Corruption and Collaborations: The Brazilian Example *

Lucas Alves Chacha

June 14, 2022

Abstract

In August 2013, Brazilian legislators enacted two important anti-bribery laws granting sanction reductions to collaborators. After this, in 2014 the Brazilian prosecutors initiated a series of big investigations unveiling big corruption schemes between the Brazilian government and important private corporations. In these cases, the prosecution strategy to investigate the defendants relied heavily on the offenders' disclosures under the newly enacted laws. In this context, one can argue that the investigations are evidence of the success of this anti-corruption policy. However, the impact of the policy over the widespread corruption in the country is not clear. This paper uses the empirical strategy from Miller (2009) to empirically test the effectiveness of the Brazilian policy against corruption. Results show that the policy was effective to both prosecute and deter corruption in Brazil.

1 Introduction

The Brazilian laws No. 12.846/13 and No. 12.850/13 were enacted in August 2013 as a response to the Brazilian protests of June 2013. Protesters demanded (among other things) stronger measures against corruption from the government. The laws were inspired by the guidelines of the United Nations Convention Against Corruption (UNCAC)[1]. They introduced the possibility for individuals and corporations involved in corruption to disclose their activities in exchange for sanction reductions. Notably, they also provide other measures to inhibit corruption besides the self-reporting mechanisms. In this sense, the law No. 12.846/13 (Anti-Corruption Law) also sets strict liability for firms on corruption crimes conducted by its employees. It institutes compliance measures

*This work was supported by the small grants scheme of the research initiative on Rebalancing Public and Private Interests and Erasmus Center of Empirical Legal Studies of Erasmus School of Law and the sector plan for law funding of the Ministry of Education, Culture and Research.

[1]Brazil signed the Agreement of the Convention on the 9th December 2003 and ratified on 15th June 2005. However, until 2013 little enforcement was observed.

as possible punishments and states other sanctions in case of corporate corruption. Furthermore, the law No. 12.850/13 (Law against organized crimes) also instituted new types of investigation procedures for Brazilian authorities. All these provisions have features that may somehow deter corruption.

After the enactment of the laws, in 2014 and later years a series of big corruption investigations were observed nationwide. Given this context, are the recent big corruption investigations an output of widespread corruption or an effective shift of the prosecution standards? More specifically, did the anti-corruption policy from 2013 help to decrease domestic bribery in Brazil?

The answers from these questions are not clear. Most of the theoretical studies on the subject are in antitrust literature. They show that on one hand the sanction reduction policy decreases the expected fines and might incentivise crimes. On the other hand, the sanction reduction scheme might induce a 'prisoner's dilemma' situation in which criminals would report their misconduct (Buccirossi and Spagnolo, 2001; Motta and Polo, 2003; Buccirossi and Spagnolo, 2006; Spagnolo, 2005; Harrington, 2008; Chen and Rey, 2013; Dufwenberg and Spagnolo, 2014). In this last scenario, sanction reduction policies are effective not only in deterring the crime, but also by making prosecution more efficient (Landes, 1971; Grossman and Katz, 1983; Kaplow and Shavell, 1994; Franzoni, 1999; Mungan and Klick, 2016). Therefore, although the policies may help to deter and prosecute corruption, they could backfire if poorly applied.

There is some empirical evidence that sanction reductions could help deter crimes. Most evidence comes from experiments (Bigoni et al., 2015; Engel et al., 2016; Abbink and Wu, 2017). However, in the antitrust literature the studies from Brenner (2009) and Miller (2009) use the number of cartel detections in Europe and the US respectively to test the efficacy of sanction reduction (leniency policies) over criminal activities. The basic idea is to extract a sign from the noisy data of criminal detections. Notably, if the data is consistent with any enhanced detection, the number of detected crimes should increase immediately after the introduction of the policies. Furthermore, if the data is consistent with long term deterrence, there should be a decrease in the number of detected crimes which should be below the previous number of detections. Berlin et al. (2018) tried this approach to test the effect of Chinese anti-corruption policies. However, their work found little evidence to support the claim. This paper uses the same methodology and finds that there is evidence that the anti-bribery policy put in place was effective against corruption.

The work here is divided as follows, in Section 2 the most important aspects from corruption and the ways to measure it are discussed. Following, in Section 3 the approach from Miller (2009) which is going to be used here is better explained. Next, in Section 4 the particular Brazilian institutional background is explained so that in Section 5 the relevant data is presented and briefly explored. Section 6 introduces the empirical strategy so in Section 7 the results are presented. Lastly, the conclusion and final remarks are drawn in Section 8.

2 Brief Contextualization

Perhaps the most challenging aspect of approaching corruption empirically is that corruption is highly unobservable. Individual preferences or willingness to pay bribes occurs only inside people's minds. Nonetheless, even when agents decide to pay a bribe, there is no good record of the aggregate transactions. In other words there is no good measure about the quantity of bribes in the economy, or even the total value of bribes paid in a certain period. In order to measure corruption levels, one has to rely on evidence of bribery.

The most popular way to measure corruption is by using the perception indexes (CPIs). A series of empirical studies using CPIs successfully explored cross-national relations between economic variables and corruption (Treisman, 2007; Rose-Ackerman, 2006). However, this approach seems to be biased whenever a crackdown in corruption happens. Thus, when corruption is largely detected, people tend to perceive the corruption as higher, since means of communication spread the knowledge about the events[2]. Therefore, CPIs may not be good assessments for corruption in a particular region over time after a policy shock. The Figure 1 shows the evolution of the Transparency International Corruption Perception Index[3] for Brazil.

