DISCUSSION PAPER SERIES

DP14634

CORRUPTION AND EXTREMISM

Tommaso Giommoni, Massimo Morelli and Antonio Nicolò

PUBLIC ECONOMICS

ISSN 0265-8003

CORRUPTION AND EXTREMISM

Tommaso Giommoni, Massimo Morelli and Antonio Nicolò

Discussion Paper DP14634
Published 17 April 2020
Submitted 15 April 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CORRUPTION AND EXTREMISM

Abstract

When should we expect an opposition group to select an extremist leader or representative? This paper shows the important role of corruption for this choice. Moreover, we show an important asymmetry in the role of corruption, in that the effect on extremism exists only within the opposition group. When the elite has greater ability to use corruption to obtain a better bargaining outcome from the opposition group leader (political corruption), then the equilibrium selection of group leader is more likely to be extreme. On the other hand, the perception of an existing rent extraction by the elite in power may determine the opposite effect within the majority group. We provide strong evidence for these novel predictions using the random audits data in Brazil as exogenous corruption signals, verifying that only within the opposition (to state-level incumbents) the signals determined an extremism drift in voting. Finally, we extend the analysis to extremism and conflict risk in divided countries.

JEL Classification: D72, D73

Keywords: Corruption, extremism, Bargaining, delegation, agency

Tommaso Giommoni - giommoni@kof.ethz.ch
ETH Zurich

Massimo Morelli - massimo.morelli@unibocconi.it
Bocconi University and CEPR

Antonio Nicolò - antonio.nicolo@manchester.ac.uk
University of Manchester, University of Padova

Corruption and Extremism*

Tommaso Giommoni†, Massimo Morelli‡, Antonio Nicolò§

April 14, 2020

Abstract

When should we expect an opposition group to select an extremist leader or representative? This paper shows the important role of corruption for this choice. Moreover, we show an important asymmetry in the role of corruption, in that the effect on extremism exists only within the opposition group. When the elite has greater ability to use corruption to obtain a better bargaining outcome from the opposition group leader (political corruption), then the equilibrium selection of group leader is more likely to be extreme. On the other hand, the perception of an existing rent extraction by the elite in power may determine the opposite effect within the majority group. We provide strong evidence for these novel predictions using the random audits data in Brazil as exogenous corruption signals, verifying that only within the opposition (to state-level incumbents) the signals determined an extremism drift in voting. Finally, we extend the analysis to extremism and conflict risk in divided countries.

Keywords: Corruption, Extremism, Bargaining, Delegation, Agency.

JEL codes: D72, D73.

*We wish to thank Marina Calculli, Marco Celentani, Livio Di Lonardo, Guido Tabellini, Yotam Margalit, Moritz Osnabrüegge, Scott Tyson, as well as seminar participants at Bocconi, ASSET 2018, SPSA, EPSA, Galatina workshop, York University, University of Sussex, Carlos III and CUNEF for useful comments. Special thanks go to Arieda Muço and Sergio Galletta, who shared with us very useful data. Massimo Morelli Acknowledges support from the European Research Council, ERC advanced grant 694583. The usual disclaimer applies.

†ETH Zurich and Dondena.

‡Bocconi University, IGIER and CEPR.

§University of Padova and University of Manchester.

1 Introduction

This paper aims to clarify two important political consequences of corruption: (1) corruption generates extremism in politics and can even trigger conflict in ethnically divided countries; (2) the extremism consequences are asymmetric: the extremism effect operates on the opposition forces only, while the majority or government supporters do not have any incentive to counter corruption with an extreme agency choice.

Corruption may be used in principle to accommodate the opposition and conciliate groups by offering private privileges in exchange for a political green light. This has been sometimes proposed as a reason to believe that corruption could serve the purpose of *greasing the wheels* (Tanzi 1998), rather than *sanding the wheels* (Mauro 1995). What our model shows is that such a conclusion would be wrong, because one needs to factor in the incentives of the opposition group. When members of an opposition group do not trust their political leader to resist bribes from the elite (what we call political corruption), they have an incentive to switch to a more extreme leader (in terms of ideology, religion, or any other dimension of this kind) to counterbalance the corruption temptations. Politically corrupt leaders accept agreements with the elite for their own personal interests and perpetuate the status quo, hence an opposition group that desires change may choose a more extreme and hence likely less corruptible figure. "Intrinsically" motivated leaders are ideologically driven and have more incentive to replace the elite in power to impose their own ideology. Religion is a good example of this strategic use of extremism. Minority groups may choose to be represented by religious leaders who want to impose their religion not because minorities necessarily share a religious view, but because religious extremism reduces the incentive to reach a corrupt agreement with the elite.[1] The bulk of the paper will be devoted to the role of political corruption in creating incentive for opposition groups to select extreme representatives in political offices, but at the end of the paper we will discuss the applications of the model to civil conflicts in divided countries.

[1]The Shining Path in Peru, the Tamil Tigers in Sri Lanka, Maoists in China, the Lord's Resistance Army in Uganda, and the Revolutionary Armed Forces of Colombia all espoused more radical ideas and goals than the average citizens in their countries. See Walter (2017) for a discussion and evidence of the systematic tendency to choose extremists as leaders. See also Jackson and Morelli (2007) for a similar form of endogenous bias in delegation even when looking at international conflict.

As a clear example of political corruption, consider the election of a governor who controls a state budget and can use some transfers in order to weaken the opposition on some policy that she cares about (and the voters of her ideology who can reelect her care about too). Suppose a voter receives a signal that the party she was considering to vote for at the next gubernatorial elections has been involved in corruption cases (even if at another political level). If this party is in power at the state level, then corruptibility would mean that they would be willing to *"give"* bribes in order to push their agenda; hence corruptibility might be a good signal for the voter if he primarily cares about that agenda at the state level. On the other hand, if at the state level this party is in the opposition, an upward update of corruptibility implies an upward updated belief that the representative will *"accept"* a bribe in exchange for weakening her position in the determination of state level policies, and in this case the voter should be tempted to vote for a more extreme representative, in the hope that she will be less corruptible and more determined to fight for the state level policy goals that the voter cares about. Thus, if the perception that corruption is widespread increases, the reaction of voters may well be asymmetric, with a more likely drift to extreme representatives for those on the receiving side than for those on the giving side.

The left-right polarization is only one of the many divides that we study in political economy, and we will use this one in the empirical analysis because extremism is easier to define and measure on the left-right dimension than in other domains. However, the same type of asymmetric impact of corruption on extremism can be conjectured also for other domains: for example, in a country divided along ethnic or religious lines, with a majority and a minority group, often occupying different subsets of the state territory, the majority and the minority leaders have to negotiate on all resource allocation issues, the biggest of which is typically the distribution of the surplus from natural resource extraction and exports. Conflict risk is always related to the way the majority divides such a surplus, and the surplus distribution that the opposition leader accepts in the bargaining process may be affected by favors or bribes that she could receive and hide.

Given the generality of the type of insight that we aim to provide, we propose a simple theoretical model that can apply to both types of contexts, and then we provide some evidence in support of the findings. In order to have a model that could easily be applied across contexts, we voluntarily avoid introducing institutional

details: the insights obtained from our abstract bargaining and agency model could be developed further for any application one might prefer to focus on.

We empirically test the main predictions of the model using Brazilian random audits data as a source of exogenous information shocks on corruptibility. We can check the impact of news on corruptibility of parties at the municipal level on the willingness of voters of that ideological side to support a more extreme party at the state elections. The audit program aimed at monitoring local public finances, enforced by the Brazilian federal agency *Controladoria Geral da União* (CGU), involves a random selection that allows us to estimate the causal effect of being audited (Ferraz and Finan, 2008). It is also possible to quantify the amount of rent-seeking detected by the auditors and to measure this as the fraction of misused public resources. We study the impact of such -municipal- valence shock on the electoral results at a higher administrative level, the state elections. In particular, we focus on the electoral consensus toward extremist parties in the 2006 Brazilian gubernatorial elections. We are particularly interested in studying whether the effect of this type of shocks is different between cities whose municipal government is aligned with the state administration and those where it is not aligned.

In terms of estimation strategy, we take advantage of the way the audit program is designed and we exploit the random timing of these scrutinies. In particular, we follow Ferraz and Finan (2008) and we compare electoral outcomes of cities whose audit reports have been made available before the 2006 gubernatorial elections (treatment group) and those where the audits release only occurred after the polls (control group): since municipalities were selected randomly, cites receiving the information shock after the election represent a valid control group. To conduct this analysis we assemble a large dataset: firstly, we make use of data on municipal audits, obtained from Brollo et al. (2013), which include information on audits timing and main outcomes. Secondly, we collect detailed information on municipal politics and gubernatorial elections for the time span under investigation. Finally, we have details on Brazilian parties position in the ideological spectrum, over time, to be able to classify them in terms of political extremism.

The main results of the analysis confirm the predictions of our model. On the one hand, the release of information about corruption does not seem to raise the consensus of extremist parties in the 2006 gubernatorial elections on average. On the other hand, when we consider the alignment of the municipal government with

the state administration, we uncover large differences. In particular, we find that the occurrence of a valence shock leads to a significant increase in extremist parties consensus in cities where the administration is not aligned with the state government, *i.e.* only in cities to the opposition in their state. The effect is large and statistically significant, as one standard deviation increase in the level of detected corruption raises, in non-aligned cities, extremist parties votes by 9.7%. Moreover, this effect is robust to the use of various measures of corruption as well as different types of dependent variables. These results validate the main predictions of the model and they represent an important causal evidence that political alignment across administrative levels mediates the effects of corruption scandals.

After the theoretical section and the empirical analysis mentioned above, we will discuss the potential applications to ethnic politics. A case that fits our theory well is the case of Hezbollah: one can time the birth and political rise of Hezbollah in Lebanon as consequent to the corruption evidence on Amal in the early 1980's, at the time when Israel obtained greater room of maneuver in the south of Lebanon also due to some allegations of collusion or corruption of Amal itself. Hezbollah movement (and later party) soon after obtained all the support from the Shia population and later from other sections of the lebanese society as well, on the basis of their perceived moral integrity in terms of the political negotiations on what mattered the most to the Shia population. The support of the Shia population of Lebanon for Hezbollah has been neither related to an inherent extremism of the population itself (only 13 percent of Shiites in Lebanon are in favor of the creation of an Islamic State) but rather it is related to the greater reliability of Hezbullah as an agent for resistance, the most credible agent of Shia interests in the relations with Israel.[2]

Intuitively, the selection of extremists as opposition leaders in countries divided in ethnic or religious groups can lead to an increase in real conflict risk, and we will show that indeed there is positive correlation between political corruption and risk of ethnic conflict.

[2]Hezbollah's ideological and regional commitment are perceived as much stronger by the Shiite base also due to the much greater efforts made by Hezbollah to guarantee more welfare and poverty relief to the south of Lebanon, previously almost ignored by the Beirut government even in public funding and surplus sharing.

2 Related literature

The paper relates to the recent and growing literature studying the relation between corruption and consensus to extremist and populist parties. There is evidence that exposure to corruption may have a long term impact on the votes for populist forces (Aassve et al. 2018), and some recent results also on short term effects (Sanz et al. 2020). Moreover, many studies underline the importance played by corruption in the political agenda of extremist and populist politicians (Hanley and Sikk 2014, Curini 2017). Our study is the first to provide a causal evidence that the release of a scandal may favour extremist forces, but this relation only emerges in cities where the parties managing the city are part of the opposition in their state.

A second related literature is the one focusing on corruption audits. Ferraz and Finan (2008) show that the disclosure of corruption scandals reduces the incumbent consensus, while Ferraz and Finan (2011) demonstrate that electoral accountability reduces corruption as rent-seeking is lower in cities where mayors have re-election incentives. There is evidence that monitoring leads to a reduction in future corruption for audited cities (Avis et al. 2019, Bobonis et al. 2016). Moreover, the implementation of an audit has effects on candidates selection (Cavalcanti et al. 2019) and there are spillover effects across different types of elections and municipal borders (Muço 2018). Our paper provides a causal evidence that political alignment across jurisdictions modifies the impact of an audit, and this can crucially affect the electoral outcomes.

