Corruption Deterrence Through Self-Reporting and Collaborations: The Brazilian Example

Lucas Alves Chacha

March 3, 2020

Abstract

In August 2013, Brazilian legislators enacted two laws allowing individuals and corporations to self-report corruption crimes and collaborate with the authorities in exchange for sanction benefits. Later on, in 2014 the Brazilian prosecutors started a series of big corruption investigations unveiling major corruption schemes between the Brazilian government and important private corporations. In these cases, the prosecution strategy to pursue the defendants relied heavily on self-reporting and collaborations. In this context, one can argue that the investigations are evidence of the success of the anti-corruption policy. However, the extent of its influence over a broader scope of corruption is not clear. In order to empirically check if the policy is effective against corruption, I propose a comparative case study. In this sense, this research proposal uses a dynamic model of corruption deterrence to extract theoretical expected time paths of corruption detection. Furthermore, the theoretical predictions are tested against Brazilian data. If the the theoretical assumptions are true then the data should confirm its predictions[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) more measures against corruption from the government. The laws originated from the guidelines of the United Nations Convention Against Corruption (UNCAC)[2]. They introduced the possibility for individuals and corporations involved in corruption to self-report 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 conduced by its employees. It institutes compliance measures 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.

[1] This is a research proposal for the empirical part of my thesis. First, the reader should know that the motivation and specific characteristics of the Brazilian anti-corruption policy are discussed in Part I. Secondly, the theoretical model of domestic corruption deterrence is discussed in Part II. Lastly, this current empirical research proposal is relative to the third part of the work.

[2] 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.

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 wide spread corruption or an effective shift of the prosecution standards? More specifically, has the anti-corruption policy from 2013 decreased domestic bribery in Brazil? The hypothesis here is that the policy decreases corruption activities. It happens by enhancing prosecutorial productivity and by decreasing the individual incentives from potential criminals to engage in briberies.

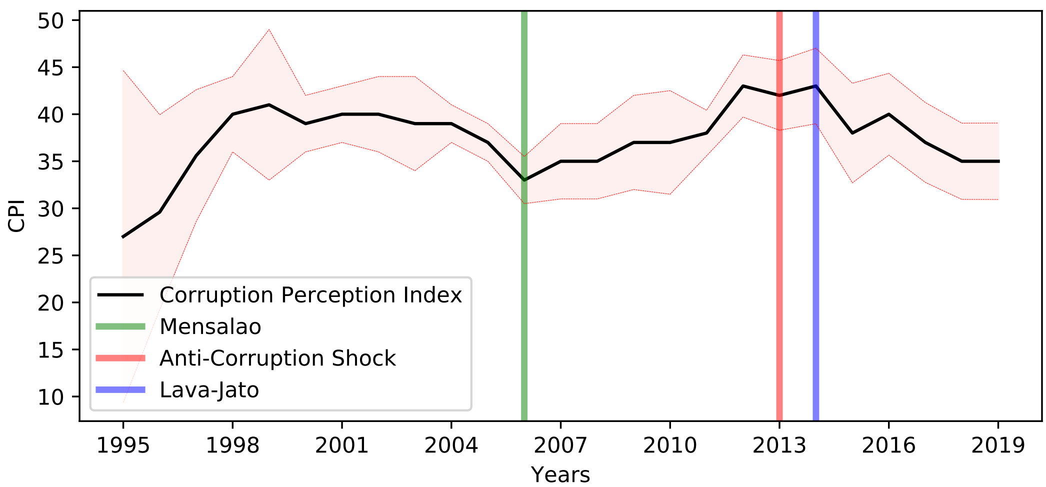
In order to test the hypothesis empirically, this research proposal presents a methodology to estimate a possible change in the unobservable corruption activities in Brazil.

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 number of corruptors or corruptees in the economy, or even the total value of bribes paid in a certain period. In order to measure corruption levels, one have to rely on evidences of the briberies.

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 crackdowns 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[3]. Therefore, CPIs may not be good assessments for corruption in a particular region over time after a policy shocks. The Figure 1 shows the evolution of the Transparency International Corruption Perception Index[4] 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

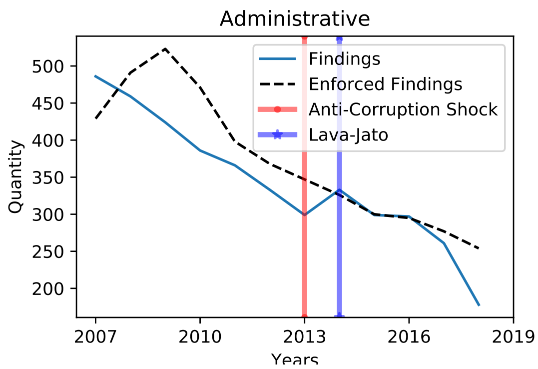
[3]There are methodological attempts to minimize this kind of bias. However, it still the case that the indexes are fundamentally measuring perceptions (Transparency International, 2019).

[4]100 = Completely clean; 0 = Completely Corrupt.

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 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)[5]. 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).