Figure 1: Transparency International Corruption Perception Index for Brazil

The Figure 1 shows that the perception about corruption in Brazil increased since the anti-corruption shock. This statement goes against the study's hypothesis. Nonetheless, in another strand of literature, authors look for objective observable evidence of changes in corruption. To cite a few, Golden and Picci (2005) derived their index of corruption from differences between the expected stock of infrastructure and the observed ones, Di Tella and Schargrodsky (2003) from differences in prices of homogeneous goods in public contracts and Ferraz

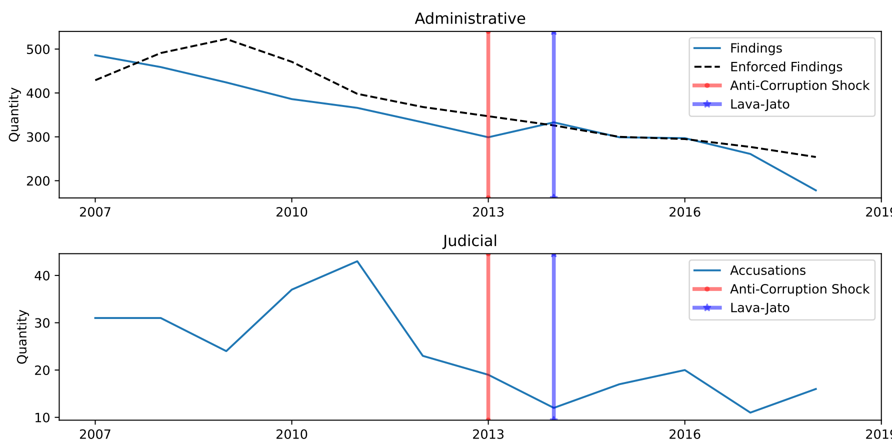
[2] There are methodological attempts to minimize this kind of bias. However, it is still the case that the indexes are fundamentally measuring perceptions (Transparency International, 2019).

[3] 100 = Completely clean; 0 = Completely Corrupt.

and Finan (2011) and Olken (2007) looked at the number of irregularities on audited public contracts.

The Figure 2 shows the recent series of parameters used by Ferraz and Finan (2011)[4]. They show that, at least since 2011, their corruption indicator is declining in Brazil, contradicting the conclusions from the CPI. It can be concluded from this example that measures of corruption can be misleading without a proper context.

Figure 2: Irregularity findings by the Brazilian External Audit Agency (TCU)



This work attempts to assess corruption from observable detection of the briberies. Moreover, the methodology proposed here derives from the antitrust literature, which tries to measure the unobservable cartel formation in the economy using crime detection (Miller, 2009; Brenner, 2009).

3 Theoretical Framework

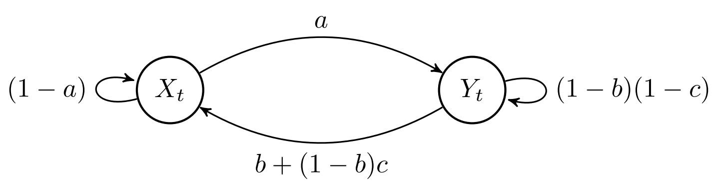
In natural sciences, whenever researchers need to infer something about an unobservable phenomenon, they often rely on matching a predicted sign to a noisy data. Once again, corruption is not observable. Nonetheless, some outputs from bribery may be observed, such as the number of inquiries and convictions, for instance. In this sense, a theory about the mechanics of bribery is needed to predict the sign. Or else, how the observed variables should behave if something occurs on the unobservable data.

Miller (2009) proposed one way to test empirically if an antitrust policy can deter cartel formation. He innovates by constructing a model showing the mechanism of transition from collusion to non-collusion for cartels. The model

[4] The data was obtained from the Annual Activities Report from the the Brazilian external audit agency (TCU), available at: `https://portal.tcu.gov.br/transparencia/relatorios/relatorios-de-atividades/relatorios-de-atividades.html`

consists of a two-state first order Markov process. Either they collude X_t or they do not collude Y_t. A firm can transit from not colluding to collude (start colluding) with a probability a, and can be detected with probability b or simply desist from collusion with probability c. Figure 3 shows the Markov transition diagram from Miller's model[5].

Figure 3: Markov Transition Diagram of Miller (2009)



The diagram shows the mechanism in which the firms change from colluding to not colluding. This is effective in showing how the observable detection of the crime may change in time. However, the model does not care for understanding the incentive mechanism of collusion[6]. In this sense, it is a good model for orientating the empirical strategy yet less effective to explain the impact of the proposed enforcement changes.

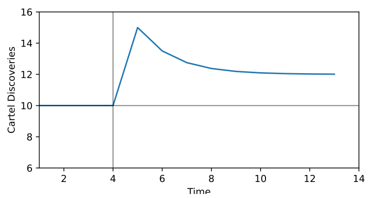
In his empirical approach, the criminal level is measured by observing the number of cartels detected by an Antitrust Authority (AA) before and after a policy shock. If a certain anti-crime enforcement is successful, it is expected that crime detection rises (spike) in the first moment due to enhanced detection efficiency. Later, after a while it drops to levels lower than before the enforcement shock (deterrence effect). Figure 4[7], exemplifies the expected path of the variable.

[5] In the second Part of my thesis, I model the decision to bribe. From the preliminary results, one can expect a similar mechanism of corruption detection.

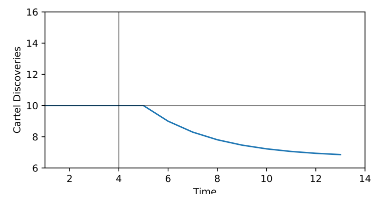
[6] There is an extensive literature in this field, see (Harrington, 2008; Spagnolo, 2005; Aubert et al., 2006; Motta and Polo, 2003). Most of them inspired the theoretical model in Part II of my thesis

[7] The path of the detected cartels on time on Figure 4 is given using arbitrary probabilities of collusion, detection and desistance.

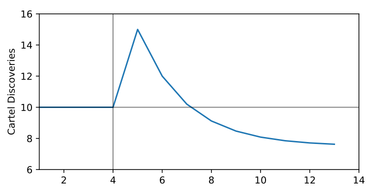
Figure 4: The Expected Number of Cartel Discoveries by Period (Extracted from Miller (2009))



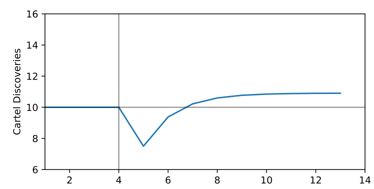
(a) Enhanced Detection



(b) Enhanced Deterrence



(c) Enhanced Detection Deterrence



(d) Weaker Detection and Deterrence

Notably, this study is not interested in cartel activity. However, the same methodology can be used to analyse corruption crimes. In this sense, Berlin et al. (2018) used Miller's model along with judiciary variables of detection of crimes of bribery in China. They found deterrence but no spikes on detection. By the same logic, this methodology could be used along with Brazilian data of corruption detection (the candidate variables are discussed below).