In comparative politics, before Walter (2017) no paper or book seriously discussed the prevalence of endogenous ideology. There are very few papers on ideology and civil war in general, and they are all about the mobilization role of ideology - see e.g. Costalli and Ruggeri (2015), Iannacone and Berman (2006), Gutierrez, Sanin and Wood (2014). Beside the above small literature on ideology and conflict, our paper also relates to the literature on the relation between corruption and civil war. Neudorfer and Theuerkauf (2014) present robust findings for the positive effect of corruption on the risk of ethnic civil war and argue that corruption increases the risk of large-scale ethnic violence. There is also some (scarse) empirical evidence on the relation between corruption and political extremism: Henderson and Kuncoro (2011) show the relation between corruption and the electoral outcomes of the Islamic party in Indonesia; Chayes (2016) explicitly link corruption with institutional

failure, terrorism, and religious extremism. A recent research by Transparency International (2017) illustrates by means of case studies how corruption strengthens extremist groups (ISIS's use of corruption in recruitment narratives; Iraq; the rise of Boko Haram in Nigeria).

3 The model

In this section we describe a model of bargaining with endogenous selection of negotiating agents, which we view as a general class of situations where we can study the impact of corruption on extremism and the asymmetric impact of corruption. We will mention at the end how this general but simple model applies in the multiple applications that we consider most relevant.

An opposition group (a party, an ethnic group or a faction) and a majority group have to agree on how to share an amount of resources R.[3] Both groups are risk neutral and elect a representative. The representative of the majority m makes a proposal to the representative of the opposition group o: a non negative amount $X \leq R$. The representative of the opposition either accepts or refuses the proposal. If the proposal is accepted, it is implemented and the game ends. If the proposal is rejected, a political turmoil arises. The political turmoil (if it occurs) determines group-specific costs c_k, $k \in \{o, m\}$. With some probability $(1-p)$, the turmoil implies a regime change that is beneficial to the opposition. For simplicity we can assume that the opposition leader rises to power and her group gets the entire amount or resource R while, with complementary probability p, m keeps the power and does not need to negotiate how to share R, which is then entirely distributed to the majority group.

We assume that $R > \frac{c_o}{1-p}$, so that the expected benefits of a political turmoil are larger than the costs if the opposition group receives a zero offer.

A leader's utility is constituted by (1) the utility of her group (the resources the group receives minus the cost of political turmoil if it occurs), (2) the private benefits that a (corrupt) leader may obtain if is on the receiving side, and (3) her ideology v, representing the non-monetary utility of being in power (due e.g. to

[3]Think of R as being the value of the entire distributable surplus of a country or the total oil and gas revenue of a state; or the amount of public funds to be allocated among different districts or for the production of different public goods when majority and opposition care differently among these public goods.

the different goods or policies that one can produce when in power). A leader with $v = \infty$ is a "fanatic": for a fanatic leader the expected utility of being or remaining in power trumps any amount of monetary utility that he can obtain, either for herself or for her group.

We consider the following simple bargaining subgame once leaders have been chosen: the majority leader m makes a "take it or leave it" offer (X, t) to the leader of the opposition group o, where t denotes a private bribe going directly to her; the receiver of the offer accepts or rejects. While this bargaining subgame is very simple, the heuristic power comes from the fact that each group, anticipating the expected continuation payoffs of the bargaining subgame, selects the leader on the basis of its perception of her extremism.

m's utility in case of agreement is

$$U_m(X, t) = R - X - \gamma(t) + v_m,$$

where $\gamma(t) : \mathbb{R}_+ \to \mathbb{R}_+$ is a function describing the cost of political corruption, i.e., the cost of bribing the leader of the opposition group. We assume that $\gamma(0) = 0$ (no bribing is costless) and that γ is weakly increasing. This function does not only capture the direct monetary cost of bribing a single leader (which may be paltry compared to the amount of resources to offer to the entire minority group to reach an agreement) but also any other non-monetary cost that the act of political bribing may generate. v_m stands for the non-monetary (ideological) value of keeping the power for the majority leader.

In case of conflict m's expected utility is $p(R + v_m) - c_m$, where c_m is the cost of conflict for the majority group. We assume that $R \geq \frac{c_m}{p}$, so that $pR > c_m$ (also the majority group prefers a political turmoil rather than concede all the surplus to the opposition group).

The time line of the whole game is as follows:

1. Assuming that candidates to leadership always exist with perceived extremism levels in the interval $[0, V]$, the majority group and the opposition group first choose their leader types, v_m, v_o;

2. The true type of the leader for group k is realized: $\hat{v}_k = v_k$ with probability $1 - \frac{v_k}{V}\varepsilon$ and $\hat{v}_k = \infty$ with complementary probability, with $\varepsilon < \frac{1}{2}$.

3. If the opposition leader turns out to be a fanatic, then political turmoil occurs and the game ends; if o is not a fanatic, then leader m observes the type of leader o and makes her an offer – a pair (t, X), $t \in \mathbb{R}_+$ and $X \leq R$.

4. The opposition group leader either accepts or rejects the offer. With probability ρ leader o is honest and rejects any offer that contains a bribe $t > 0$. The honesty of o is not observable by m at the time she makes the proposal.

5. If leader o rejects (for whatever reason, honesty or calculus), a political turmoil between the opposition group and the majority arises and the majority wins with probability p. If the proposal is accepted by o, then it is implemented and the game ends.

The cost of bribing the political leader of the opposition group, $\gamma(t)$, could depend on the institutional settings to prevent and fight corruption, like the degree of freedom of the press or the independence of the judicial system from political power. For simplicity, we assume that there exists a finite amount $T > 0$ such that $\gamma(t) = 0$ for all $t \leq T$ and $\gamma(t) > R$ for all $t > T$. Therefore m never finds profitable to offer a bribe larger than T. Hence T will be our simple measure of *corruption*.[4]

4 The prediction: corruption and opposition extremism

In this section we evaluate the impact of corruption on (1) the level of extremism chosen by the opposition and the majority groups when selecting their representative leaders and (2) the risk of political turmoil. We obviously solve the game by backward induction. Once groups have chosen their leaders, m makes its proposal, which is accepted by o only if it is preferred to entering in political turmoil. By assumption any bribe $t \leq T$ is costless for m, while larger bribes are too costly. Therefore if m offers a bribe, she offers a bribe equal to T. The minimum share of the surplus that a corruptible o accepts is

$$X_{\min} = (1 - p)(R + v_o) - c_o - T. \tag{1}$$

[4]See Appendix 1 to see that our results continue to hold if we assume a more general convex cost function $\gamma(t)$.

We make the following assumptions:

Assumption 1: $\rho \leq \frac{T}{T+c_o+c_m+V(1-p)}$.

This assumption says that the probability that a politician is incorruptible is negligible, and it guarantees that for any non-fanatic majority leader (and consequently for the majority group itself) it is always preferable to try to bribe the opposition leader.[5]

The trade-off that the opposition group has to face is the following: given a certain level of political corruption T, a higher level of leader's extremism increases the amount of resources offered to the minority $X_{\min}$ (see (1)), but it also increases the probability that the leader will be a fanatic, who chooses political turmoil even in case a conflict is not the best response for the opposition group. The higher is the level of political corruption T, the lower is the opportunity cost of conflict for the opposition group, (because a peaceful agreement is less profitable), and the higher will be the chosen level of leader's extremism. The effects are strictly increasing below a certain threshold

$$\hat{T} \equiv (1-p)(V-v_m) - c_m - c_o. \tag{2}$$

For the majority group instead the reasoning is very different: a more ideologically driven majority leader cares more to remain in power than a less extremist one, so it is more willing to give resources to the opposition group. Moreover, a fanatic majority leader is willing to offer $X_{\min} + T$ because she wants to minimize the probability of a conflict, even if the probability of having an incorruptible opposition leader is so negligible that for the majority group it is not profitable to make such an offer. It follows that the majority group does not want to select a leader type such that is an extremist with positive probability, in particular when the level of corruption is high.

We will denote v_o^* the level of leader's extremism chosen in equilibrium by the opposition.

[5] To see this, consider a negotiation between a majority leader who has the highest non-monetary incentive to avoid any political turmoil ($v_m = V$) and an opposition leader who only cares about the resources allocated to his group ($v_o = 0$): if corruptible, o accepts any offer larger than $(1-p)R - c_o - T$, while an honest opposition leader accepts any offer larger than $(1-p)R - c_o$. The majority leader prefers the former offer to the latter if $(1-\rho)(R-(1-p)R+c_o+T)+\rho(pR-c_m) \geq R-(1-p)R+c_o$, that is if $\rho \leq \frac{T}{T+c_o+c_m+V(1-p)}$.

Proposition 1 *The majority group always selects $v_m = 0$. The extremism of the opposition group and the probability of political turmoil are (weakly) increasing in the level of corruption. The equilibrium choices of the opposition are the following:*

1. *If $T \geq \hat{T}$, then $v_o^* = V$ and prob of war is equal to ε;*
2. *If $T < \hat{T}$ then*
   (a) *equilibrium extremism is*

$$v_o^* = \frac{c_m + c_o + T}{(1-p)},$$

   increasing in T; and
   (b) *the probability of war is equal to $\frac{c_m+c_o+T}{(1-p)}\frac{\varepsilon}{V}$, which is increasing in T.*

Proof: See Appendix 1.

To reiterate, both opposition's extremism and probability of political turmoil are (weakly) increasing in the level of corruption T. Ideological extremism provides a non-monetary payoff to the opposition leader that makes her tougher in the negotiation on the sharing of total resources, forcing the majority leader to make a more generous offer in order to avoid political turmoil.

For the majority group the argument is almost the opposite. When the level of corruption is large, the majority group prefers a leader m who bribes the opposition leader o, and therefore does not want to have an extremist leader who does not bribe o due to the worry that o could be honest. When the level of corruption is low and the opposition group chooses a leader with some positive level of ideology (as a commitment device) to extract the surplus from the majority group, having $v_m = 0$ ensures to the majority the same payoff as it gets in case of conflict, while having $v_m > 0$ gives a lower payoff than conflict in case of agreement.

In Appendix 1 we also include a second proposition showing that the results hold also when using a more general convex cost function $\gamma(t)$, and in that case an increase in corruption can be captured by a reduction of the marginal cost.

5 Empirical analysis

In this section we provide evidence in support of the asymmetric corruption effects on extremism derived in the above model. The random corruption audits at the municipal level in Brazil offer voters of an audited city a signal about the corruptibility of the parties managing their city. If our theory is correct, any such signal could convince a voter of the same ideology of that local government to select a more extreme representative for state level elections, but only if the voter's "side" in the political spectrum is not controlling the state at the moment.

5.1 Institutional background

5.1.1 Local Institutions and elections

Brazil is a federal republic divided into 5,570 municipalities and 27 federal units, which include 26 states and the Federal District (where Brasilia is located). The executive and the legislative branches of each administrative levels are chosen by citizens through direct elections. The different levels of government, states and municipalities, are important providers of public goods (e.g., primary and secondary education, health care, housing, transportation, local police and local infrastructure). Moreover, local authorities are important collectors of fiscal revenues: state tax revenues derive mostly from sales taxes, while most important municipal taxes are property and service taxes. Further, transfers from upper levels of government cover a significant proportion of these expenditures: large municipalities can generate up to 40 percent of their revenues with taxes, while smaller municipalities (those with less than 50,000 inhabitants) receive around 80 percent of their revenues via federal transfers (Farvacque-Vitkovic and Kopanyi 2014).

At the state level, the executive branch is headed by a governor (*Governador*) who is elected directly via a run-off system every four years,[6] and faces a two-term limit. The state assembly (*Assembleia legislativa*) is elected at the same time as the governor. At the municipal level, the mayor (*Prefeito*) has executive power while the city council (*Câmara de Vereadores*) holds the legislative one.[7] The term lasts

[6]Participation in elections is compulsory in Brazil for literate citizens aged between 18-70. It is non-compulsory for Brazilians of age lower than 18 or higher than 70 or illiterate citizens of any age.