In the next sections the dataset is presented and discussed, then the theoretical predictions are set, subsequently the empirical strategy is defined and lastly some preliminary findings are shown along with some suggestions of empirical extensions for this research.

Theoretical Framework

As discussed above, in order to estimate the effect of an unobservable variable one can only rely on evidence shown by other observable variables. In this sense, a theory about the mechanics of the corruption crimes is needed.

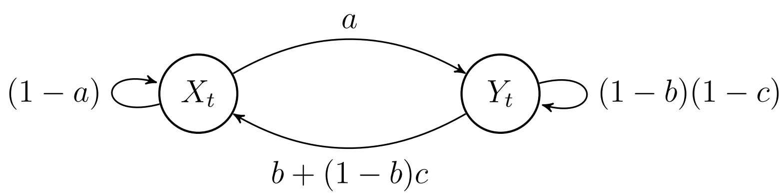
Miller (2009) proposed one way to test empirically if an antitrust policy can deter cartel formation. He innovates by constructing a model showing the mechanism of transition from collusion to non-collusion for cartels. The model consists in a two-state first order Markov process, in which firms are either in one of the two phases in time. 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 transition diagram from Miller's model[6].

The diagram shows the mechanism in which the firms change from colluding to not colluding. This is effective on showing how the observable detection of the collusions may

[5]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`

[6]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.

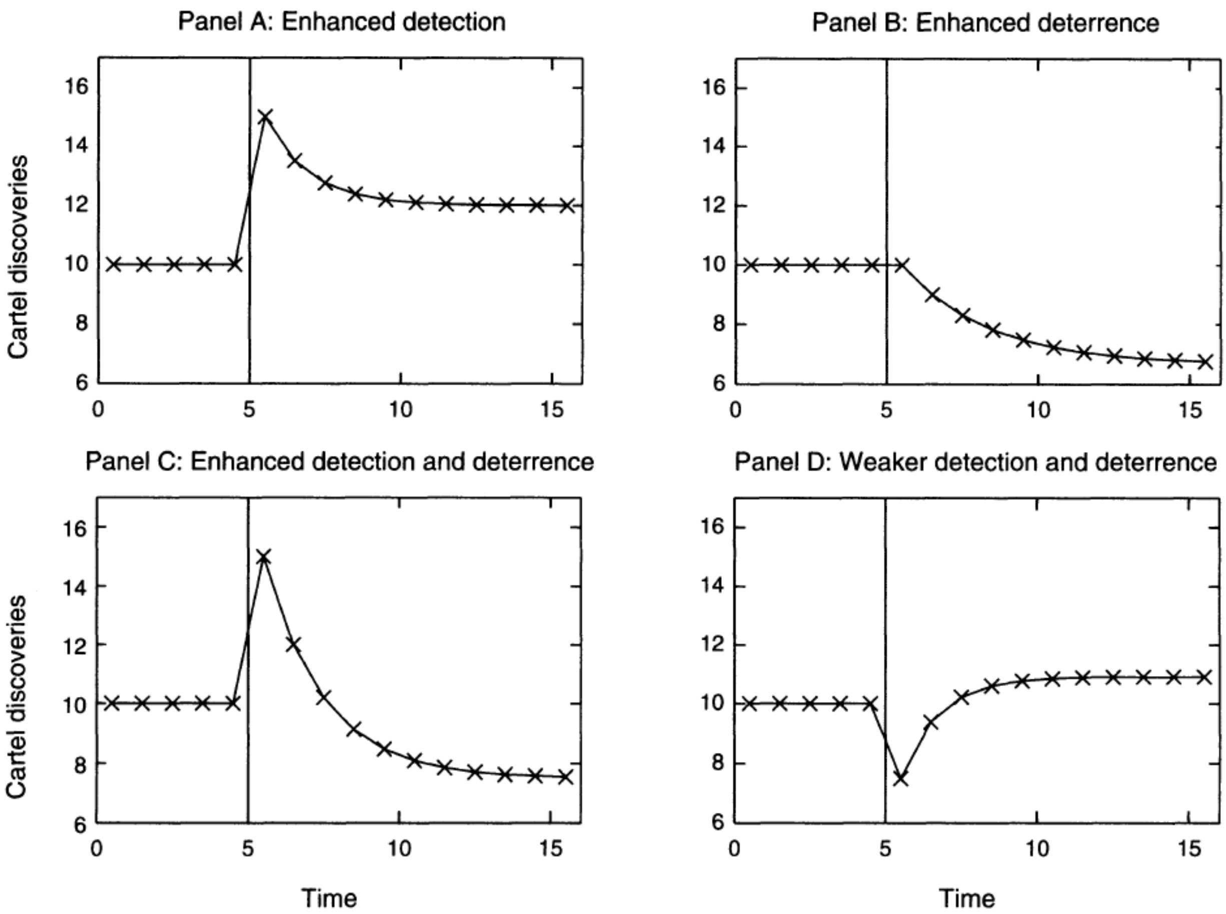
Figure 3: Markov Transition Diagram of Miller (2009)



change in time. However, the model does not care for understanding the incentive mechanism of collusion[7]. In this sense, it is a good model for orientating the empirical strategy yet less effective to explain 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 a first moment due to enhanced of detection efficiency. Later, after a while it drops to levels lower than before the enforcement shock (deterrence effect). Figure 4[8], exemplifies the expected path of the variable.