4 Institutional Background

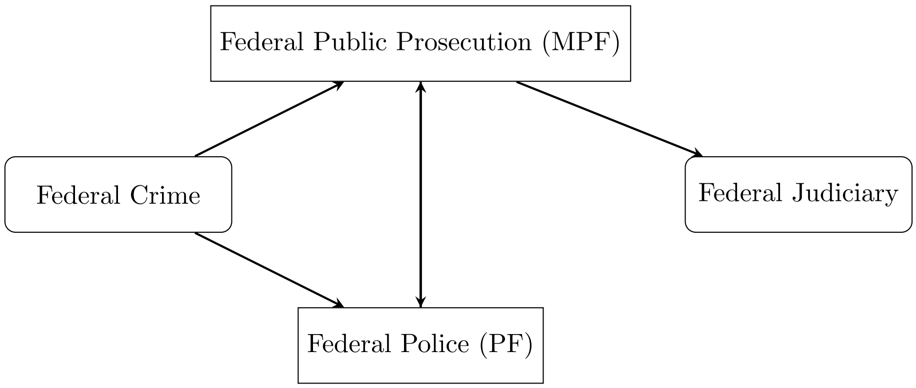
In order to apply Miller's method to the case of corruption in Brazil, it is first necessary to understand the country's processual criminal procedures.

Following Federal Law No. 3.689/41, both the Brazilian General Attorney's Office (Public Prosecution) and the Federal Police alone are **competent** to investigate corruption crimes of public interest. However, the standard procedure (*de facto*) is for both institutions to carry out investigations together. Normally, the corruption investigation starts at the Federal Police (Art. 5, I), which concludes it and delivers it to the Public Prosecution to proceed with the accusation (Art. 24). However, the investigation can start at the Prosecution and then be forwarded to the police to conduct parallel investigations conjointly (Art. 5, II). Only in exceptional cases the Prosecution carries out one investigation only by itself[8].

The Diagram in Figure 5 shows the flux of process in the Brazilian system.

[8]It was only in 2015 the Brazilian Supreme Court formalized the understanding that the Brazilian Prosecution could investigate a case alone ((HC) 89837/STF). Before this decision,

Figure 5: Brazilian Criminal Investigation Procedures Flowchart



It is reasonable to conclude that the Federal Police is generally the first to detect potential corruption crimes in Brazil. Even though the Public Prosecution can sometimes also be the first to detect. However, in the standard investigation cases, the prosecution will only know about an eventual corruption crime after the police finishes its investigations.

5 The Data

The data used here was downloaded from the online processual search engine of the Brazilian Public prosecution (MPF)[9]. It consists of 885.675 investigations conducted by the Brazilian Federal Police, sorted by date of the beginning of the investigation from January 2009 to January 2020.

It is possible to find data from before 2009 and after 2020 in the database. However, the periods before 2009 can be biased for minor states which did not use the informational system at the time. Moreover, the period from 2020 onwards was excluded from the list due to the covid-19 health crisis. Because of the exceptional use of emergency procurements and contracts, which lead to a rent-seeking supply shock. This could have altered the number of denounces and investigations on corruption in Brazil during the crisis.

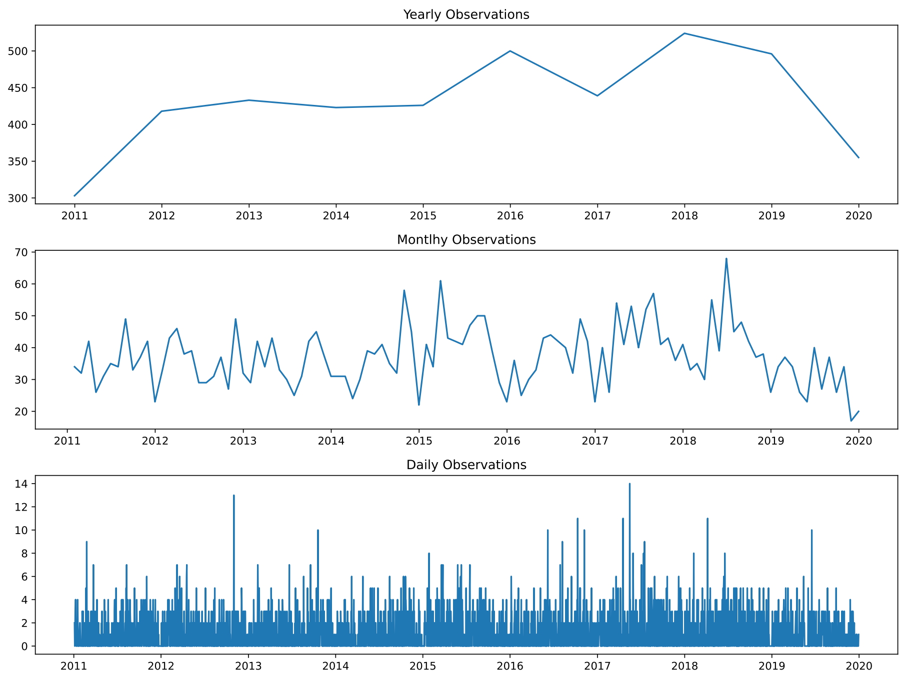
Figure 6 shows the number of inquiries regarding crimes of passive and active corruption[10] from 2010 to the end of 2019.

criminal investigation was a Police monopoly.

[9]Available at: http://apps.mpf.mp.br/aptusmpf/portal?servidor=portal and downloaded between February and March of 2020. The code I created for scraping the data is available at: https://github.com/caxaxa/MPF_Crime_Scraper.

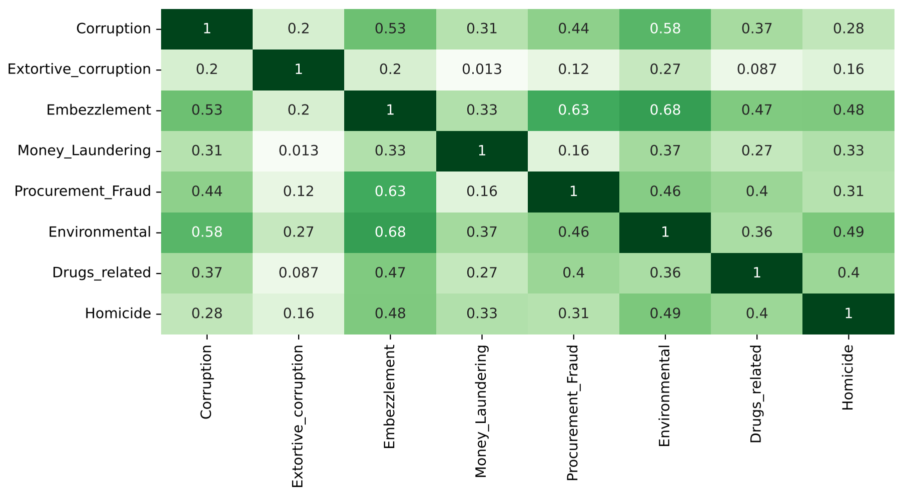
[10]Articles 317 and 333 from the Brazilian Penal Code, Federal Decree-Law No. 2.848/40.