[7]The number of seats in the city council depends on the population of the municipality. The number ranges from a minimum of 9 in cities with population lower than 15,000, to a maximum

for four years and mayors face a two-term limit, differently from councillors who are not constrained. The mayor is the head of the local administration and the council has the duty of overseeing the local government as well as voting the municipal resolutions. Municipal polls elect the mayor and the city councillors, and the local government coalition has to form a majority.[8]

Brazil has a multi-party system at the federal, state and municipal level: candidates at any administrative level must be supported by a party or a coalition of parties. Many parties participate in local and state elections.[9]

5.1.2 The Brazilian anti-corruption program

The municipal audits program was initiated in 2003, with the goal of fighting corruption in local governments and increasing transparency of public spending. The program is implemented by the Brazilian federal agency *Controladoria Geral da União* (CGU) and consists in the random selection of municipalities to be audited, with the exception of municipalities with more than 500,000 inhabitants.[10] The examination mostly covers the transfers allocated by the federal government to each municipality. The lottery takes place every two or three months in the *Caxia Econômica Federal*, in Brasilia: excluding the pilot audits in 2003, where one municipality per state was audited (23 municipalities), the number of selected municipalities raised from 50 to 60 per round after 2005.

In selected municipalities, a team of auditors is sent with the goal of inspecting the use of federal transfers, following the guidelines of CGU. In particular, the types of funds under inspection are also randomly selected and communicated jointly with eligible municipalities. After the investigations, the auditors complete a report describing the full details of the irregularities found which is publicly released within a short period of time. There is large anecdotal evidence suggesting that local media widely cover audits' outcomes (Muço 2018), and there is robust evidence that the coverage of these audits crucially affects their impact (Ferraz and Finan, 2008).

of 55 in cities with more than 8 millions inhabitants

[8]The method is different for municipalities with more than 200,000 inhabitants, whose municipal electoral system requires a second round.

[9]Only few parties are relevant at the federal level. In particular, only two parties were dominant in presidential elections since 1994, *Partido da Social Democracia Brasileira* (PSDB) and *Partido dos Trabalhadores* (PT).

[10]The lotteries are run independently for each state, therefore the probability for a municipality to be selected in a given year varies by state.

5.2 Data

To conduct our empirical analysis we assemble different datasets, covering the gubernatorial elections of October 2006, which took place during the second year of the 2005–2008 municipal electoral term.

5.2.1 Municipal Corruption Data

First, we make use of data on municipal corruption, obtained from Brollo et al. (2013). This dataset contains a set of corruption measures for the first 29 lotteries covering the period July 2003–March 2010. The dataset includes 1481 municipalities that have been randomly selected for the audit, the majority of them, 1396 municipalities, have been audited only once, while a small group, 85 municipalities, received two audits in this time span.[11] In the empirical analysis, we focus on a subset of these audits as we only include those that took place within a window of one year from the 2006 gubernatorial elections: cities audited in the pre-election period (from October 2005 to September 2006) represents the treatment group and those monitored in the post-elections months (From October 2006 to September 2007) are part of the control group.

To measure corruption we use a narrow definition, in order to capture only illegal practices, excluding cases of bad administration that not necessarily imply corruptive phenomena. This measure includes the following categories of irregularities: (i) severe illegal procurement practices; (ii) fraud; (iii) favouritism; and (iv) over-invoicing.[12] More precisely, in order to account for the intensity of detected corruption, we use the variable *Narrow corruption*, which measures the fraction of misappropriated funds on the total amount audited at the level municipality-term.[13]

[11] In these few cases, we only consider the outcomes of the first audit. We follow this approach as there is evidence that being audited significantly influences the behaviour of the political actors, affecting the outcome of subsequent audits in the same cities, as documented by Avis et al. (2018) and Bobonis et al. (2016).

[12] The measure we use to capture corruption is called *Narrow* in the original dataset by Brollo et al. (2013). The dataset contains an additional indicator of corruption, *Broad* corruption, measuring the phenomenon in a broader sense. More precisely, this second measure includes the following violations: (i) general illegal procurement practices; (ii) fraud; (iii) favoritism in the good receipt; (iv) over-invoicing; (v) diversion of funds; and (vi) paid but not proven. Many irregularities in the categories "diversion of funds", "paid but not proven" as well as some "general illegal procurement practices" do not necessarily imply corruption as they may result from poor public management. Nevertheless, we conduct the analysis also with this measure of broad corruption and the main results emerge also in this case.

[13] The original name of the variable in Brollo et al. (2013) is *Narrow fraction of the amount*

Table 3 in Appendix 2, Panel A, shows the descriptive statistics of the corruption measures used in the analysis. Our main measure of corruption, *Narrow corruption*, has an average value of 0.013, suggesting that an audit discovers, on average, a fraction of 1.3% of misappropriated funds. Instead, the other measure of corruption, *Broad corruption*, is larger by construction, with an average values of 0.044. Besides this continuous measure of corruption, we also use two additional measures, namely a variable that captures when the level of corruption detected by the audit is above, respectively, the median level and the fourth quartile amount, both based on the *Narrow corruption* version.

5.2.2 Political Data

Second, we collect data on local politics considering the different administrative levels under analysis: municipalities and states. We collect information on incumbent mayors and governors, such as party affiliation and coalition of parties. Moreover, we gather detailed data on municipal and gubernatorial elections, including information on party affiliation and coalition of candidates, their vote share and the number of valid votes. To be able to classify parties in the ideology spectrum and in terms of extremism level, we rely on the categorization by Power and Rodrigues-Silveira (2019). They make use of the Brazilian Legislative Surveys (BLS), which are based on written questionnaires administered to members of the National Congress, to construct a dynamic classification of Brazilian parties. In this manner, they manage to locate each political party in a scale which ranges between -1 (left) and +1 (right). We rely on this classification to characterize party extremism for gubernatorial elections of 2006: in particular, we define *extreme* the party p if:

$$I_p > |0.5|$$

where I_p is the ideology score of party p, according to Power and Rodrigues-Silveira (2018). In other words, we are capturing the left and the right tail of the ideology

This approach is similar to Cavalcanti et al. (2019) as we use an indicator of the severity of corruptive phenomena instead of a binary measure. The intensity of illegal practices may be more informative than the presence of corruption per se. An important assumption that we make is that when an audit yields detected corruption at the municipal level this makes voters of that municipality update their beliefs about corruptibility also of the corresponding party candidates for the state elections.

distribution, considering the limit value of 0.5.[14] We use in the analysis two measures of extreme voting, i) the vote share to extreme parties and ii) the difference in vote share to extreme parties, compared to the last gubernatorial elections of 2002. It is important to note that our main definition of extremism does not distinguish between extreme left or right. We account for that as we conduct a series of additional analysis to study whether the effect is differential between votes towards extreme left or right parties.

Furthermore, as we are interested in studying the effect on voters whose local government is not aligned with the state government, we compute a measure of political mis-alignment across different administrative units, the municipality and the state. To do so we define a city as not-aligned with the state if the party controlling the municipal government is not included in the coalition that supports the governor in the corresponding state.[15]

Table 3, Panel B, shows the descriptive statistics for all these political variables for municipal and gubernatorial elections. *Extreme vote share - Difference* and *Extreme vote share - Levels* are the main dependent variables that capture the consensus to extreme parties in gubernatorial elections of 2006: the latter variable has an average value of -0.276, suggesting that, on average, the support toward these forces has decreased between 2002 and 2006; while the former variable has an average value of 0.202. Moreover, the fraction of municipalities not aligned with the state government is 61.3% according to our categorization (*Not-aligned - party*). Finally, it is important to note that the majority of Brazilian cities has a right-wing municipal administration, the 74.4%, and that the average turnout is very high, 79.7%.

Finally, we collect a series of additional municipal information, that we draw from the 2000 and 2006 Brazilian census. In particular, we collect data on the population, the monthly per-capita income, the share of employed and the average education level. The descriptive statistics of these variables are shown in Panel C

[14] We also generate an alternative measure of ideology using the limit values of 0.55, 0.6 and 0.65, and the main results are confirmed.

[15] We also provide an alternative definition of mis-alignment which is based on the left/right wing division. In particular, we rely on the classification by Power and Rodrigues-Silveira (2018) to classify municipal and state governments between left-wing and right-wing and we define a city as not-aligned with the state government if the local majority does not share the same ideology with the state government. This measure is less accurate than the one based on the governor coalition as it leads to a number of wrong classifications. However, the results using this alternative measure of alignment are similar to the main ones, these outputs are not shown for the sake of space.

of Table 3.

5.3 Empirical strategy

5.3.1 Estimation

Our goal is to study the effect of corruption disclosure, which represents an information shock, on voters' behaviour in gubernatorial elections. More precisely, we want to analyse i) whether this type of information shock affects the propensity of voters to support extremist parties and ii) if the effect is different depending on whether the municipality is aligned or not with the state government, *i.e.* whether the city is minority or not within its own state.

The random timing of Brazilian audits allows us to study the causal impact of the corruption information shock on extremist voting at the gubernatorial elections of 2006: for any given level of corruption, the comparison of the electoral outcomes in cities that received information on the audit with those where no information was published, represents an estimate of the causal effect of the information release. In particular, we follow the approach used by Ferraz and Finan (2008) in the definition of the treatment and the control group: the treatment group includes cities whose audit reports were made available before October 1st 2006 (*Pre-elections*), while cities in the control group are those where audits' release only occurred after this date (*Post-elections*). Our identification strategy, therefore, relies on the comparison between cities receiving the information shock before 2006 elections with those treated after the polls. Because municipalities were selected at random, the set of cities whose audit release took place after the elections represent a valid control group. An important remark is that the content of the audits may cover irregularities in public funds in different municipal electoral terms. This may represent a problem if we were not able to identify the exact timing of the misuse as, in that case, for the voters may not be easy to identify the responsible political administration. We are able, fortunately, to identify the term where the violation took place.

Figure 1, in Appendix 2, graphically illustrates the timing of the audits release, which divides cities between treatment and control groups: in the analysis we include cities where the publication of the audit took place within a year from the 2006 gubernatorial election, *i.e.* in the interval October 2005–September 2007. We focus on this time window as the municipal administration in office during the 2006

gubernatorial elections began in January 2005. Therefore, we cannot enlarge further our time window without including the old administration which may have a different political alignment with respect to the state government. The auditing rounds included are seven, four before and three after the polls, involving, respectively 234 and 170 municipalities. This approach allows us to consider cities that receive the information very close to the elections, where the shock to cities in the treatment and control group are not very distant in time.

Table 4, in Appendix 2, shows balance tables where we report the summary statistics for the control and the treatment municipalities. The table shows several political, economic and social characteristics of cities. All these variables are balanced between the two groups, with the exception of the corruption measures, *Narrow corruption* and *Broad corruption*. In particular, the amount of discovered corruption is always higher in the control group than in the treatment one, and the difference is statistically significant. However, for the variable *Narrow corruption*, which is the main indicator of our analysis, this difference completely depends on the incidence of lottery released on August 2006 -part of the treatment group-, that detected a very low level of corruption. If we remove cities audited in that lottery the variable becomes balanced between treatment and control group.[16] Nonetheless, we do not think that this outcome can undermine the reliability of our empirical analysis. Indeed, this evidence would be problematic only in case the auditors behave differently based on the timing of disclosure, in particular in reference to the date of 2006 gubernatorial elections, while this would not be an issue if this happened by chance. In support of our interpretation, it is worth mentioning that there is an extensive literature showing that auditors in Brazil do not behave strategically (Ferraz and Finan, 2008; Litschig and Zamboni, 2018) and that auditors do not control the date of publication, as this may happen between 53 and 355 days after the audit (Cavalcanti et al. 2019).