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



Notably, this study is not interested in cartel activity. However, the same methodology

[7]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

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

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 bellow).

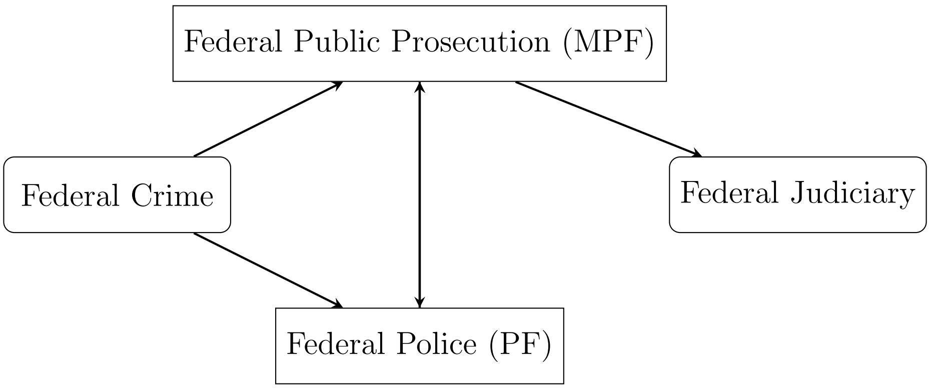
Institutional Background

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 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 conduce parallel investigations conjointly (Art. 5, II). Only in exceptional cases the Prosecution carries out one investigation only by itself[9].

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

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.

The Data

The data used here was downloaded from the online processual search engine of the Brazilian Public prosecution (MPF)[10]. It consists of 885.675 investigations conduced by the Brazilian Federal Police, sorted by date of the beginning of the investigation from January 2009 to January 2020[11]. The data also provides the crime, and the state in which the investigation

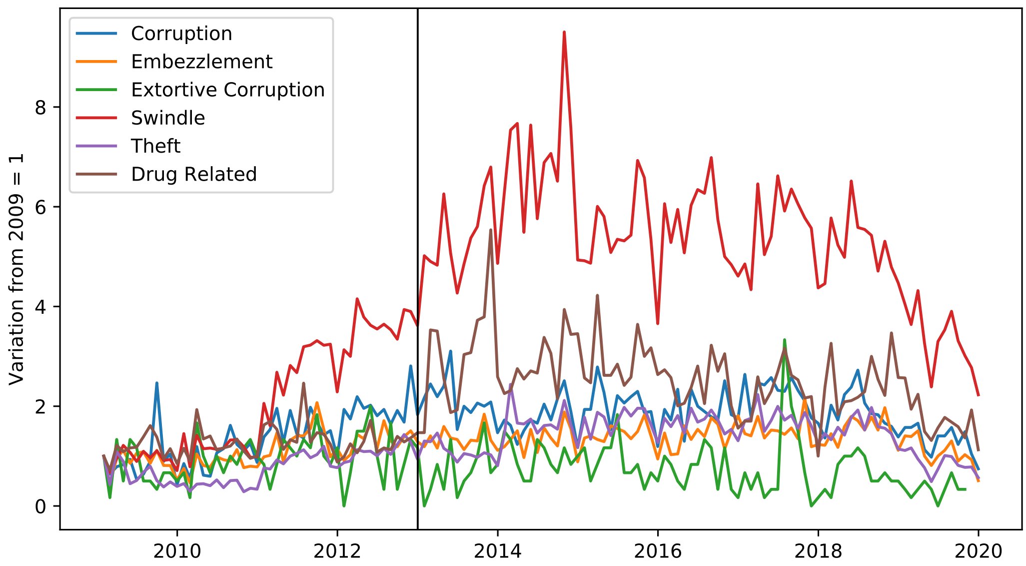
[9]It was only in 2015 the Brazilian Supreme Court formalized the understanding that the Brazilian Prosecution could investigate a case alone ((HC) 89837/STF). Before this decision, criminal investigation was a Police monopoly.

[10]Available at: `http://apps.mpf.mp.br/aptusmpf/portal?servidor=portal`.

[11]Periods before 2009 can be biased for minor states which did not use the informational system at the time.

is conduced. The Tables 1 and 2 show some aggregate values from dataset, also Figure 6 shows the variation of some crimes over the past ten years.

Figure 6: Investigation Detection Variation (2009-01 = 1)



Empirical Strategy

In this section candidates for empirical strategies are individually set. The preliminary results of each strategy is provided in the next section.

Analysing the Detection Series Alone (Regression Discontinuity)

The first 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 criminal detections, precisely the intention of this work. The authors' approaches can be adapted to this case and formalized as follows:

$$Y_t = \beta_0 + \beta_1 D_t + \beta_2 T1_t^n + \beta_3 T2_t^n + \beta_4 X_t + \varepsilon_t \tag{1}$$

for,
t= Monthly (or daily) observations; and
n= Order of the polynomial.

Where
Y_t= Number of 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 polynomials that might fit the expected detection curve, the variables $T1$ and $T2$ can be tested on different order polynomials to check 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.