Figure 6: Corruption Inquiries by Starting Dates



The database also provides the crime, and the state in which the investigation is conducted. The Tables **??** and **??** show some aggregate values from the dataset, also Figure 7 shows the correlation between multiple different crimes in Brazil.

Figure 7: Crime Correlation Matrix



From Figure 7 it is possible to observe some level of correlation between crimes. Mainly between embezzlement and environmental crimes. The reason might be because corruption occurs when those crimes happen. Or, because the law enforcement at large responds to exogenous other variables. Therefore, this paper also controls for a set of external variables. The chosen ones are the quarterly Brazilian real GDP growth[11], the monthly Brazilian unemployment rate[12], the monthly average of the main national treasury bonds real interest rate (SELIC)[13] and all the other crime investigations except the corruption crimes.

6 Empirical Strategy

The main empirical strategy resembles the works from Miller (2009), Brenner (2009) and Berlin et al. (2018). The first two based on cartel detections, but the last one focused on corruption, precisely the objective in this work. The authors' approaches can be adapted to this case and formalized as follows:

[11] Source Instituto de Pesquisa Economica Aplicada (IPEA), available at: http://ipeadata.gov.br.

[12] Before 2012, the source is Fundacao Sistema Estadual de Analise de Dados (SEAD), available at: https://www.seade.gov.br/. After 2012 the data is from Instituto Brasileiro de Geografia e Estatistica (IBGE), avalable at: https://www.ibge.gov.br/estatisticas/sociais/educacao/9127-pesquisa-nacional-por-amostra-de-domicilios.html?=&t=series-historicas. The data was corrected using the global monthly weighted average.

[13] The nominal rate was gathered from the Brazilian Central Bank, available at: https://www.bcb.gov.br/estatisticas/txjuros, and deflated by the IPCA inflation index, available at: http://ipeadata.gov.br.

$$Y_t = \beta_0 + \beta_1 D_t + \beta_2 T1_t^n + \beta_3 T2_t^n + \beta_4 X_t + \varepsilon_t \quad (1)$$

for,
t= Observation period (month or day); and
n= Order of the polynomial.

Where
Y_t= Number of new corruption inquiries. Or, detected corruption crimes;
D_t= Dummy for the impact of the enforcements, being 0 before August 2013 and 1 after;
$T1$= Time effect of all sample. Being 1 at the first observation 2 at the next and so on;
$T2$= Time effect from the beginning of the policy shock. Being 1 after august 2013, 2 in the next month and so on; and
X_t = Vector of control variables
ε_t= Is the error term from functional predictions, it is expected to be normally distributed and i.i.d.

In order to test different shaped curves that might fit the expected detection curve, the matrices $T1$ and $T2$ assume values of different order polynomials to check for distinct goodness of fit. Consequently, if the regression shows a polynomial with a spike after the policy shock followed by a decrease on the estimated mean, this may be evidence of an effective policy, both in terms of detection and deterrence.

Lastly, Berlin et al. (2018) also uses a Quandt-Likelihood Ratio (QLR) to detect arbitrary breaks in the series. The test consists in testing Chow-like breaks in every point of the curve to check for placebo effects. Consequently, this test might show if there is another interesting breakpointing in time that may be important for the data. Perhaps this test can be effective for determining an eventual change in the trend after the policy shock.

7 Preliminary Results

Regression Without Controls

7.0.1 Poisson Regression

The first set of regressions show how different order polynomial fit the data of crimes of corruption[14]. The Table 1 shows the regressions (1) to (6), containing the coefficients for the intercept and the policy dummy for different order polynomials.

[14] Here the chosen method uses global polynomials, i.e. there is a functional polynomial that predicts the entire series. Most of the time these functions tend to be very unstable. Therefore, one possible improvement is to try locally weighted scatter-plot smoothers (LOWESS) in future regressions.

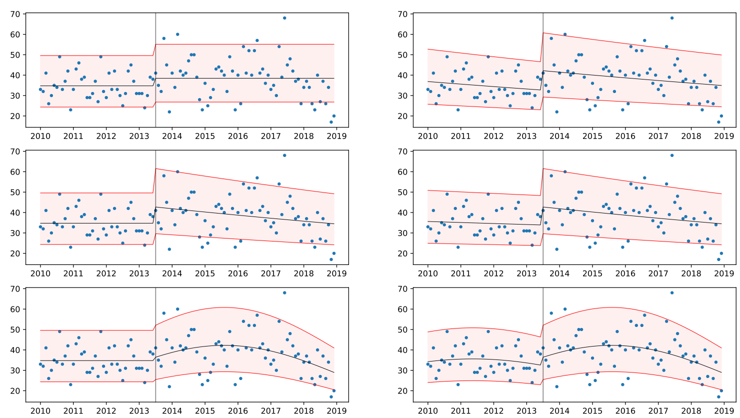
Table 1: Poisson Regression without Controls

	Dependent variable: Number of New Corruption Inquiries					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.549***	3.607***	3.549***	3.573***	3.549***	3.533***
	(0.026)	(0.032)	(0.026)	(0.051)	(0.026)	(0.075)
Dummy	0.101***	0.256***	0.206***	0.232***	0.047	0.119
	(0.033)	(0.060)	(0.046)	(0.066)	(0.064)	(0.102)
Full Seri Polynomial	None	1st Order	None	1st Order	None	2nd Order
Post-policy Polynomial	None	None	1st Order	1st Order	2nd Order	2nd Order
Observations	108	108	108	108	108	108
Pseudo R^2	0.011	0.0229	0.0235	0.0238	0.0408	0.0418
Residual Std. Error	1.000(df = 106)	1.000(df = 105)	1.000(df = 105)	1.000(df = 104)	1.000(df = 104)	1.000(df = 102)
F Statistic	(df = 1.0; 106.0)	(df = 2.0; 105.0)	(df = 2.0; 105.0)	(df = 3.0; 104.0)	(df = 3.0; 104.0)	(df = 5.0; 102.0)

Note: *p<0.1; **p<0.05; ***p<0.01

The results show that there is a consistent upward shift after the policy shock. This is given by the consistently positive and significant value of the Dummy variable. It means that data is consistent with a spike in detections after the policy. Therefore, consistent with an increase in detections of corruption. The mean predicted values from the last periods are lower than the pre-policy level for models (3), (5) and (6). However, they are not statistically significant. Nonetheless, the downward trend is significant. Therefore, from the models above, it is not possible to infer that there was a decrease in the overall Brazilian corruption. Or else, that the policy deterred criminals. Figure 8 shows the predicted values and a 95% confidence interval from mean predicted parameters.