5.3.2 Specification

The baseline specification relies on cross sectional variation as we compare the electoral outcome in 2006 gubernatorial elections between cities in the treatment and

[16]Moreover, the main results of the empirical analysis still hold if we exclude from the sample the cities audited in that lottery. These outputs are shown in Table 7, and are discussed in the "Additional analysis" section.

control group. The estimated model is as follows:

$$y_{ms} = \alpha + \beta_1 A_{sm} + \beta_2 C_{ms} + \beta_3 A_{sm} * C_{ms} + \gamma X_{ms} + \delta_s + \epsilon_{ms} \quad (3)$$

Where y_{ms} is the dependent variable for municipality m in state s that captures the electoral consensus to extreme parties, measured both with the vote share and with the difference in vote share, compared to the previous elections. A_{ms} is the *Pre-election* dummy that indicates whether the publication of the audit occurred before the gubernatorial elections of 2006, C_{ms} captures the intensity of corruption found in the audit, namely it is the share of corrupted resources, either the variable *Narrow corruption* or *Broad corruption*. We make use of different versions of this variable: the simple fraction of detected corruption, the indicator whether the fraction is above the median and above the fourth quartile. δ_s is the set of state fixed effect, necessary in this setting as lottery randomization is stratified at the state level. X_{ms} is a vector of time invariant, municipal-specific, controls that we include to improve the precision of estimates in case the randomization leads to any form of selection between treatment and control group. This vector includes information on the political background (number of voters and turnout of the last municipal election) and on the social background (average income level, share of employed, average education and population in the city in 2006). Moreover, we include a set of dummy variables that capture the distance (in months) of the elections from the audit release, differential across states. We include this covariate in order to control for the length of the exposure to the information shock. Finally, ϵ_{ms} are robust standard errors, clustered at the state level.

The coefficient β_1 captures the causal effect of the publication of the audit results on electoral outcomes. We expect this coefficient to be significant if the audit process *per se* may affect the propensity of voters to support extreme parties in gubernatorial elections. Nevertheless, as discussed by Ferraz and Finan (2008), this is not capturing the fact that the impact of an information shock also depends on the prior beliefs of voters on the incumbent's corruption activity. Therefore, we interact the Pre-election dummy with the measure of detected corruption: the coefficient β_3 represents the average effect of the release of information, conditional on the level of corruption detected by the audit. This coefficients captures whether electoral outcomes also depend on the type of information released rather than the simple disclosure.

Furthermore, in this analysis we are interested in studying whether the impact on extreme voting also depends on the extent to which a municipality is aligned or not with the state government, *i.e.* whether the municipality is controlled by parties that belong to the opposition at state level. To capture this aspect, we refine our baseline specification taking into account the political alignment of municipal administration compared to the state. The estimated model is as follows:

$$\begin{aligned} y_{ms} = \alpha + \beta_1 A_{sm} + \beta_2 C_{ms} + \beta_3 NA_{ms} + \\ + \beta_4 A_{sm} * C_{ms} + \beta_5 A_{sm} * NA_{ms} + \beta_6 C_{ms} * NA_{ms} + \\ + \beta_7 NA_{ms} * A_{sm} * C_{ms} + \gamma X_{ms} + \delta_s + \epsilon_{ms} \end{aligned} \quad (4)$$

Where NA_{ms} is a dummy equal to one when the municipality is not aligned with the state government and it belongs to the minority at the state level[17]. The coefficient β_7, therefore, captures the extent to which the impact of the information shock may vary based on the municipality being aligned or not with the state government, and it allows us to empirically test the main prediction of our model.

5.4 Results

Table 1 shows the main results of our analysis using as dependent variable the difference in vote share to extreme parties (*Extreme vote share - Difference*). As already discussed, we use three measures of corruption: the continuous measure, the indicator of corruption above the median and above the fourth quartile. Columns (1), (3) and (5) of Table 1 show the results for the estimation of model 3. We report, for the sake of space, only the coefficient for the interaction term, that measures the impact of the audit on the dependent variable, conditional on the level of corruption found in the municipality. The disclosure of information just before elections has a mixed effect on the consensus of extreme parties and the effect is not statistically significant for any measure of corruption. Moreover, the same result emerges if we conduct this test using as dependent variable the vote share to extreme parties (*Extreme vote share - Levels*). These results are shown in Table 2, columns (1), (3)

[17]The variable NA_{ms} is constructed using as benchmark the ideology of state government in office during the term 2002-2006, regardless of when the city received the audit. This approach allows us to capture voters' beliefs when they elected the governor in 2006.

and (5): in this case the effect is always positive but not statistically different from zero. These results suggest that the release of information about corruption does not affect the performance of extreme parties at the polls on average.

Most important, we estimate model 4 and we raise the issue whether the alignment of city government with the state administration may affect the impact of the audit. Columns (2), (4) and (6) of Table 1 show these results for the dependent variable *Extreme vote share - Difference*: for the sake of space, we only show the interaction terms. The outputs from this analysis suggest that city political alignment strongly affects the impact of the audits on votes to extremist forces, as the triple interaction term is always positive and statistically significant. This suggests that the release of information about corruption fosters extremist parties only when the municipality is minority into the state as it is not aligned with the state government. The effect is large and very precisely estimated: if we consider the continuous measure, one standard deviation increase in the level of detected corruption, in not-aligned cities, raises the dependent variable by 9.7%, which amounts to about 27.3% of a standard deviation. Further, the effect is also strong for the other two measures of corruption: being above the median level (above the fourth quartile) raises extreme voting by an amount equal to 18.4% (19.1%). Similar, but weaker, results emerge also if we consider as dependent variable *Extreme vote share - Levels*: the impact is always positive but the magnitude of the coefficients is lower. These results are displayed by columns (2), (4) and (6) of Table 2.

Table 1: Impact of corruption on extreme voting – narrow corruption, difference in vote share

Dep. var.: Vote share to extreme - difference	Linear		Quartile		Median	
	(1)	(2)	(3)	(4)	(5)	(6)
Not-aligned*Pre-election		0.0230		0.0172		0.0185
		(0.0433)		(0.0407)		(0.0424)
Pre-election*corruption	-0.396	-2.083***				
	(0.410)	(0.586)				
Not-aligned*corruption		-0.0704				
		(0.643)				
Not-aligned*Pre-election*corruption		2.139***				
		(0.761)				
Pre-election*corruption (above fourth Q)			0.00797	-0.114***		
			(0.0621)	(0.0368)		
Not-aligned*corruption (above fourth Q)				0.0127		
				(0.0341)		
Not-aligned*Pre-election*corruption (above fourth Q)				0.191***		
				(0.0557)		
Pre-election*corruption (above median)					0.0183	-0.103**
					(0.0569)	(0.0422)
Not-aligned*corruption (above median)						0.0120
						(0.0317)
Not-aligned*Pre-election*corruption (above median)						0.184***
						(0.0501)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367	367	367	367	367	367

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections in difference with respect to the 2002 elections. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit – version *Narrow corruption* –, expressed as the percentage of the budget involved in general violations, *corruption (above median)* is a dummy equal to one whether the share of irregularities is above the median level and *corruption (above fourth Q)* is a dummy equal to one whether the share of irregularities is above the fourth quartile. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

Table 2: Impact of corruption on extreme voting – narrow corruption, vote share

Dep. var.: Vote share to extreme - difference	Linear		Quartile		Median	
	(1)	(2)	(3)	(4)	(5)	(6)
Not-aligned*Pre-election		0.00953		0.00457		0.00605
		(0.0217)		(0.0175)		(0.0173)
Pre-election*corruption	0.308	-0.343				
	(0.378)	(0.328)				
Not-aligned*corruption		0.164				
		(0.495)				
Not-aligned*Pre-election*corruption		0.755				
		(0.562)				
Pre-election*corruption (above fourth Q)			0.0395	-0.0202		
			(0.0495)	(0.0304)		
Not-aligned*corruption (above fourth Q)				0.00572		
				(0.0339)		
Not-aligned*Pre-election*corruption (above fourth Q)				0.0970*		
				(0.0516)		
Pre-election*corruption (above median)					0.0365	-0.0202
					(0.0458)	(0.0323)
Not-aligned*corruption (above median)						0.00948
						(0.0319)
Not-aligned*Pre-election*corruption (above median)						0.0875*
						(0.0500)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367	367	367	367	367	367

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit – version *Narrow corruption* –, expressed as the percentage of the budget involved in general violations, *corruption (above median)* is a dummy equal to one whether the share of irregularities is above the median level and *corruption (above fourth Q)* is a dummy equal to one whether the share of irregularities is above the fourth quartile. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

In addition, we also run the main analysis using the alternative measure of corruption, *Broad corruption*. Tables 5 and 6 (in Appendix 2) show the outputs for this analysis using, respectively, *Extreme vote share - Difference* and *Extreme vote share - Levels* as dependent variables. The results are similar to the main analysis conducted on the *Narrow corruption*, although a bit weaker: the disclosure of information on corruption only affects the consensus of extreme parties in case the municipality is not aligned with the state government. And this emerges using both dependent variables.

The results from this empirical analysis seem to validate the main prediction of our model, as we show that the political alignment of municipalities with respect to

the state government is a key determinant for the impact of corruption disclosure on extremist voting. This important heterogeneous result supports our theory and represents the first causal evidence that the reaction to a corruption scandal may be affected by the political connections of a city with higher administrative bodies.

5.5 Additional analysis

In this section we provide additional analysis in order to better understand the main results and to test the robustness of these empirical findings.

Firstly, we want to distinguish between consensus to the extreme left and right parties and we want to differentiate the effect for cities with a left and a right wing administration. Therefore, we estimate model 4 including an additional interaction term that captures the political affiliation of the municipal administration, left-wing or right-wing.[18] Figure 2 in Appendix 2 shows these outcomes. Panel a) shows the results using as dependent variable the votes to the extreme left parties and panel b) to the extreme right parties: in both panels we show the overall effect and the interaction term. What emerges is that the disclosure of corruption leads non-aligned cities with left-wing administration to vote more for extreme left parties (weak effect) and right-wing administrations to vote more for extreme right wing parties. This suggests that the information shock leads voters to radicalise their position rather than changing side.

Secondly, we want to change the definition of extremism in order to study whether there are heterogeneous effects as we select a different set of extremist forces. With this aim, we use alternative definitions of extreme parties in order to increase progressively the degree of extremism of our measure and we apply different ideology limit values: in particular, we only include parties whose ideology score

[18] In particular, we add an additional interaction term PM_{ms} to the empirical model. The estimated model is as follows:

$$y_{ms} = \alpha + \beta_1 A_{sm} + \beta_2 C_{ms} + \beta_3 NA_{ms} + \beta_4 PM_{ms} + \beta_5 A_{sm} * C_{ms} + \beta_6 A_{sm} * NA_{ms} + \beta_7 C_{ms} * NA_{ms} + \beta_8 A_{sm} * PM_{ms} + \beta_9 C_{ms} * PM_{ms} + \beta_{10} NA_{ms} * PM_{ms} + \beta_{11} NA_{ms} * A_{sm} * C_{ms} * PM_{ms} + \gamma X_{ms} + \delta_s + \epsilon_{ms} \quad (5)$$

where PM_{ms} captures the party of the municipal government in city m in state s and C_{ms} is the indicator of corruption above the median level. We estimate two versions: the first one where PM_{ms} captures left-wing governments, the second one where it captures right-wing governments. All other terms are defined as in the main model 4.

is higher than, respectively, 0.55, 0.6 or 0.65. Figure 3 in Appendix 2 shows the main outcomes for the dependent variables difference in vote share: these figures show that the magnitude of the coefficient does not change much as we increase the degree of extremism and we select a smaller set of political parties. This result suggests that the main effect is driven by the most radical forces and that, as we exclude the less extreme ones, the dynamics remains rather unchanged.

Thirdly, we study the heterogeneity of the effect depending on the time of disclosure. In particular, we want to analyse whether the reporting of scandals close to the elections are more effective than those reported far from them.[19] The results are shown in Figure 4 in Appendix 2, where we plot the triple interaction term for the scandals reported in semester one or two. What emerges is that the audits disclosed close to the gubernatorial elections lead to a larger increase in extreme voting, for non-aligned cities, compared to the other ones. But the difference is small and the estimates remain relatively stables. Moreover, the coefficients are both statistically different from zero. This result suggests that the impact of corruption disclosure is stronger the closer it is to the elections.