Since it is expected two different effects on the observable variables, it is possible to test two distinct regression discontinuities. The estimation can assume the following functional form:

$$Y_t = \beta_0 + \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 X_t + \varepsilon_t \tag{2}$$

where D_{1t} and D_{2t} are respectively the dummies for the policy shock and for a reasonable period after it. This approach assumes that, passed a reasonable period after the shock, one could start observing a deterrent effect from the policy.

Lastly, Berlin et al. (2018) also use a Quandt-Likehood 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 is there 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.

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:

$$Y_t = \beta_0 + \beta_1 T1_t^1 + \beta_2 D_{1t} + \beta_3 D_{1t} * (treatment) + \beta_4 X_t + \varepsilon_t \tag{3}$$

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[12]. 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 Jr, 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.

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.

Preliminary Results

In the following subsection I present some preliminary results from the above mentioned research strategies. It is important to notice that, more than scrutinizing the statistical properties of the regressions, this research proposal intends to discuss the research design of each strategy. In this way, the relevant conclusions are drawn from the relevance and feasibility of the underlying assumptions along with some evidence of how the data behaves in certain situations.

Most of the regressions are aggregated at the monthly level. Daily or case level regressions have to many zero observations, this is known to distort the normality property of the OLS

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

regression residuals. Nonetheless, in future regressions it is possible to explore the data at this level with the proper precautions. Also, most of the data is indexed at January 2009, this is done to avoid problems of interpretation of parameters when levels are too distinct. Lastly, the regressed dependent variable consists in the sum of the crimes of corruption, embezzlement and extortive corruption[13]. From this point on the term 'crimes of corruption' refers to this aggregation.

Regression discontinuity

The first set of regressions show how different order polynomial fit the data of crimes of corruption[14]. The Table 3 shows the regressions (1) to (6), containing the coefficients for the intercept and the policy dummy for different order polynomials. Each polynomial is drawn independently both before and after the shock. The Figure 7 show the fitted curve over the data.

Figure 7: Regression Discontinuity

From Table 3 it is possible to observe that the dummy representing the policy shock is significant only for regressions (1), (3) and (5). Notably, the first and the fifth show that, after the policy intervention there is evidence of a significant spike on detections. However,

[13]Extortive corruption (concussao, Law No. 2.848/40, Art. 316) happens when someone requires to himself or herself bribe without any *quid-pro-quo* counterpart.

[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.

the third one shows that, considering that the spike happened a little bit after the policy[15], the overall trend has gone down. Moreover, regarding the deterrence effect, all the higher order polynomial functions point out to a more deterrent present and future. This result is in line with Miller (2009), and shows that the policy may be effective against corruption after all.

The second set of regressions added other seemingly exogenous variables to the regression. The chosen ones are the quarterly Brazilian real GDP growth[16], the monthly Brazilian unemployment rate[17], the monthly average of the main national treasury bonds real interest rate (SELIC)[18] and all the other crime investigations except the corruption crimes.

The underlying motives for choosing these variables are distinct. The GDP growth accounts for people's quality of life (even though is not the *per capita* measure) and may perhaps affect crime rate. By the same logic, unemployment may have a liquidity effect and incentivize people to commit crimes. Additionally, since bribery is mainly a rent seeking activity, it may be affected when real interest rates change. Lastly, there may be an overall trend in criminal detection that can be explained by the investigation of other crimes. Note that no discussion is made about what are the time gaps between the change in the variable, the performance of the crime and its eventual detection. In the absence of a good theory for the timing mechanism, everything is regressed just in time[19].

The Table 4 shows the regressions (1) to (4), containing the additional variables. The regressions (1) and (2) do not have the polynomials whereas the regressions (3) and (4) have no polynomials before the shock and a 3rd order polynomial after the policy. It can be observed that there is a significant positive impact of the of the policy in the corruption detections in the specifications without using trend ((1) and (2)). However, when the trend is added, the effects are no longer significant((3) and (4)). Figure 8 shows how the estimated curves fit the data.

[15]There is reason to believe that this is not a sharp discontinuity. However, a better theory for the timing mechanism is needed.

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

[17]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.

[18]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`.

[19]Any ideas are welcome.

