welfare? Evidence from cigarette tax hikes," *American Economic Journal: Applied Economics*, 2016, *8* (4), 128–59.

Svensson, Jakob, "Who must pay bribes and how much? Evidence from a cross section of firms," *The quarterly journal of economics*, 2003, *118* (1), 207–230.

7 Appendix

Tables A1 and A2 present the results of estimating specifications (1) and (2), respectively, but without taking logs to the number of infractions per capita. Results are very similar.

Table A1: Effect of the Disclosure of Corruption on Traffic Offenses - Linear Model

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. No Weather Covariates.										
Disclosure of Corruption	1.7662	0.5708	0.2491	0.0228	0.0238	-0.2475	0.7147	0.5910	-0.0002	-0.0883
	(5.2703)	(0.6085)	(0.3917)	(0.049)	(0.0424)	(0.3841)	(1.6809)	(1.1541)	(0.0001)	(1.1821)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.0109	0.00283	0.00659	0.00277	0.0169	0.00864	0.0129	0.0103	0.00107	0.00789
Semi-elasticity (%)	0.80	1.32	1.80	0.77	0.95	-1.32	0.91	1.60	-0.40	-0.38
	(1.61)	(2.26)	(3.61)	(3.71)	(1.36)	(1.36)	(1.82)	(3.11)	(0.80)	(0.88)
Panel B. Weather Covariates Included.										
Disclosure of Corruption	1.7660	0.5708	0.2492	0.0228	0.0238	-0.2473	0.7147	0.5910	-0.0002	-0.0883
	(5.2701)	(0.6085)	(0.3915)	(0.049)	(0.0424)	(0.3844)	(1.6809)	(1.1541)	(0.0001)	(1.1821)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.0112	0.00288	0.00709	0.00280	0.0170	0.00905	0.0132	0.0104	0.00109	0.00802
Semi-elasticity (%)	0.81	1.32	1.80	0.77	0.95	-1.32	0.91	1.55	-0.40	-0.38
	(1.63)	(2.26)	(3.61)	(3.71)	(1.36)	(1.34)	(1.80)	(3.12)	(0.81)	(0.88)

Notes: this table shows the results of estimating regression model (1) where the dependent variable is the number of traffic offenses per 100,000 inhabitants. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects and time fixed effects. Standard errors (in parenthesis) are clustered at the state level. Panel A shows the effect of interest without controlling for weather characteristics, while regressions in Panel B include weather characteristics as covariates (these are average cloudiness, average total rain, average maximum temperature and average minimum temperature). The semi-elasticity is computed as (Estimate/100,000)/(Mean numbers of infractions per capita)$\times$100. In this manner, these estimates of semi-elasticities are comparable to the ones presented in Table 4.

Table A2: The Effect of the Number of Corruption Cases Disclosed on Traffic Offenses (2SLS Regressions) - Linear Model

	(1) All Infractions	(2) Speeding	(3) No Seat Belt	(4) Illegal Parking	(5) Illegal Substance Usage	(6) Illegal Equipment	(7) Illegal Driving	(8) Illegal ID	(9) Red-Light Crossing	(10) Other Infractions
Panel A. No Weather Covariates.										
Corruption Cases	1.12598	0.21188	0.0830	0.01243	0.01029	-0.12000	0.43197	0.18302	-0.00004	-0.02048
	(2.12139)	(0.43681)	(0.13395)	(0.03916)	(0.01963)	(0.22069)	(1.12079)	(0.40602)	(0.00005)	(0.04753)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.0110	0.00278	0.00657	0.00277	0.0169	0.00872	0.0129	0.0102	0.00104	0.00788
Semi-elasticity (%)	0.51	0.49	0.60	0.42	0.41	-0.64	0.55	0.48	-0.09	-0.09
	(0.96)	(1.10)	(0.97)	(1.13)	(0.78)	(1.12)	(1.43)	(1.06)	(0.10)	(0.21)
Panel B. Weather Covariates Included.										
Corruption Cases	1.12598	0.20755	0.08581	0.01214	0.01004	-0.11813	0.43197	0.18302	-0.00004	-0.02048
	(2.24379)	(0.52698)	(0.13368)	(0.01334)	(0.01200)	(0.24823)	(1.03725)	(0.35153)	(0.00008)	(0.04011)
Observations	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333	127,333
R-squared	0.0112	0.00283	0.00707	0.00280	0.0171	0.00913	0.0131	0.0104	0.00106	0.00801
Semi-elasticity (%)	0.51	0.48	0.62	0.41	0.40	-0.63	0.55	0.48	-0.09	-0.09
	(1.02)	(1.22)	(0.97)	(0.45)	(0.48)	(1.32)	(1.32)	(0.92)	(0.15)	(0.18)

Notes: this table shows the results of estimating regression model (2) by 2SLS using the random selection of municipalities as an instrument where the dependent variable is the number of traffic offenses per 100,000 inhabitants. In each column specific type of traffic offenses are considered. All regressions include municipality fixed effects and time fixed effects. Standard errors (in parenthesis) are clustered at the state level. Panel A shows the effect of interest without controlling for weather characteristics, while regressions in Panel B include weather characteristics as covariates (these are average cloudiness, average total rain, average maximum temperature and average minimum temperature). The semi-elasticity is computed as (Estimate/100,000)/(Mean numbers of infractions per capita)$\times$100. In this manner, these estimates of semi-elasticities are comparable to the ones presented in Table 4.