Figure 8: New Corruption Inquiries Poisson Regression

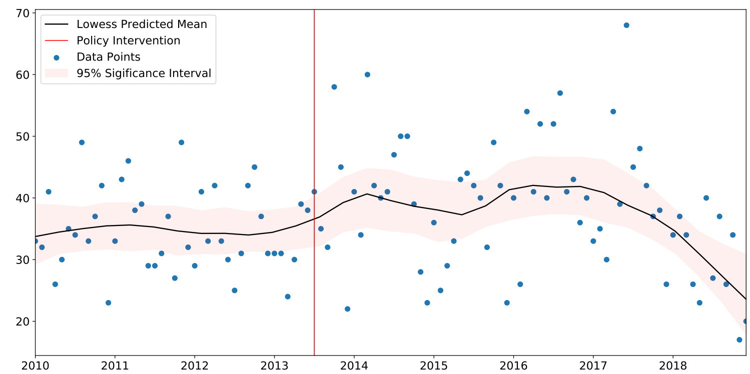


The result above is true for other regression methods such as OLS and Negative Binomial regressions. Notably, Binomial Regressions have more significant coefficients as shown in Table 5. Furthermore, OLS regressions allow the model to be regressed with higher order polynomials. This was not possible with Poisson and Negative Binomials. Parameters seem to be perfectly correlated for higher order parameters. This might imply that OLS models could be overfitting the data.

7.0.2 Locally Weighted Scatterplot Smoothing

Fitting a polynomial curve over noisy data has some known problems. Notably the predicted function misbehaves as it predicts points outside the data. Therefore, it might be overfitting the data, or, 'chasing noise'. Recently, as the computational power increases, methods of local regressions are taking place. Notably, locally weighted scatterplot smoothing from Cleveland (1979) is a powerful method. The Figure 9 shows the predicted LOWESS curve over the data of the Brazilian corruption inquiries.

Figure 9: New Corruption Inquiries LOWESS



Note that, differently from the past examples, the mean predicted values at the end of the whole period have statistically significant differences from the observations on the pre-policy periods. Therefore, it is evidence of enhanced deterrence after the policy shock. The average mean predicted values for the pre-policy period is 34.71 new cases per month. At the end of the period, the value is down to 23.57. A 95% confidence level drop of 32.11%. Conversely, there are no statistically significant peaks. Or else, none of the higher points is statistically different from the points on the pre-policy period[15].

Regression With Controls

Crime literature shows that potential criminals are influenced by the environment Becker (1968). Notably, since this work uses macro data, it may be affected by other macroeconomic variables. First, the regression is controlled for unemployment.This variable should affect newly detected corruption crimes by the supply side, as rent and opportunities determine the willingness to perform crimes (Becker, 1968; Mauro, 1995). Secondly, GDP is a proxy for the government budget. Therefore, bigger GDP would lead to more public contracts, corruption opportunities and to a better equipped public prosecution. Lastly, as corruption is a rent seeking activity, it should be bigger as real interests fall. The Table 2 show the results from Table 1 controlled for the above mentioned variables.

Results are a bit more significant and more prevalent. Now models (3), (4), (5) and (6) show a lower predicted value at the end of the period. Yet they are still not significant. Nonetheless, the dummy variable is now positive and

[15]The results using daily data are similar, as Figure 11

Table 2: Poisson Regression with Controls

	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.705***	3.619***	3.255***	3.213***	3.721***	3.726***
	(0.108)	(0.109)	(0.132)	(0.158)	(0.243)	(0.300)
Dummy	0.137**	0.378***	0.350***	0.331***	0.243***	0.298**
	(0.068)	(0.086)	(0.076)	(0.086)	(0.090)	(0.116)
Unemployment	-0.012	0.014	0.041***	0.043***	-0.022	-0.026
	(0.012)	(0.013)	(0.015)	(0.016)	(0.032)	(0.034)
Real Interest	-6.941**	-9.051***	-7.960**	-7.686**	-9.928***	-9.897***
	(3.352)	(3.397)	(3.375)	(3.421)	(3.492)	(3.632)
GDP	-0.181	0.575	2.029**	2.170**	2.613**	2.650**
	(0.915)	(0.926)	(0.988)	(1.029)	(1.017)	(1.050)
Full Seri Polynomial	None	1st Order	None	1st Order	None	2nd Order
Post-policy Polynomial	None	None	1st Order	1st Order	2nd Order	2nd Order
Observations	108	108	108	108	108	108
R^2						
Adjusted R^2						
Residual Std. Error	1.000(df = 103)	1.000(df = 102)	1.000(df = 102)	1.000(df = 101)	1.000(df = 101)	1.000(df = 99)
F Statistic	(df = 4.0; 103.0)	(df = 5.0; 102.0)	(df = 5.0; 102.0)	(df = 6.0; 101.0)	(df = 6.0; 101.0)	(df = 8.0; 99.0)
Note:						*p<0.1; **p<0.05; ***p<0.01

significant for all model, reinforcing the evidence for an enhanced detection after the policy.

As a small robustness test. The same models are regressed for one year lagged control variables. The idea here is that crimes are affected by the current environment, but are only detected in the future. Therefore, the detections today, or the quantity of detectable crimes today, depend on the environment in the past. As Figure 6 shows, there is little difference in the results.

8 Final Remarks

This paper reproduced the methodology present in Miller (2009), Brenner (2009) and Berlin et al. (2018) to test the efficacy of sanction reduction interventions over crimes of corruption in Brazil. The paper uses the data of the number of monthly and daily inquiries on corruption from 2010 to 2020.

Results show that, for all Poisson regressions, there is a consistently significant positive predicted value for a dummy after the policy intervention. Notably, this is evidence that there was an increase in detections after the policy.

Furthermore, using local weighted scatterplot smoothing, the data shows that predicted values at the end of the period are statistically lower than on the period before the intervention. This is evidence of increased deterrence.

In summary, the introduction of the sanction reduction policies in Brazil in August 2013 were effective for prosecuting an deterring crimes of corruption in Brazil[16].