Finally, we conduct a series of robustness checks to control for the validity of our results, where we make use of the continuous measure of corruption. First, we conduct the main analysis -model 6- excluding every lottery one-by-one. This test allows us to check whether the main results depend on a single specific lottery rather than on the entire sample in analysis. Table 7, in the Appendix 2, shows the outputs: the main results emerge almost always. Second, a potential concern on the empirical strategy may be that the measure of detected corruption can be correlated

[19] We estimate an alternative model where we distinguish the effects based on whether the audit occurred within one semester or two from the elections (in the main analysis we only focus on a time window of twelve months). The model is as follows:

$$\begin{aligned} y_{ms} = \alpha + \sum_{t=1}^{2} \beta_{1,t} A_{smt} + \beta_2 C_{mst} + \beta_3 NA_{ms} + \\ + \sum_{t=1}^{2} \beta_{4,t} A_{smt} * C_{mst} + \sum_{t=1}^{2} \beta_{5,t} A_{smt} * NA_{ms} + \beta_6 C_{mst} * NA_{ms} + \\ + \sum_{t=1}^{2} \beta_{7,t} A_{smt} * NA_{ms} * C_{mst} + \gamma X_{ms} + \delta_s + \epsilon_{ms} \end{aligned} \quad (6)$$

Where A_{sm1} and A_{sm2} represent, respectively, the first semester before the elections and the second semester before the elections and and C_{ms} is the indicator of corruption above the median level. All other terms are defined as in the main model 4.

with other city characteristics. The level of detected corruption, differently for the occurrence of a audit, is not randomly assigned and may be correlated with municipal characteristics, such as capital spending or low level of trust, and this may be a concern for the interpretation of our results. To avoid to capture these differences, we replicate our baseline analysis including an interaction term where we multiply the pre-election dummy, interacted with the non-aligned dummy, by a set of control variables, potentially correlated with corruption level. Namely, we use the following variables: population level (in 2006), income level, average education and share of population in public administration. The main results are shown in Table 8, in the Appendix 2: on the one hand, the interaction term of interest (the triple interaction: non-aligned*pre-election*corruption) remains always positive and statistically significant and the coefficient is stable across the different analysis; on the other hand, the coefficient of the *placebo* triple interaction term is never statistically significant, the only exception is the variable population but this dose not weaken the interaction term of our interest. These results limit our concern suggesting that the measure of corruption is not likely to be a valid proxy for other municipal characteristics.

6 Insights for civil conflict

We believe that our simple theory has uncovered a general asymmetric link between corruption and extremism, beyond the political selection studied for the specific case of Brazil. The six lowest-scoring countries on the Corruption Perception Index (CPI) 2017 data, namely Somalia, South Sudan, Syria, Afghanistan, Yemen and Libya, are all facing intense conflicts, often characterized by violent extremism and terrorism. Chayes (2016) views corruption as an important understudied cause of insecurity, but the rationalist theories of conflict do not address the role of corruption. Our model suggests that a potentially important channel through which corruption may cause more conflict is the derived incentive to select extremist leaders – endogenous extremism as a best response to *political corruption*.

Walter (2017) hints that extremists could be preferred as opposition leaders for a different reason: "... Average citizens may have incentives to join or collaborate with an extreme rebel group if they feel that such a group is more likely to win a war and resist corrupting influences once in power." However, because most minority

groups do not have any chance to obtain full control of power in their country, the reasoning behind a choice to select an extremist as representative cannot be limited to the considerations of what the leader would do if she obtained power: there are already significant agency problems even without any change in the relative power of groups. Corruption of a political leader or representative of a minority group can generally affect the interests of the minority group through negotiations and renegotiations of all kinds of political allocation decisions, resources and policy decisions, changes of entitlements or priorities.[20]

Our model can be extended to include asymmetric information (hence adding classic reasons to observe conflict in equilibrium – Fearon, 1995), and the insights of Proposition 1 continue to hold. In other words, it is a robust finding that corruption causes extremism of the opposition and a higher conflict risk. We can also show (results available upon request) that the awareness by the minority group of the possibility of corruption in negotiations can sometimes also be a cause of terrorist attacks, since terrorism can have similar effects on the negotiations as the choice of an extremist representative. Both these forms of endogenous reactions eliminate the possibility of a "greasing the wheel" effect, and hence political corruption can never be good. The only cases where political corruption can reduce the probability of conflict are (1) when it is not anticipated as a possibility by the minority group when choosing their representative in negotiations; (2) when it is so high that the minority group is completely resigned to repression and avoids delegation altogether (results available upon request).

As far as the prediction of higher conflict risk is concerned, we can take a final look at a comprehensive dataset that covers most countries in the world, 151 nations, for more than 50 years (1965-2017). We measure political corruption relying on the *Varieties of Democracy* dataset that provides a set of time varying, country-specific, indicators. We make use of two indicators: the *Executive corruption* index, capturing corrupt behaviours involving members of the executive power, and the *Legislature corruption* index, focusing on malfeasance perpetrated by the legislative body. For incidence and intensity of conflicts we use the *Political Instability Task*

[20]The first example that may come to most people mind is the allegations of corruption of Arafat, perhaps responsible for the increasing popularity and support of Hamas. The allegations of Arafat's corruption in the negotiation process may have turned Palestinians to support Hamas for multiple reasons, but the primary reason seems to be the fear of not being well represented in the negotiations themselves. See also the discussion of the Hezbollah case in the introduction.

Force Worldwide Atrocities Dataset (Neudorfer, Theuerkauf 2014) and we make use of three indicators: the variable ethnic conflict -dummy equal to one when a country experiences a conflict in a specific year-, the number of rebels fighting and the count of fatalities.[21] Table 9 in Appendix 2 shows the descriptive statistics of this dataset: the two measures of political corruption have similar average values and the probability for a country to experience an ethnic conflict in a given year is around 10%.[22]

To study the relation between corruption and conflict we make use of an estimation strategy that relies on fixed effects and controls.[23] The main results confirm that political corruption is highly correlated with civil war incidence and intensity. Table 10 in Appendix 2 shows these findings: first, columns (1) and (2) focus on the dependent variable ethnic conflict and a positive and significant result emerges without and with controls (Panels A and B). Second, columns (3-6) show these results using as dependent variables the number of rebel fighters and the count of fatalities and the same results emerge. These results are coherent with our bargaining model and the key idea that corruption of political leaders may fuel conflicts, perhaps through an endogenous increase in extremism.

[21]The PITF dataset defines an ethnic civil conflict as an armed dispute between the government and ethnic challengers which result in at least 1,000 direct fatalities over the full course of the armed conflict, exceed 100 conflict-related deaths in at least 1 year and during which each party has mobilized at least 1,000 people, including armed agents, demonstrators, and troops (Neudorfer, Theuerkauf 2014, Marshall et al. 2009).

[22]This indicator is very volatile having an average variation equal to 30%: countries like South Sudan, Myanmar and the Philippines experience ethnic conflicts more than 80% of the years while many countries never experience this type of conflict.

[23]The model that we estimate is as follows:

$$y_{it} = \beta_0 + \beta_1 Corr_{it} + \beta_2' X_{it} + \gamma_{ct} + \delta_i + \epsilon_{it}$$

y_{it} is the dependent variable of the analysis for country i in the year t; $Corr_{it}$ is one corruption indices varying at the country/year level and it represents the main explanatory variable of the analysis; X_{it} is the set of control variables which includes population, GDP, area, elevation (average value and standard deviation), total past conflict onsets and total past years of peace. Moreover, γ_{ct} is the set of continent-year fixed effects and δi are country fixed effects. Then, standard errors are robust and clustered at the country level. The sample includes most countries in the world, 151 nations, for the time span 1965-2017.

The average impact of corruption on ethnic conflict is captured by the variable $Corr_{it}$, which is contemporaneous to the outcome variable. Nonetheless, the same results emerge as we use lagged indicators for corruption ($Corr_{it-1}$, $Corr_{it-2}$, $Corr_{it-3}$), these results are not shown and are available upon request.

7 Concluding remarks

This paper sheds light on important and neglected features and consequences of corruption, studying it from a political economy and agency perspective. While it is common sense that administrative corruption leads to inefficient allocation of resources and unequal access to resources or contracts (*e.g.* see Shleifer, Vishny 1993), what we uncover is the importance and the asymmetric effects of political corruption.

Corruptibility of representatives of opposition groups makes such groups want to switch to more extreme agents. We have shown that this mechanism has been definitely at play in one of the most studied countries for this type of phenomenon (due to availability of randomized data), but we have also hinted that the same general incentive exists in many other domains, including those where political turmoil risk becomes actual deadly conflict risk. In future research one could aim to expand the empirical analysis on civil conflict risk by creating time varying extremism indeces, and looking for a setting where the extremist agency mechanism can be isolated there too.

References

[1] Aassve, A. , G. Daniele, M. Le Moglie (2018). "Never forget the first time: The persistent effects of corruption and the rise of populism in Italy", BAFFI CAREFIN Centre Research Paper.

[2] Avis, E., C. Ferraz, and F. Finan (2018): "Do government audits reduce corruption? estimating the impacts of exposing corrupt politicians," *Journal of Political Economy* 126 (5), 1912-1964.

[3] Bobonis, G.J., L.R.C. Fuertes and R. Schwabe (2016): "Monitoring Corruptible Politicians," *American Economic Review*, 106 (8), 2371-405.

[4] Brollo, F., T. Nannicini, R. Perotti and G. Tabellini (2013): "The Political Resource Curse,' *American Economic Review*, 103(5), 1759-96.

[5] Cavalcanti, F., G. Daniele and S. Galletta (2018): "Popularity Shocks and Political Selection," *Journal of Public Economics*, 165(C), 201-216.

[6] Chayes, S. (2016): *Thieves of State. Why Corruption threatens Global Security,* Norton Publ. New York.

[7] Costalli, S. and A. Ruggeri (2015): "Indignation, Ideologies, and Armed Mobilization: Civil War in Italy, 1943-45," *International Security,* Vol. 40, No. 2 pp. 119-157.

[8] Curini, Luigi (2018): "Corruption, Ideology, and Populism: The Rise of Valence Political Campaigning". London: Palgrave, 2018.

[9] Farvacque-Vitkovic, C.D., and M. Kopanyi (2014): "Municipal Finances: A Handbook for Local Governments," *World Bank Publications.*

[10] Fearon J. D., (1995): "Rationalist Explanations for War" *International Organization*, Vol. 49, No. 3 (Summer, 1995), pp. 379-414.

[11] Ferraz, C., and F. Finan (2008): "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics*, 123(2): 703-45.

[12] Ferraz, C., F. Finan (2011): "Electoral accountability and corruption: Evidence from the audits of local governments," *American Economic Review,* 101(4), 1274-311.

[13] Gutierrez Sanin, F. and E.J. Wood (2014): "Ideology in Civil War: Instrumental Adoption and Beyond," *Journal of Peace Research,* Vol. 51, No. 2, pp. 213-26.

[14] Hanley, S. and A. Sikk (2016). "Economy, corruption or floating voters? Explaining the breakthroughs of anti-establishment reform parties in eastern Europe," *Party Politics*, 22(4), 522-33.

[15] Henderson, J. V., Kuncoro, A., (2011). "Corruption and local democratization in Indonesia: The role of Islamic parties," *Journal of Development Economics*, vol. 94(2), 164-80.

[16] Iannacone, L.R. and E. Berman (2006): "Religious Extremism: The Good, the Bad, and the Deadly," *Public Choice,* Vol. 128, No. 1, pp. 109-29.

[17] Jackson, M.O., and M. Morelli (2007): "Political Bias and War," *American Economic Review*, Vol. 97-4, 1353-73.

[18] Litschig, S. and Y. Zamboni. (2018). "Audit risk and rent extraction: Evidence from a randomized evaluation in Brazil," *Journal of Development Economics*, Elsevier, vol. 134(C), pages 133-149.

[19] Marshall, M. G., Gurr, T. R., and B. Harff (2009): "PITF state failure problem set: Internal wars and failures of governance, 1955-2008. Dataset and coding guidelines" [Computer file]. Arlington, VA: George Mason University.