Figure 8: Regression Discontinuity

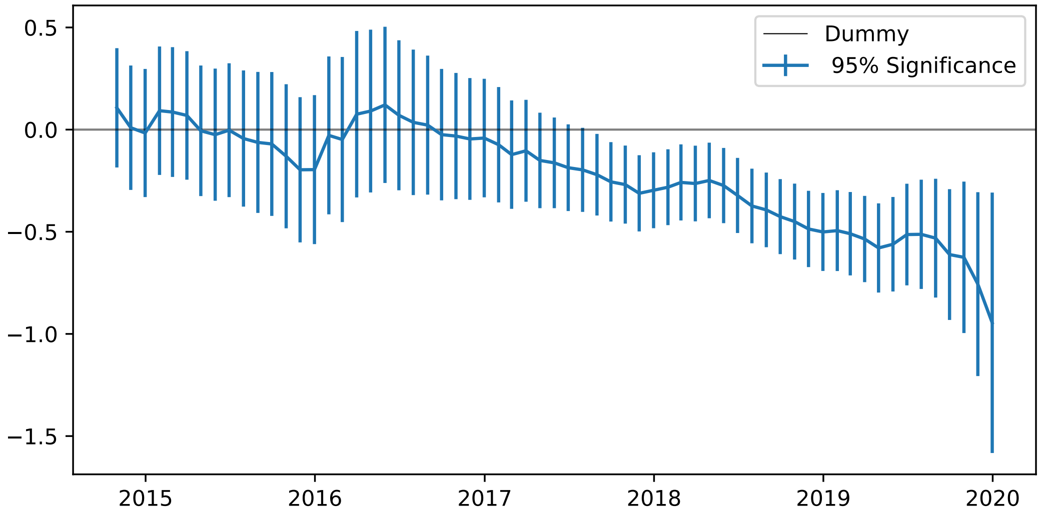
Note that, without a higher order trend in the functional form, it is not possible to differentiate the spike (gain of productivity) from the lower final plateau (deterrence effect) just by looking at the policy dummy sign. In this sense, to overcome this issue, one can check for two discontinuities. In this way, it is expected a positive value for the first dummy and a negative for the other. Another way to illustrate the strategy, is by setting two different treatments, one after the policy shock and other at a reasonable time after de shock.

The problem of using a second treatment and consequently another discontinuity to this particular issue is that there is no theory on where to define the threshold. In other words, where do we draw the line for the start of the deterrence effects? Or else, what is a reasonable period for the deterrence effect to start being observed? Combating corruption in a big country as Brazil can take several years.

One way to solve this issue, is by looking at the data. It is possible to simulate the above specification until the point that it is likely to have a break in the trend. The Figure 9 shows all the values and standard errors for a dummy in each point in time[20]. It is possible to see that, after 2018 there is a persistent change in the trend that leads to significant dummies. In this way, for simplicity, I defined January 2018 to be the threshold for the second treatment.

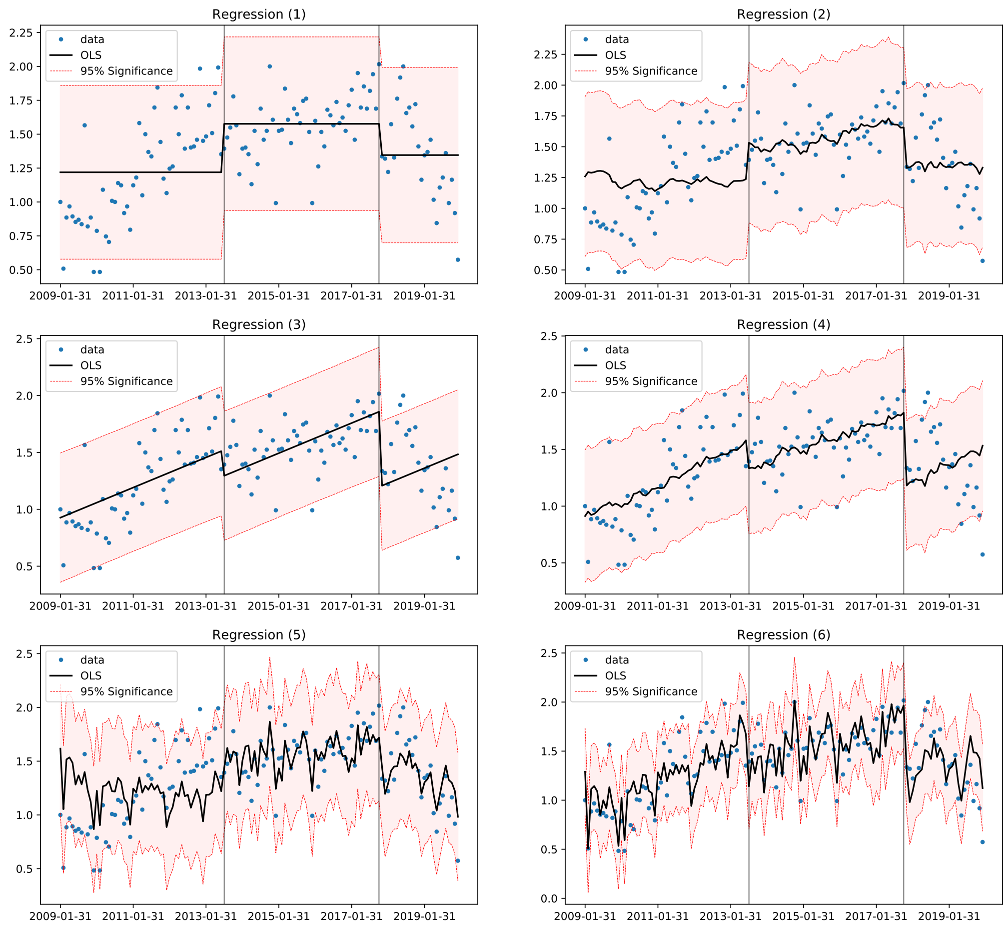
[20]The functional specifications has a constant, GDP, Unemployment, real interests and the policy dummy. I chose to let out other crimes in order lo leave out the criminal trend.