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of American Statistical Association*.

Abbink, K. and Wu, K. (2017). Reward self-reporting to deter corruption: An experiment onmitigating collusive bribery. *Journal of Economic Behavior and Organization*.

Aubert, C., Rey, P., and Kovacic, W. E. (2006). The impact of leniency and whistle-blowing programs on cartels. *International Journal of Industrial Organization*.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*.

Berlin, M. P., Qin, B., and Spagnolo, G. (2018). Leniency, asymmetric punishment and corruption: Evidence from china. *CEIS Tor Vergata Research Paper Series*.

[16]The code and other statistical results are available at: `https://github.com/caxaxa/PhD_Empirics/blob/main/Data_Analysis.ipynb`

Bigoni, M., Fridolfsson, S., Le Coq, C., and Spagnolo, G. (2015). Trust, leniency and deterrence. *Journal of Law Economics and Organization*.

Brenner, S. (2009). An empirical study of the european corporate leniency program. *International Journal of Industrial Organization*.

Buccirossi, P. and Spagnolo, G. (2001). The effects of leniency on illegal transactions: How (not) to fight corruption. *Working Paper*.

Buccirossi, P. and Spagnolo, G. (2006). Leniency policies and illegal transactions. *Journal of Public Economics*.

Chang, M.-H. and Harrington, J. E. (2015). When can we expect a corporate leniency program to result in fewer cartels? *Journal of Law and Economics*.

Chen, Z. and Rey, P. (2013). On the design of leniency programs. *Journal of Law and Economics*.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.

Di Tella, R. and Schargrodsky, E. (2003). The role of wages and auditing during a crackdown on corruption in the city of buenos aires. *Journal of Law and Economics*.

Dufwenberg, M. and Spagnolo, G. (2014). Legalize bribe giving. *Economic Inquiry*.

Engel, C., Goerg, S., and Yu, G. (2016). Symmetric vs. asymmetric punishment regimes for collusive bribery. *American Law and Economics Review*.

Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption. *American Economic Review*.

Franzoni, L. A. (1999). Negotiated enforcement and credible deterrence. *The Economic Journal*.

Golden, M. and Picci, L. (2005). Proposal for a new measure of corruption, illustrated with italian data. *Economics and Politics*.

Grossman, G. and Katz, M. (1983). Plea bargain and social welfare. *American Economic Review*.

Harrington, J. J. E. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics*.

Kaplow, L. and Shavell, S. (1994). Optimal law enforcement with self reporting of behavior. *Journal of Political Economy*.

Landes, W. (1971). An economic analysis of the courts. *Journal of Law and Economics*.

Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*.

Miller, N. H. (2009). Leniency and cartel enforcement. *American Economic Review*.

Motta, M. and Polo, M. (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization*.

Mungan, M. C. and Klick, J. (2016). Reducing false guilty pleas and wrongful convictions through exoneree compensation. *Journal of Law and Economics*.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy*.

Rose-Ackerman, S. (2006). *International Handbook on the Economics of Corruption*. Edward Elgar Publishing Limited.

Spagnolo, G. (2005). Divide et impera optimal deterrence mechanisms against cartels and organized crime. *Working Paper*.

Transparency International (2019). Corruption perceptions index 2019: Technical methodology note. *Transparency International*.

Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-national empirical research? *Anual Review of Political Sciences*.

A Supporting Plots and Tables

Figure 10: New Corruption Inquiries Multivariate Poisson Regression

Table 5: Negative Binomial Regression

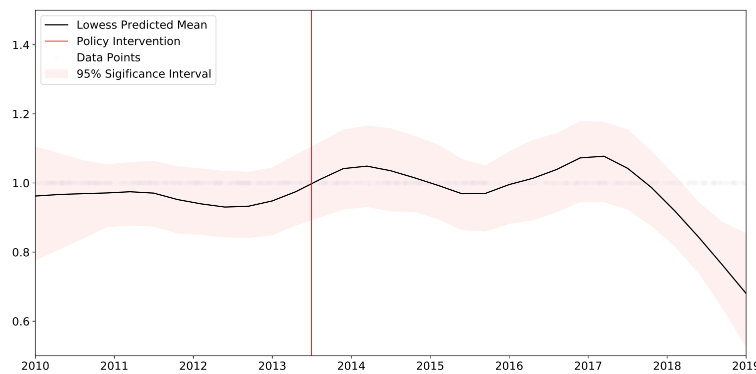
	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
cons	3.549***	3.609***	3.549***	3.573***	3.549***	3.533***
	(0.038)	(0.047)	(0.037)	(0.073)	(nan)	(nan)
dummy	0.101**	0.259***	0.210***	0.236**	0.047***	0.119***
	(0.048)	(0.087)	(0.068)	(0.095)	(nan)	(nan)
t2			-0.003**	-0.002	0.012***	0.019***
			(0.002)	(0.003)	(nan)	(nan)
s2					-0.000***	-0.000***
					(nan)	(nan)
Full Seri Polynomial	None	1st Order	None	1st Order	None	2nd Order
Post-policy Polynomial	None	None	1st Order	1st Order	2nd Order	2nd Order
Observations	108	108	108	108	108	108
R^2						
Adjusted R^2						
Residual Std. Error	1.000(df = 106)	1.000(df = 105)	1.000(df = 105)	1.000(df = 104)	1.000(df = 104)	1.000(df = 102)
F Statistic	(df = 1.0; 106.0)	(df = 2.0; 105.0)	(df = 2.0; 105.0)	(df = 3.0; 104.0)	(df = 3.0; 104.0)	(df = 5.0; 102.0)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6: Poisson Regression with Lagged Variables

	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.559***	3.538***	3.187***	3.135***	3.742***	3.745***
	(0.111)	(0.110)	(0.126)	(0.151)	(0.253)	(0.312)
Dummy	0.099	0.358***	0.350***	0.325***	0.238***	0.295**
	(0.067)	(0.088)	(0.078)	(0.087)	(0.090)	(0.116)
Unemployment	-0.000	0.021	0.048***	0.050***	-0.026	-0.030
	(0.012)	(0.013)	(0.015)	(0.015)	(0.033)	(0.036)
Real Interest	-0.785	-5.463	-6.214*	-5.921*	-9.771***	-9.737**
	(3.319)	(3.493)	(3.456)	(3.486)	(3.736)	(3.901)
GDP	-0.050	0.593	2.135**	2.314**	2.844***	2.888***
	(0.919)	(0.927)	(0.989)	(1.028)	(1.023)	(1.052)
Full Seri Polynomial	None	1st Order	None	1st Order	None	2nd Order
Post-policy Polynomial	None	None	1st Order	1st Order	2nd Order	2nd Order
Observations	108	108	108	108	108	108
R^2						
Adjusted R^2						
Residual Std. Error	1.000(df = 103)	1.000(df = 102)	1.000(df = 102)	1.000(df = 101)	1.000(df = 101)	1.000(df = 99)
F Statistic	(df = 4.0; 103.0)	(df = 5.0; 102.0)	(df = 5.0; 102.0)	(df = 6.0; 101.0)	(df = 6.0; 101.0)	(df = 8.0; 99.0)
Note:						*p<0.1; **p<0.05; ***p<0.01