[20] Mauro, P. (1995): "Corruption and Growth," *Quarterly Journal of Economics*, Volume 110, Issue 3,681-712.

[21] Muço, A. (2018): "Learn from thy Neighbour: Do Voters Associate Corruption with Political Parties?" Mimeo.

[22] Neudorfer, N. S. and U. G. Theuerkauf (2014): "Buying war not peace: The influence of corruption on the risk of ethnic war." *Comparative Political Studies* 47.13 (2014): 1856-1886.

[23] Power, T.J., and R. Rodrigues-Silveira (2019): "Mapping Ideological Preferences in Brazilian Elections, 1994-2018: A Municipal-Level Study." *Bras. Political Science Review*, 13(1), e0001. Epub February 07, 2019.

[24] Sanz, C., A. Solé-Ollé, and P. Sorribas-Navarro (2020): "Betrayed by the Elites: How Corruption Amplifies the Political Effects of Recessions." Mimeo.

[25] Shleifer, A., and R. W. Vishny (1993): "Corruption." *Quarterly Journal of Economics* 108 (3): 599-617.

[26] Vito, T. (1998): "Corruption Around the World: Causes, Consequences, Scope, and Cures". IMF Working Paper, 1-39, 1998.

[27] Walter, B.F. (2017): "The Extremists' Advantage in Civil Wars," *International Security*, Vol. 42, No. 2 (Fall 2017), 7–39.

Appendix 1

Proof of Proposition 1

We work by backward induction: first we determine the Nash equilibrium of the bargaining game between m and o and then we analyze the choice of a leader made by the opposition group and finally by the majority group. First of all if o is fanatic, then every proposal is rejected. If o is not a fanatic leader, then leader m makes an offer to o. Suppose first that leader m is not a fanatic leader; m can either
(i) make an offer that is accepted only by a corrupt opposition leader, $X_{\min}(v_o) = (1-p)(R+v_o) - c_o - T$; or
(ii) an offer that is accepted also by an honest opposition leader, $X_{\min}(v_o) + T$; or
(iii) an offer that is rejected by every opposition leader o and causes political turmoil for sure.

Leader m prefers to offer $X_{\min}(v_o)$ instead of $X_{\min}(v_o) + T$ (which is accepted also by an honest o) if

$$(R+v_m-(1-p)(R+v_o)+c_o+T)(1-\rho)+\rho(p(R+v_m)-c_m) \geq R+v_m-(1-p)(R+v_o)+c_o$$

which implies

$$v_o \geq v_m + \frac{c_m + c_o}{1-p} - \frac{T(1-\rho)}{\rho(1-p)} \equiv v_1. \tag{7}$$

Leader m prefers to offer $X_{\min}(v_o) + T$ instead of offering strictly less than $X_{\min}$ if

$$(R + v_m - (1-p)(R+v_o) + c_o) \geq p(R+v_m) - c_m$$

which implies

$$v_o \leq v_m + \frac{c_m + c_o}{1-p} \equiv v_2 \tag{8}$$

Leader m prefers to offer $X_{\min}(v_o)$ instead of going to conflict if

$$(R + v_m - (1-p)(R+v_o) + c_o + T) \geq p(R+v_m) - c_m$$

which implies

$$v_o \leq v_m + \frac{c_m + c_o}{1-p} + \frac{T}{(1-p)} \equiv v_3.$$

It follows that leader m offers
(i) $X_{\min}(v_o) + T$ if $v_o \leq v_1$
(ii) $X_{\min}(v_o)$ if $v_1 < v_o \leq v_3$
(iii) $X = 0$ if $v_o > v_3$

If m offers $X_{\min}(v_o)$ to the opposition, a peaceful agreement with a non-fanatic opposition leader is reached with probability $(1-\rho)$, while if m offers less than $X_{\min}(v_o)$ there will be political turmoil. Note that Assumption 1 guarantees that $v_1 < 0$ and therefore a non-fanatic majority leader will never offer more than $X_{\min}(v_o)$.
Suppose now that the majority leader is a fanatic one. A fanatic m aims to maximize the probability to keep the power or, equivalently, to minimize the probability of conflict, so a fanatic leader offers $\min\{R, X_{\min}(v_o) + T\}$.
Consider now the choice of the minority group. Let us first compute the optimal delegation type v_o^* under the assumption that the majority leader m will offer $X_{\min}(v_o)$, and we will then obtain explicit conditions under which this v_o^* is less than the level of ideology inducing conflict, (v_3).

Assuming that m offers $X_{\min}(v)$, the maximization problem for the opposition group is:

$\max_{\{v_o\}} (1-\frac{v_o}{V}\varepsilon)(1-\rho)\left((1-p)(R+v_o) - c_o - T\right) + (\frac{v_o}{V}\varepsilon + (1-\frac{v_o}{V}\varepsilon)\rho)((1-p)R - c_o)$
s.t. $0 \leq v_o \leq V$
F.O.C.:

$$(1-p)(1-\frac{v_o}{V}\varepsilon)(1-\rho) - \frac{1}{V}\varepsilon(1-\rho)\left((1-p)(R+v_o) - c_o - T\right) + \frac{\varepsilon}{V}(1-\rho)((1-p)R - c_0) = 0, \tag{9}$$

$\leftrightarrow$

$$(1-p) - 2(1-p)\frac{v_o}{V}\epsilon + \frac{\varepsilon}{V}[T + (1-p)v_o] = 0,$$

$\leftrightarrow$

$$(1-p) + \frac{\varepsilon}{V}[T + (1-p)v_o] = 2(1-p)\frac{v}{V}\varepsilon$$

$\leftrightarrow$ which yields

$$v_o^* = \min\{V, \frac{V(1-p)/\varepsilon + (1-p)v_o + T}{2(1-p)}\} \tag{10}$$

which is always equal to V for every $\varepsilon < 1/2$. We need now an explicit condition

under which $v_o^* = V$ is less than v_3

$$V \leq \frac{c_m + c_o + T + (1-p)v_m}{(1-p)}$$

which can be rewritten as

$$T \geq \hat{T} \equiv (1-p)(V - v_m) - c_m - c_o \tag{11}$$

Thus: If $T \geq \hat{T}$ then the equilibrium agency ideology is V and there is war with probability ε. Now consider the case $T < \hat{T}$. In such a case the minority group has to choose between the constrained (war minimizing) choice v_3 and any other greater v that would trigger war for sure. The minority group's expected payoff from choosing v_3 is

$$U_o(v_3) = (1 - \frac{v_3}{V}\varepsilon)(1-\rho)((1-p)(R+v_3) - c_o - T) + (\rho(1 - \frac{v_3}{V}\varepsilon) + \frac{v_3}{V}\varepsilon)((1-p)R - c_o)$$

In case the opposition group chooses $v > v_3$ it gets $(1-p)R - c_o$, which is the payoff in case of a conflict and therefore choosing v_3 is preferred if and only if

$$((1-p)(R + v_m + \frac{c_m + c_o}{1-p} + \frac{T}{(1-p)}) - c_o - T) \geq (1-p)R - c_o,$$

or

$$v_m \geq -\frac{c_m + c_o}{1-p},$$

which is obviously satisfied.

We now consider the choice of the majority group. Suppose first that the level of corruption is such that $T > (1-p)V - c_o - c_m$. For all $v_m \geq 0$ the opposition group will choose $v_o = V$ and later m offers $(1-p)(R+V) - c_o - T$. A leader m who is perfectly aligned with her group ($v_m = 0$) by Assumption 1 gets a higher payoff offering $X_{\min}(v_o)$ than $X_{\min}(v_o) + T$; therefore the majority group chooses v_m to minimize the probability of having a fanatic leader and chooses $v_m = 0$.
Consider now the case in which $T \leq (1-p)V - c_o - c_m$. Consider any v_m small enough such that the opposition group will select $v_o = v_3$, the utility of the majority

group if m is not a fanatic is equal to

$$U_m = (1 - \frac{v_3}{V}\varepsilon)(1-\rho)(R-(1-p)(R+v_3)+c_o+T) + (\rho(1-\frac{v_3}{V}\varepsilon)+\frac{v_3}{V}\varepsilon)(pR-c_m)$$

or

$$U_m = (1 - \frac{v_3}{V}\varepsilon)(1-\rho)(pR-(1-p)(v_3)+c_o+T) + (\rho(1-\frac{v_3}{V}\varepsilon)+\frac{v_3}{V}\varepsilon)(pR-c_m)$$

noticing that $v_3 = v_m + \frac{c_o+c_m+T}{1-p}$ and in case of agreement the majority's payoff is equal to $pR-(1-p)(v_m)-c_m$, it follows that majority's utility is maximized when $v_m = 0$ and the expected payoff of the majority group is equal to $pR - c_m$. Suppose the majority group chooses v_m such that $T > (1-p)(V-v_m) - c_o - c_m$. It follows that

$$U_m = (1-\varepsilon)(1-\rho)(R-(1-p)(R+V)+c_o+T) + (\rho(1-\varepsilon)+\varepsilon)(pR-c_m)$$

Notice that majority's payoff in case of agreement is $pR-(1-p)V+c_o+T$ and $V = v_m + \frac{c_o+c_m+T}{1-p}$ and therefore the payoff is equal to $(pR - v_m - c_m) < pR - c_m$ It follows that the majority group chooses $v_m = 0$.

QED.

Continuous cost of corruption

In this section we modify our assumption on the cost of bribing and we assume that it is an increasing and convex function. Suppose for sake of simplicity that it has a quadratic functional form with $\gamma(t) = \frac{1}{2}\phi t^2$ and $\phi > 0$. What bribe the majority leader decides to offer to the opposition leader depends now on marginal cost of bribing ϕ; the smaller is ϕ, the higher will be the bribe offered in equilibrium, ceteris paribus. In fact a proposal (X,t) costs $-X - \frac{1}{2}\phi t^2$ and therefore the optimal bribe (the cost minimizing one) is $t^* = \frac{1}{\phi}$.
Assumption 1a below is a simple adaptation of Assumption 1 to this case.
Assumption 1a: $\rho \leq \frac{1}{1+\phi(c_0+c_m+V(1-p))}$.
This assumption says that the probability that a politician is incorruptible is negligible. Namely, consider a negotiation between a majority and opposition leader such that $v_o = 0$ and $v_m = V$, that is between a majority leader who has the highest non-monetary incentive to avoid any political turmoil and an opposition leader who only

cares about the resources allocated to his group. A corruptible leader of the opposition group accepts any offer larger than $(1-p)R-c_o-\frac{1}{\phi}$, while an honest opposition leader accepts any offer larger than $(1-p)R-c_o$. The majority leader prefers to make the former offer than the latter if $(1-\rho)(R-(1-p)R+c_o+\frac{1}{\phi})+\rho(pR-c_m \geq R-(1-p)R+c_o$, that is if $\rho \leq \frac{1}{1+\phi(c_0+c_m+V(1-p))}$. Therefore Assumption 1a guarantees that for any non-fanatic majority leader (and consequently for the majority group itself) it is always preferable to try to bribe the opposition leader.

Proposition 2 *Both extremism and conflict risk are (weakly) decreasing in the marginal cost of corruption:*

1. *If* $\phi \leq \hat{\phi} \equiv \frac{1}{2((1-p)(V-v_m)-c_m-c_o)}$, *then* $v_o^* = V$ *and prob of war* $= \epsilon$;
2. *If* $\phi > \hat{\phi}$ *then*
   (a) $v_o^* = \frac{c_m+c_o+\frac{1}{\phi}+(1-p)v_m}{(1-p)}$, *decreasing in* ϕ, *and*
   (b) *the probability of war is* $\frac{c_m+c_o+\frac{1}{\phi}+(1-p)v_m}{(1-p)V}\epsilon$, *which is decreasing in* ϕ.

The majority group always selects $v_m = 0$.