Figure 9: Placebo Tests



The third set of regression discontinuities results can be checked on Table 5. It is possible to observe that the dummies have the intended signs((1),(2) and (5)). Or else, there is a positive spike of productivity after the anti-corruption policy, and after some time, it diminishes as a possible decrease of the corruption activity. However, when the regression has a linear time trend, the positive effect on the first dummy vanishes ((3),(4) and (6)), staying only the deterrence effect of the second dummy. Note that it does not immediately invalidate the gain of productivity hypothesis, since the time trend is positive, i.e. after the policy there is still an increase in corruption detections. The Figure 10 helps visualize the effects separately.

Figure 10: Regression Discontinuity



The results shown here seem to indicate that the expected effects may be present. However, there is room for different interpretations.

Diff-in-Diff

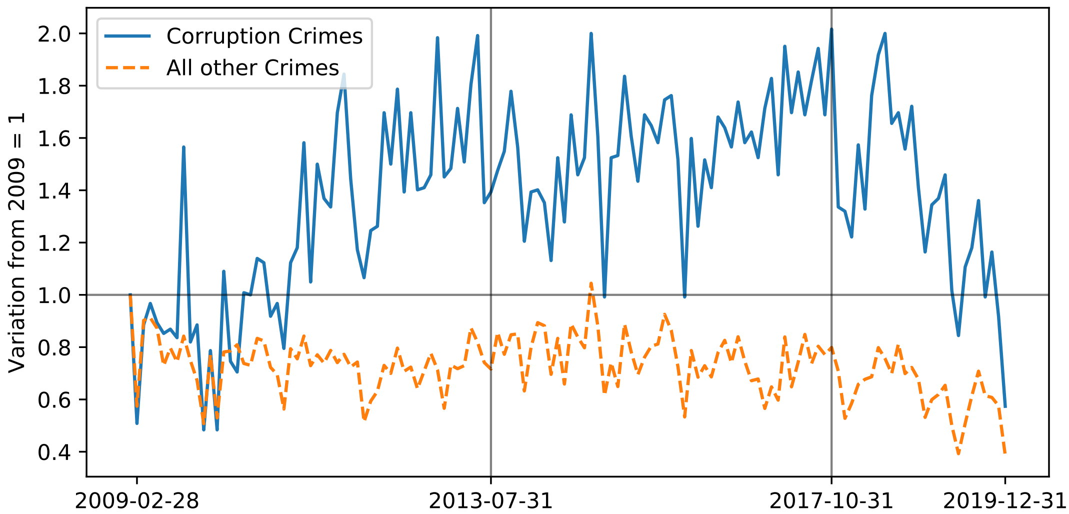
The models tried above care little about causal inference, since there is no other parameter for comparison. In other words, every potential criminal before the policy threshold is in the control group and everyone after is in the treatment group. In this sense, if there are any unobservable variables 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[21]. For instance, individuals investigated for corruption crimes

[21]Note that, estimation (2) in Table 3 is a diff-in-diff estimation considering the groups before and after the policy and the time trend. The estimated interaction term is negative. Therefore it shows that the treatment

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 11 shows the variation of each type of criminal detection over time.

Figure 11: 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 11 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 richness 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 identi-

lowers the expected criminal detection considering the upper time trend. One can call it a deterrence effect, since short term spikes are considered in the trend.

fication numbers from the Brazilian judiciary. However, a better definition of the sample must be done before sending the request[22].

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 12 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 7 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.

Final Remarks

After presenting the preliminary results, it is possible to conclude about the research design:

1. There are controversial evidences that the Brazilian anti-corruption policy decreased unobservable corruption;

2. The policy shock seems to be fuzzy. In other words, it seems that there is a time gap between the enactment of the law and its expected effects;

3. For a more precise causal inference it is necessary to identify better fitted control and treatment group;

4. Characteristics from samples from the detected corruption crimes can tell something about the anti-corruption policies; and

5. Corruption detection in different Brazilian states vary considerably.

Lastly, the discussion is open for any suggestion on how to have more precise results about the anti-corruption policy effects on corruption[23]

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

[23]For anyone interested in more details from the data. In the followings links you can find:

Extensive description of the data and some different regressions: `https://github.com/caxaxa/Chacha_PhD_Projects/blob/master/MPF_ANALYSING_STRUCTURED_DATA.ipynb`.

The research proposal's regressions: `https://github.com/caxaxa/Chacha_PhD_Projects/blob/master/MPF_ESTIMATED_MODELS.ipynb`.

The data in panel format and estimations: `https://github.com/caxaxa/Chacha_PhD_Projects/blob/master/MPF_ANALYSING_STRUCTURED_DATA_PANEL_DATA.ipynb`

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.*

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

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

Brenner, S. (2009). An empirical study of the european corporate leniency program.