Figure 11: LOWESS Daily Frequency



B THINGS THAT DID NOT WORK

Finding a Counter Factual (Difference-in-Differences)

Generally in comparative case analysis, in order to be able to state causal inferences, it is needed at least one treated and one control group. Of course, the ideal set for a difference-in-difference analysis would require two groups of potential corruptors in which one is affected by the policy and the other is not. In this case, unfortunately, every potential criminal for corruption is exposed at the same time by the policy. However, there are other heterogeneities to be explored, such as different types of crimes. Therefore, assuming that corruption criminals are the treated group and all the other criminals are the control group, it is possible to test the difference between both groups using a difference-in-difference estimator such as:

$$Y_t = \beta_0 + \beta_1 T1_t^1 + \beta_2 D_{1t} + \beta_3 D_{1t} * (treatment) + \beta_4 X_t + \varepsilon_t \tag{2}$$

Where, Y_t represents all crimes and $D_{1t} * (treatment)$ are the corruption crimes (interaction term).

If the control and treatment groups respect the parallel trends assumption on the pre-treatment period, then the diff-in-diff estimator is consistent. However, if series are not parallel before the policy shock, then a better counterfactual or control group is necessary.

One way to overcome this problem using the richness of the data is to create synthetic controls from the other crimes (Abadie et al., 2010). The underlying assumption is that detection of corruption crimes resembles some similarities with other detected crimes. In this sense, it is possible to generate a synthetic series, using a weighted average of all convex combinations of all other crimes that minimize the difference between the synthetic series and the treatment group in the pre-treatment trend. After this, the synthetic generated series is used as a counterfactual.

Testing other Interesting Heterogeneities (More Diff-in-Diff)

There is also the possibility to collet more information from each case if the investigation is already made public[17]. In this sense, it is possible to sample random cases from before and after the policy and analyse their characteristics. Some interesting candidates are the values involved in the crime, the number of participants, the duration of the scheme and others.

It is expected from the corruption crimes detected before and after the anti-corruption policy to have different characteristics. For instance, after the policy shock, few people would engage corruption, in this way, the average duration of the bribery schemes detected after the treatment should be higher as we depart from the date of the shock (Chang and Harrington, 2015).

However, some interesting findings could be also inferred from the data itself. For example, if values involved in the bribery schemes or any other distinguishable characteristics are significantly different, perhaps some conclusions about features of the policy can be drawn.

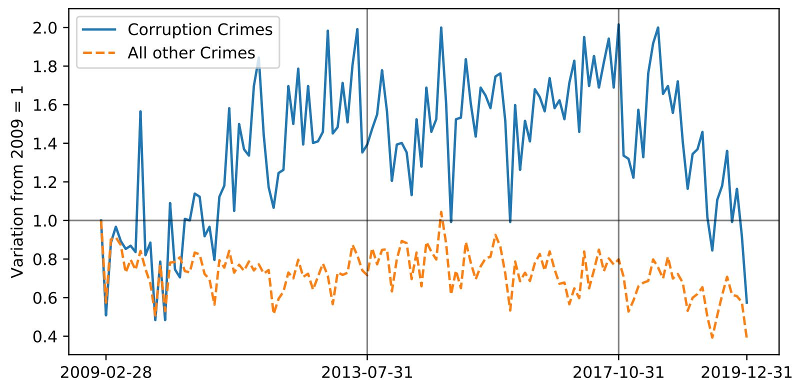
Diff-in-Diff

The models tried above care little about causal inference, since there is no other parameter for comparison. In other words, every potential criminal before the policy threshold is in the control group and everyone after is in the treatment group. In this sense, if there are any unobservable variables affecting individuals over time, besides the one we are interested in, they should not be accounted for. One way to overcome this problem is by looking for eventual differences in similar distinct groups (control and treatment) before and after the policy shocks, i.e. differences in differences.

In order to be able to estimate differences in differences, first is necessary to assign individuals to distinct groups. For instance, individuals investigated for corruption crimes are the treatment group and individuals investigated for all other crimes are the control group. After this definition, it is necessary to assure that both groups have a similar pre-policy trend. The Figure 12 shows the variation of each type of criminal detection over time.

[17] The Brazilian Constitution imposes that all judicial decision (and non confidential processes) need to be publicly available.

Figure 12: Difference in Differences : Checking for the parallel pre-treatment trend



Note that the series behave very differently in the period before the shock. This shows that probably these groups are not so similar regarding unobservable characteristics. Perhaps comparing corruption criminals with all other criminals is not a good idea.

In any case, since the intention of this research proposal is to explore potential empirical strategies, let the Table 8 show the estimated parameters for these candidates. The 'interaction' parameters represent policy dummy (1 , 2, 3 and 5) or the post policy dummy (4 and 6) times the treatment group. One can note that the interaction parameters show the sign of the differences. Analysing the significant ones, the 'Interactions 2 and 5' show that the policy increased the number of corruption investigations in the entire period. Whereas the 'interaction 4' shows that there is a decrease in new corruption detections (deterrence effect).

In front of these contradictory findings, and the lack of pre-treatment parallel trends. **We should not use this model**. However, there is a way to use the heterogeneity of the data to extract a better counterfactual for a diff-in-diff estimation. It is possible to use synthetic controls calculated from the combination of different crimes. The implicit assumption here is that the corruption crimes may have some unobservable similarities with other crimes. More precisely, detection of different crimes may resemble the corruption detections in different ways. Consequently, one can derive a vector of weighted averages from different crimes and use them as a control group. This approach may perhaps allow this research to better use diff-in-diff estimators.

Lastly, it is possible to sample individual cases from before and after the policy intervention. However, sometimes it is not easy to find information about the processes online without their identification number from the judiciary. It is possible to request the identification numbers from the Brazilian judiciary.