Proof:

Given v_o, the utility of the conflict for o is equal to

$$(1-p)(R+v_o)-c_o.$$

For m an offer (X,t) costs $-X-\frac{1}{2}\phi t^2$, therefore the cost minimizing offer is such that $t=\frac{1}{\phi}$. The minimum offer that a corruptible representative o accepts is

$$X_{\min} = (1-p)(R+v_o)-c_o-\frac{1}{\phi}.$$

Leader m prefers to offer $X_{\min}(v_o)$ instead of $X_{\min}(v_o)+\frac{1}{\phi}$ which is accepted also by an honest o if

$$(R+v_m-(1-p)(R+v_o)+c_o+\frac{1}{\phi})(1-\rho)+\rho(p(R+v_m)-c_m) \geq R+v_m-(1-p)(R+v_o)+c_o$$

which implies

$$v_o \geq v_m + \frac{c_m+c_o}{1-p} - \frac{1}{\phi}\frac{(1-\rho)}{\rho(1-p)} \equiv v_1. \tag{12}$$

Leader m prefers to offer $X_{\min}(v_o) + \frac{1}{\phi}$ instead of offering strictly less than $X_{\min}$ if

$$(R + v_m - (1-p)(R + v_o) + c_o) \geq p(R + v_m) - c_m,$$

which implies

$$v_o \leq v_m + \frac{c_m + c_o}{1-p} \equiv v_2. \tag{13}$$

Leader m prefers to offer $X_{\min}(v_o)$ instead of going to conflict if

$$(R + v_m - (1-p)(R + v_o) + c_o + \frac{1}{\phi}) \geq p(R + v_m) - c_m,$$

which implies

$$v_o \leq v_m + \frac{c_m + c_o}{1-p} + \frac{1}{\phi(1-p)} \equiv v_3.$$

It follows that leader m offers
(i) $X_{\min}(v_o) + \frac{1}{\phi}$ if $v_o \leq v_1$
(ii) $X_{\min}(v_o)$ if $v_1 < v_o \leq v_3$
(iii) $X = 0$ if $v_o > v_3$

If m offers $X_{\min}(v_o)$, a peaceful agreement with a non-fanatic opposition leader is reached with probability $(1-\rho)$, while if m offers less than $X_{\min}(v_o)$ there will be political turmoil. Note that Assumption 1a guarantees that $v_1 < 0$ and therefore a non-fanatic majority leader will never offer more than $X_{\min}(v_o)$.
Suppose now that the majority leader is a fanatic one. A fanatic m aims to maximize the probability to keep the power or, equivalently, to minimize the probability of conflict, so a fanatic leader offers $\min R, X_{\min}(v_o) + \frac{1}{\phi}$.

Consider now the agency choice of the minority group.

Let us first compute the optimal delegation type v_o^* under the assumption that the majority leader m will offer $X_{\min}(v_o)$, and we will then obtain explicit conditions under which this v_o^* is less than the level of ideology inducing conflict, (v_3).

Assuming that m offers $X_{\min}(v_o)$, the maximization problem for the opposition group is:

$\max_{\{v_o\}} (1 - \frac{v_o}{V}\varepsilon)(1-\rho)\left((1-p)(R+v_o) - c_o - \frac{1}{\phi}\right) + (\frac{v_o}{V}\varepsilon + (1 - \frac{v_o}{V}\varepsilon)\rho)((1-p)R - c_o)$
s.t. $0 \leq v_o \leq V$

F.O.C.:

$$(1-p)(1-\frac{v_o}{V}\varepsilon)(1-\rho)-\frac{1}{V}\varepsilon(1-\rho)\left((1-p)(R+v_o)-c_o-\frac{1}{\phi}\right)+\frac{\varepsilon}{V}(1-\rho)((1-p)R-c_0)=0, \tag{14}$$

$\leftrightarrow$

$$(1-p)-2(1-p)\frac{v_o}{V}\epsilon+\frac{\varepsilon}{V}[\frac{1}{\phi}+(1-p)v_o]=0,$$

$\leftrightarrow$

$$(1-p)+\frac{\varepsilon}{V}[\frac{1}{\phi}+(1-p)v_o]=2(1-p)\frac{v}{V}\varepsilon$$

which yields

$$v_o^*=\min\{V,\frac{V(1-p)/\varepsilon+(1-p)v_o+\frac{1}{\phi}}{2(1-p)}\} \tag{15}$$

which is always equal to V for every $\varepsilon<1/2$. We need now an explicit condition under which $v_o^*=V$ is less than v_3

$$V\leq\frac{c_m+c_o+\frac{1}{\phi}+(1-p)v_m}{(1-p)}$$

which can be rewritten as

$$\frac{1}{\phi}\geq\hat{T}\equiv(1-p)(V-v_m)-c_m-c_o \tag{16}$$

Thus: If $\frac{1}{\phi}\geq\hat{T}$ then the equilibrium agency ideology is V and there is war with probability ε, proving part 1 of the proposition.

Now consider the case $\frac{1}{\phi}<\hat{T}$. In such a case the minority group has to choose between the constrained (war minimizing) choice v_3 and any other greater v that would trigger war for sure. The minority group's expected payoff from choosing v_3 is

$$U_o(v_3)=(1-\frac{v_3}{V}\varepsilon)(1-\rho)((1-p)(R+v_3)-c_o-\frac{1}{\phi})+(\rho(1-\frac{v_3}{V}\varepsilon)+\frac{v_3}{V}\varepsilon)((1-p)R-c_o)$$

In case the opposition group chooses $v>v_3$ it gets $(1-p)R-c_o$, which is the

payoff in case of a conflict and therefore choosing v_3 is preferred if and only if

$$((1-p)(R+v_m+\frac{c_m+c_o}{1-p}+\frac{1}{\phi(1-p)})-c_o-\frac{1}{\phi})\geq(1-p)R-c_o,$$

or

$$v_m\geq-\frac{c_m+c_o}{1-p},$$

which is obviously satisfied, proving part 2 of Proposition 2.

We now consider the choice of the majority group. Suppose first that the level of corruption is such that $\frac{1}{\phi}>(1-p)V-c_o-c_m$. For all $v_m\geq 0$ the opposition group will choose $v_o=V$ and later m offers $(1-p)(R+V)-c_o-\frac{1}{\phi}$. A leader m who is perfectly aligned with her group ($v_m=0$) by Assumption 1a gets a higher payoff offering $X_{\min}(v_o)$ than $X_{\min}(v_o)+\frac{1}{\phi}$; therefore the majority group chooses v_m to minimize the probability of having a fanatic leader and chooses $v_m=0$.
Consider now the case in which $\frac{1}{\phi}\leq(1-p)V-c_o-c_m$. Consider any v_m small enough such that the opposition group will select $v_o=v_3$, the utility of the majority group if m is not a fanatic is equal to

$$U_m=(1-\frac{v_3}{V}\varepsilon)(1-\rho)(R-(1-p)(R+v_3)+c_o+\frac{1}{\phi})+(\rho(1-\frac{v_3}{V}\varepsilon)+\frac{v_3}{V}\varepsilon)(pR-c_m)$$

or

$$U_m=(1-\frac{v_3}{V}\varepsilon)(1-\rho)(pR-(1-p)(v_3)+c_o+\frac{1}{\phi})+(\rho(1-\frac{v_3}{V}\varepsilon)+\frac{v_3}{V}\varepsilon)(pR-c_m)$$

noticing that $v_3=v_m+\frac{c_o+c_m+\frac{1}{\phi}}{1-p}$ and in case of agreement the majority's payoff is equal to $pR-(1-p)(v_m)-c_m$, it follows that majority's utility is maximized when $v_m=0$ and the expected payoff of the majority group is equal to $pR-c_m$. Suppose the majority group chooses v_m such that $\frac{1}{\phi}>(1-p)(V-v_m)-c_o-c_m$. It follows that

$$U_m=(1-\varepsilon)(1-\rho)(R-(1-p)(R+V)+c_o+\frac{1}{\phi})+(\rho(1-\varepsilon)+\varepsilon)(pR-c_m)$$

Notice that majority's payoff in case of agreement is $pR-(1-p)V+c_o+\frac{1}{\phi}$ and $V=v_m+\frac{c_o+c_m+\frac{1}{\phi}}{1-p}$ and therefore the payoff is equal to $(pR-v_m-c_m)<pR-c_m$

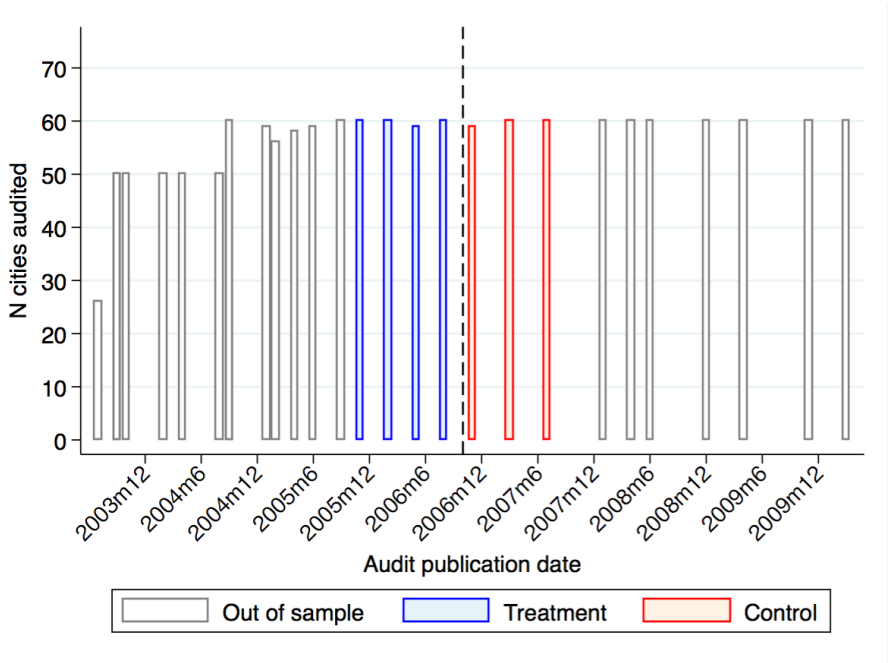
It follows that the majority group chooses $v_m = 0$.

QED.

Appendix 2

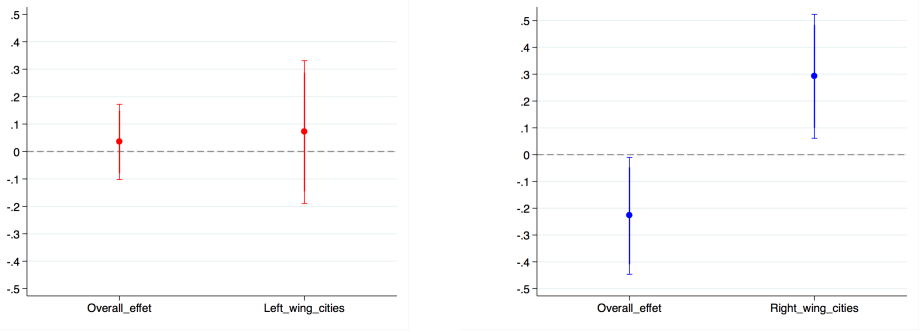
Figures

Figure 1: Audits timing – Treatment and control groups



The plot shows the different audit waves over time in relation to the gubernatorial elections of 2006 (October 1st 2006). The audits included in the treatment group are those whose results have been published between October 2005 and September 2006, while those in the control group have been reported between October 2006 and September 2007. The other audits are excluded from the sample.