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

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.*

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

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

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

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.*

Table 1: Number of Investigations of Corruption Crimes per Brazilian State (Corruption, Embezzlement and Extortive Corruption)

date region	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018-	2019	2020-12-31
AC	31.0	19.0	23.0	11.0	18.0	21.0	20.0	30.0	21.0	22.0	20.0	0.0
AL	9.0	3.0	26.0	22.0	22.0	13.0	24.0	28.0	30.0	31.0	16.0	0.0
AM	32.0	28.0	36.0	68.0	47.0	48.0	67.0	38.0	44.0	31.0	34.0	3.0
AP	34.0	31.0	47.0	53.0	48.0	23.0	30.0	35.0	35.0	28.0	16.0	0.0
BA	48.0	62.0	64.0	89.0	102.0	80.0	74.0	81.0	89.0	86.0	53.0	2.0
CE	24.0	39.0	57.0	60.0	76.0	69.0	64.0	48.0	51.0	90.0	63.0	0.0
DF	50.0	57.0	97.0	128.0	162.0	155.0	171.0	162.0	188.0	192.0	157.0	2.0
ES	0.0	17.0	16.0	17.0	26.0	14.0	18.0	33.0	33.0	38.0	19.0	1.0
GO	32.0	24.0	40.0	42.0	43.0	35.0	43.0	37.0	54.0	49.0	27.0	0.0
MA	22.0	31.0	24.0	41.0	37.0	53.0	57.0	42.0	65.0	77.0	58.0	1.0
MG	61.0	51.0	80.0	102.0	102.0	113.0	92.0	82.0	102.0	85.0	71.0	1.0
MS	23.0	36.0	55.0	78.0	84.0	81.0	84.0	73.0	51.0	43.0	26.0	0.0
MT	49.0	39.0	50.0	62.0	41.0	35.0	53.0	44.0	28.0	29.0	15.0	0.0
PA	28.0	34.0	76.0	68.0	45.0	50.0	84.0	82.0	67.0	64.0	32.0	5.0
PB	13.0	20.0	16.0	24.0	21.0	28.0	18.0	35.0	34.0	43.0	34.0	0.0
PE	32.0	29.0	43.0	69.0	65.0	51.0	46.0	56.0	57.0	39.0	23.0	1.0
PI	18.0	24.0	36.0	35.0	40.0	38.0	32.0	23.0	42.0	30.0	26.0	0.0
PR	69.0	103.0	147.0	143.0	166.0	75.0	78.0	41.0	60.0	25.0	50.0	2.0
RJ	74.0	74.0	113.0	124.0	82.0	82.0	90.0	103.0	83.0	108.0	101.0	4.0
RN	13.0	14.0	36.0	24.0	43.0	52.0	43.0	32.0	52.0	43.0	29.0	0.0
RO	16.0	12.0	23.0	20.0	23.0	30.0	23.0	28.0	22.0	18.0	22.0	0.0
RR	10.0	17.0	26.0	21.0	31.0	24.0	18.0	19.0	35.0	34.0	29.0	1.0
RS	12.0	35.0	50.0	56.0	54.0	65.0	76.0	69.0	75.0	54.0	51.0	2.0
SC	28.0	49.0	67.0	56.0	56.0	37.0	34.0	33.0	25.0	42.0	32.0	0.0
SE	2.0	6.0	12.0	18.0	8.0	14.0	9.0	17.0	15.0	26.0	23.0	0.0
SP	107.0	103.0	131.0	132.0	129.0	151.0	183.0	182.0	189.0	162.0	117.0	2.0
TO	52.0	24.0	31.0	23.0	28.0	26.0	20.0	34.0	30.0	20.0	22.0	0.0
Sum	889.0	981.0	1422.0	1586.0	1599.0	1463.0	1551.0	1487.0	1577.0	1509.0	1166.0	27.0

Table 2: Number of Annual Brazilian Investigations per Crime

Date	Total	Corruption	Embezz- lement	Environ- mental	Extortive Corruption	Swindle	Theft	Drugs	Against Property	Financial Financial	Authority Abuse	Procure ment Fraud
2009	84612	521	734	16	52	3217	2695	1835	9674	764	101	775
2010	79977	569	700	35	67	3923	1745	1960	8773	938	101	612
2011	81835	867	1088	179	77	9081	3969	2327	4051	1267	161	1026
2012	76483	1103	1099	186	74	12060	4663	2036	975	1089	145	1080
2013	85489	1191	1109	234	50	17926	4637	4907	801	1155	105	1323
2014	90158	985	1108	273	70	23121	7106	4645	505	1116	105	1180
2015	84525	1137	1177	533	62	18213	7249	4662	287	1289	73	1073
2016	80539	1088	1191	291	53	19022	7060	3906	254	1050	86	1125
2017	80188	1266	1241	236	61	18626	7346	3579	193	1074	75	1339
2018	76688	1073	1263	399	49	17522	6625	4012	177	968	71	1272
2019	62349	743	890	317	20	11078	3593	2758	63	787	45	894