However, a better definition of the sample must be done before sending the request[18].

Exploring State Heterogeneity (Panel Regressions)

Since the data is disaggregated at the case level, it is possible to build a panel data and verify if is there any difference between states when it comes to combating corruption. One possible interesting issue is to verify if is there any difference in detection of corruption between states that enforced the law (dummy or number of awarded collaborations) and the others that did not. The following model can be explored in such a way:

$$Y_{it} = \beta_0 + \beta_{1i} + \beta_2 D_t + \beta_3 X_{it} + \varepsilon_{it},$$

for,
i= Different Brazilian states; and
t= Monthly or daily observations.
Where
Y_{it}= Number of detected corruption crimes;
β_{1i} = State fix effects;
D_t= Dummy for the impact of the enforcements; and
ε_{it}= Is the error from functional predictions, it is expected to be normally distributed and i.i.d.

Off course, this model tells little about causal effects. However, the strategies proposed above may also be applied at the state level.

Panel Regressions

It is possible to explore differences between states with the data. However it is not possible to regress all the corruption crimes using states as exogenous variables. This happens because the dependent variable is a sum of the crimes in every state. Therefore it is a completely identified problem. Nonetheless, it is possible to regress one specific state against all others.

The Figure 13 shows how the investigations of corruption crimes varied from state to state over time. It is possible to observe that states are very different from each other regarding corruption investigations. This happens because some smaller states have big variances. However, bigger states, which investigate more crimes also behave very differently.

Table **??** shows the results of the regressions for the states of Rio de Janeiro, Distrito Federal and Paraná. Note that the results differ from each other considerably. For instance, Sao Paulo shows signs of an increased productivity phase followed by a deterrence. Whereas Distrito Federal showed only increases in detections and Paraná decreased detections.

[18] Define the features of the sample (number o observations, which ones, which other eventual crimes etc.), guarantee its randomness and other desirable features.

Figure 13: Investigations of Corruption Crimes per State

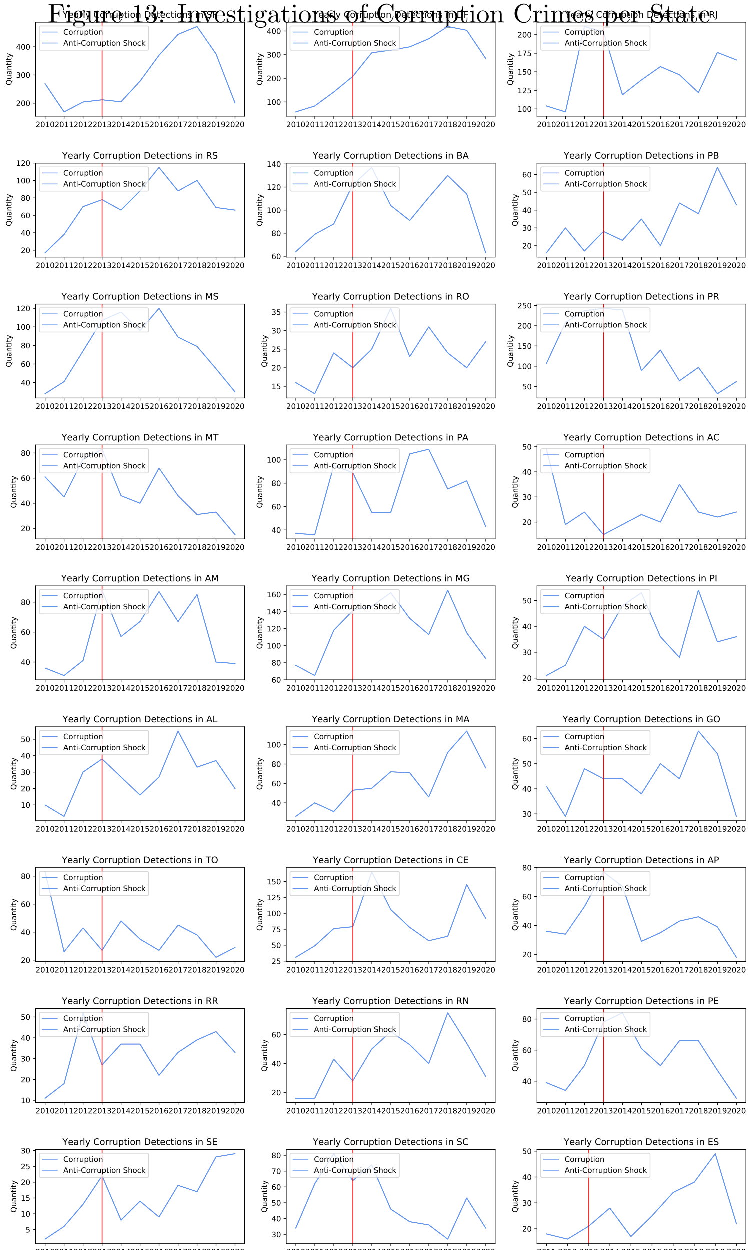


Table 8: Difference in Differences Estimations

	Dependent variable: All Crimes			
	(1)	(2)	(3)	(4)
Constant	0.724***	0.741***	0.679***	0.743***
	(0.066)	(0.088)	(0.068)	(0.061)
Interaction 1	0.006			
	(0.016)			
Interaction 2		0.002***		
		(0.001)		
Interaction 3			0.012	
			(0.016)	
Interaction 4			-0.043**	
			(0.019)	
Interaction 5				0.001***
				(0.0)
Interaction 6				0.0
				(0.0)
Linear Trend		-0.004***		-0.004***
		(0.001)		(0.0)
GDP	-49.313	-128.455*	6.331	-66.102
	(91.824)	(75.309)	(93.789)	(73.658)
Unemployment	-0.022***	-0.007	-0.015***	-0.01*
	(0.005)	(0.006)	(0.006)	(0.005)
Interest Rate	-1.381	-0.341	-1.338	-1.001
	(1.702)	(1.465)	(1.676)	(1.344)
Corruption Crimes	0.159***	0.1**	0.152***	0.168***
	(0.028)	(0.046)	(0.028)	(0.023)
Observations	132.0	132.0	132.0	132.0
Adjusted R2	0.311	0.493	0.332	0.578
Residual Std. Error	0.094	0.081	0.093	0.074
F Statistic	12.806***	22.26***	11.837***	26.639***
Note:				*p<0.1; **p<0.05; ***p<0.01