Table 3: Regression Discontinuity

	Dependent variable: Monthly Corruption Crimes Investigations					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.219***	0.697***	0.732***	0.918***	1.219***	1.219***
	(0.045)	(0.074)	(0.102)	(0.133)	(0.041)	(0.041)
Policy Dummy	0.281***	-0.107	-0.49***	-0.068	0.317**	0.19
	(0.059)	(0.094)	(0.127)	(0.159)	(0.148)	(0.192)
Trend Before	None	1st Order	2nd Order	3rd Order	None	None
Trend After	None	1st Order	2nd Order	3rd Order	3rd Order	5th Order
Observations	132.0	132.0	132.0	132.0	132.0	132.0
R2	0.15	0.453	0.56	0.616	0.32	0.331
Adjusted R2	0.144	0.44	0.543	0.594	0.298	0.298
Residual	0.331	0.267	0.242	0.228	0.299	0.299
Std. Error	(df = 130.0)	(df = 128.0)	(df = 126.0)	(df = 124.0)	(df = 127.0)	(df = 125.0)
F Statistic	22.999***	35.306***	32.112***	28.376***	14.915***	10.289***
	(df = 1.0; 130.0)	(df = 3.0; 128.0)	(df = 5.0; 126.0)	(df = 7.0; 124.0)	(df = 4.0; 127.0)	(df = 6.0; 125.0)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4: Regression Discontinuity 2

Corruption Crimes:	Without Polynomials		With Polynomials	
	(1)	(2)	(3)	(4)
Constant	1.341***	-0.038	2.155***	1.338***
	(0.183)	(0.269)	(0.367)	(0.347)
Policy Dummy	0.233***	0.228***	0.21	0.041
	(0.08)	(0.069)	(0.16)	(0.142)
Real Interest	-3.664	-1.606	-8.334	-6.962
	(5.966)	(5.217)	(5.68)	(4.981)
GDP	-509.169	-222.577	-602.764*	-588.064*
	(323.015)	(285.511)	(357.073)	(312.791)
Unemployment	-0.005	0.023	-0.102**	-0.139***
	(0.018)	(0.016)	(0.042)	(0.037)
Other Crimes		1.508***		1.477***
		(0.236)		(0.238)
Observations	132.0	132.0	132.0	132.0
R2	0.167	0.37	0.357	0.511
Adjusted R2	0.141	0.345	0.321	0.479
Residual Std. Error	0.331(df = 127.0)	0.289(df = 126.0)	0.294(df = 124.0)	0.258(df = 123.0)
F Statistic	6.365***	14.82***	9.848***	16.055***
	(df = 4.0; 127.0)	(df = 5.0; 126.0)	(df = 7.0; 124.0)	(df = 8.0; 123.0)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5: Regression Discontinuity 3

	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.219***	1.076***	0.916***	1.142***	-0.033	-0.36*
	(0.044)	(0.194)	(0.061)	(0.172)	(0.268)	(0.199)
Treatment 1	0.358***	0.307***	-0.227**	-0.249**	0.261***	-0.496***
	(0.062)	(0.08)	(0.107)	(0.116)	(0.073)	(0.09)
Treatment 2	-0.231***	-0.312***	-0.661***	-0.646***	-0.136	-0.514***
	(0.077)	(0.095)	(0.095)	(0.1)	(0.092)	(0.077)
Linear Trend			0.011***	0.012***		0.016***
			(0.002)	(0.002)		(0.002)
Interest Rate		-4.422		-5.477	-2.114	-2.655
		(5.754)		(5.083)	(5.204)	(3.826)
GDP		-167.77		-189.216	-98.952	-101.753
		(328.272)		(289.845)	(296.197)	(217.755)
Unemployment		0.025		-0.023	0.033*	-0.028**
		(0.019)		(0.019)	(0.017)	(0.014)
Other Crimes					1.377***	1.892***
					(0.251)	(0.191)
Observations	132.0	132.0	132.0	132.0	132.0	132.0
R2	0.206	0.232	0.397	0.406	0.381	0.668
Adjusted R2	0.193	0.202	0.383	0.378	0.351	0.649
Residual Std. Error	0.321(df = 129.0)	0.319(df = 126.0)	0.281(df = 128.0)	0.282(df = 125.0)	0.288(df = 125.0)	0.212(df = 124.0)
F Statistic	16.686***	7.63***	28.069***	14.266***	12.832***	35.676***
	(df = 2.0; 129.0)	(df = 5.0; 126.0)	(df = 3.0; 128.0)	(df = 6.0; 125.0)	(df = 6.0; 125.0)	(df = 7.0; 124.0)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 6: 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

Figure 12: Investigations of Corruption Crimes per State

Table 7: Panels Estimations with States Fixed Effects

Dependent variable:	DF		SP		PR	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-2.365***	-2.368***	-1.948***	-1.913***	-1.662***	-1.615***
	(0.069)	(0.070)	(0.062)	(0.012)	(0.064)	(0.064)
Dummy 1	0.412***	0.4059***	0.059	0.152**	-0.879***	-0.654***
	(0.061)	(0.064)	(0.057)	(0.060)	(0.069)	(0.074)
Dummy 2		0.020		-0.331 ***		-0.826***
		(0.068)		(0.015)		(0.129)
Observations	4350.0	4350.0	4350.0	4350.0	4350.0	4350.0
Pseudo R2	0.1349	0.1349	0.01333	0.1123	0.07595	0.08574
Note:				*p<0.1; **p<0.05; ***p<0.01		