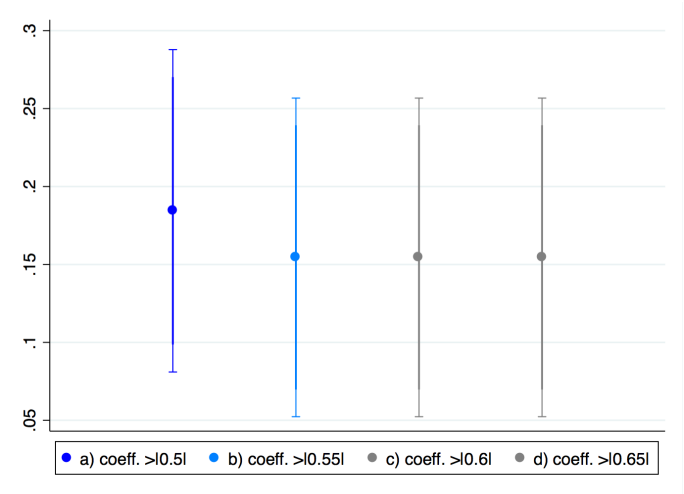
Figure 2: Impact of corruption on extreme voting in non-aligned cities – Votes to extreme left/right parties



(a) Diff. vote sh. (extreme left parties) (b) Diff. vote sh. (extreme right parties)

The plots show the triple interaction term *not-aligned*pre-election*corruption* (indicated as *Overall_effect*) and the interaction term *not-aligned*pre-election*corruption*left/right mayor* (indicated as *Left_wing_cities* -panel a- and as *Right_wing_cities* -panel b-), according to model 5. We use the indicator for corruption above the median level. The dependent variable used is *Difference on vote share (extreme left parties)*.

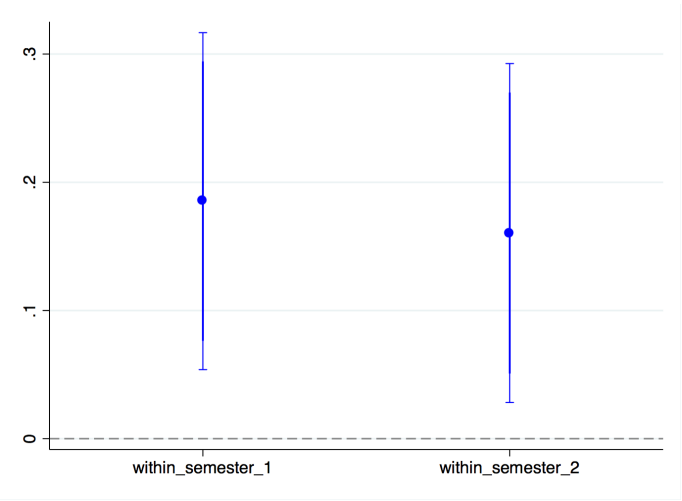
Figure 3: Impact of corruption on extreme voting in non-aligned cities – Different definition of extremism



(a) Difference on vote share (extreme)

The plot shows the triple interaction term *not-aligned*pre-election*corruption*, using the indicator for corruption above the median level, for the main analysis, increasing progressively the ideology threshold to define a party as extreme. In particular, the four points represent distinct analysis conducted using, respectively, as ideology threshold the value of 0.5 (result a), the value 0.55 (result b), the value 0.6 (result c) and the value 0.65 (result d). The dependent variable is *Difference on vote share (extreme)*. The specification is the one of model 4.

Figure 4: Impact of corruption on extreme voting in non-aligned cities – Including semester-specific variables



(a) Difference on vote share (extreme)

The plot shows the triple interaction terms *not-aligned*pre-election(1st semester)*corruption* and *not-aligned*pre-election(2nd semester)*corruption*, according to model 6. The indicator of corruption used divides cities according to the median level and the dependent variable is *Difference on vote share (extreme)*.

Tables

Table 3: Descriptive statistics

	mean	sd
Panel A: Corruption measures		
Narrow corruption	.013	.0455
Broad corruption	.045	.103
Panel B: Political variables		
Extreme vote share - Difference	-.276	.357
Extreme vote share - Levels	.202	.258
Not-aligned - party	.614	.488
Mun. left-wing	.25	.434
Mun. right-wing	.745	.436
Turnout	79.78	13.89
Number voters	19,027.57	39,889.37
Voters age<18 (fraction)	4.51	1.33
Voters 18<age<59 (fraction)	81.83	2.79
Voters age>59 (fraction)	13.66	2.73
Panel C: Other variables		
Population (2006)	27138.34	61948.63
Av. monthly income	579.46	308.66
Share of employed	38.25	8.16
Av. number of years of education	3.59	1.09
N	396	

Notes: The variables *Narrow corruption* and *Broad corruption* capture the fraction of misappropriated funds on the total amount audited, including different types of violations. The variables *Extreme vote share - Difference* and *Extreme vote share - Levels* measures the share of votes to the extreme parties in 2006 gubernatorial elections, respectively, in difference with respect to the 2002 gubernatorial polls or in levels. The parties classified as extreme are those with a value of ideology score higher than 0.5, according to Power, Rodrigues-Silveira (2018). *Not-aligned - party* is a dummy variable indicating whether the municipality is not aligned with the gubernatorial administration, based on the coalition of parties of the ruling governor. *Av. monthly income* is expressed in Brazilian Reais.

Table 4: Balance test

	Mean(Pre-election=1) (1)	Mean(Pre-election=0) (2)	Diff. (3)	Std. Error (4)	Obs. (5)
Narrow corruption	0.009	0.0190	0.01**	0.005	396
Broad corruption	0.019	0.079	0.060***	0.010	396
Population (2006)	23141.73	32562.29	9420.57	6288.89	396
Av. monthly income	593.99	559.99	-33.99	31.55	391
Av. education	3.63	3.53	-0.102	0.111	391
Turnout	79.85	79.69	-0.16	1.47	367
Number voters	16699.57	22283.75	5584.18	4218.83	367
Voters age<18 (fraction)	4.46	4.57	0.108	0.141	367
Voters 18<age<59 (fraction)	81.77	81.92	0.149	0.295	367
Voters 18<age>59 (fraction)	13.76	13.51	-0.257	0.289	367
Mun. left-wing	0.241	0.2619	0.0207	0.0441	396
Mun. right-wing	0.754	0.732	-0.022	0.044	396
Not-aligned - party	0.636	0.583	-0.053	0.049	396

Notes: This table shows the characteristics of the 396 municipalities under analysis - those whose audits have been reported in the period October 2005-September 2007. The treatment group (column 1) is composed by 228 municipalities where the disclosure occurred between October 2005 and September 2006. Instead, the control group (column 2) is composed by 168 municipalities where the disclosure happened between the October 2006 and September 2007. Column (3) shows the difference of the means and the level of significance, column (4) displays the corresponding standard error and column (5) show the sample size.

Table 5: Impact of corruption on extreme voting – broad corruption, difference in vote share

Dep. var.: Vote share to extreme - difference	Linear		Quartile		Median	
	(1)	(2)	(3)	(4)	(5)	(6)
Not-aligned*Pre-election		0.0215		0.0203		0.0416
		(0.0459)		(0.0421)		(0.0557)
Pre-election*corruption	-0.376**	-1.563***				
	(0.142)	(0.282)				
Not-aligned*corruption		-0.0821				
		(0.236)				
Not-aligned*Pre-election*corruption		1.458***				
		(0.341)				
Pre-election*corruption (above fourth Q)			-0.0554	-0.134*		
			(0.0404)	(0.0649)		
Not-aligned*corruption (above fourth Q)				-0.0186		
				(0.0516)		
Not-aligned*Pre-election*corruption (above fourth Q)				0.131		
				(0.0789)		
Pre-election*corruption (above median)					0.00424	-0.0325
					(0.0538)	(0.0863)
Not-aligned*corruption (above median)						0.0356
						(0.0610)
Not-aligned*Pre-election*corruption (above median)						0.0657
						(0.0641)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367	367	367	367	367	367

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections in difference with respect to the 2002 elections. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit – version *Broad corruption* –, expressed as the percentage of the budget involved in general violations, *corruption (above median)* is a dummy equal to one whether the share of irregularities is above the median level and *corruption (above fourth Q)* is a dummy equal to one whether the share of irregularities is above the fourth quartile. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

Table 6: Impact of corruption on extreme voting – broad corruption, vote share

Dep. var.: Vote share to extreme - difference	Linear		Quartile		Median	
	(1)	(2)	(3)	(4)	(5)	(6)
Not-aligned*Pre-election		-0.000353		0.00509		-0.00265
		(0.0192)		(0.0183)		(0.0238)
Pre-election*corruption	0.0844	-0.378				
	(0.159)	(0.268)				
Not-aligned*corruption		-0.249*				
		(0.140)				
Not-aligned*Pre-election*corruption		0.728**				
		(0.307)				
Pre-election*corruption (above fourth Q)			0.00365	-0.0323		
			(0.0292)	(0.0370)		
Not-aligned*corruption (above fourth Q)				-0.0143		
				(0.0117)		
Not-aligned*Pre-election*corruption (above fourth Q)				0.0579		
				(0.0349)		
Pre-election*corruption (above median)					0.0393	0.0107
					(0.0363)	(0.0489)
Not-aligned*corruption (above median)						-0.0206
						(0.0167)
Not-aligned*Pre-election*corruption (above median)						0.0448
						(0.0368)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367	367	367	367	367	367

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit – version *Broad corruption* –, expressed as the percentage of the budget involved in general violations, *corruption (above median)* is a dummy equal to one whether the share of irregularities is above the median level and *corruption (above fourth Q)* is a dummy equal to one whether the share of irregularities is above the fourth quartile. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

Table 7: Impact of corruption on extreme voting – Excluding single lotteries

Dep. var.: Vote share to extreme - difference	Linear						
Excluding lottery released on	November 2005	February 2006	May 2006	August 2006	December 2006	March 2007	July 2007
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Not-aligned*Pre-election	0.0218	0.0183	0.0186	0.0363	0.0104	0.00851	0.0346
	(0.0421)	(0.0535)	(0.0382)	(0.0448)	(0.0566)	(0.0411)	(0.0505)
Pre-election*corruption	-2.746***	-1.959***	-1.444	-1.972***	-2.344***	-1.960***	-1.917
	(0.856)	(0.602)	(0.860)	(0.570)	(0.518)	(0.585)	(1.360)
Not-aligned*corruption	-0.0895	-0.0408	0.0334	-0.0298	-1.414	0.665	0.0501
	(0.668)	(0.651)	(0.641)	(0.651)	(1.384)	(0.542)	(1.372)
Not-aligned*Pre-election*corruption	2.630**	2.095**	1.598	1.947**	3.470**	1.531**	1.954
	(1.082)	(0.768)	(0.987)	(0.806)	(1.289)	(0.723)	(1.411)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	313	312	314	315	313	315	320

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections in difference with respect to the 2002 elections. Every columns excludes the observation from a single audit. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit –version *Narrow corruption*–, expressed as the percentage of the budget involved in general violations. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

Table 8: Impact of corruption on extreme voting – Placebo: features correlated with corruption

Dep. var.: Vote share to extreme - difference	Linear			
	(1)	(2)	(3)	(4)
Non-aligned*corruption	-0.0563	-0.0409	-0.0706	-0.238
	(0.648)	(0.634)	(0.631)	(0.617)
Pre-election*corruption	-2.095***	-1.606***	-1.789***	-2.270***
	(0.619)	(0.487)	(0.558)	(0.657)
Non-aligned*Pre-election*corruption	2.150***	1.557**	1.775**	2.405***
	(0.764)	(0.713)	(0.639)	(0.713)
Non-aligned*Population (2006)	-0.000000160			
	(0.000000267)			
Pre-election*Population (2006)	-0.000000908***			
	(0.000000303)			
Non-aligned*Pre-election*Population (2006)	0.000000873*			
	(0.000000503)			
Non-aligned*Income		-0.0000250		
		(0.0000972)		
Pre-election*Income		0.000117		
		(0.000104)		
Non-aligned*Pre-election*Income		-0.000169		
		(0.000141)		
Non-aligned*Education			0.00394	
			(0.0258)	
Pre-election*Education			0.0259	
			(0.0299)	
Non-aligned*Pre-election*Education			-0.0353	
			(0.0416)	
Non-aligned*Public administration				0.0150
				(0.0292)
Pre-election*Public administration				0.00650
				(0.0162)
Non-aligned*Pre-election*Public administration				-0.0252
				(0.0370)
State fixed effects	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes
N	367	367	367	367

Notes: The dependent variable is the percentage of votes to extreme parties in 2006 gubernatorial elections in difference with respect to the 2002 elections. The level of extremism of parties is defined according to Power, Rdrigues-Silveira (2019), using a bandwidth of 0.5. *not-aligned* is a dummy variable capturing if the municipal government is politically aligned with the gubernatorial administration, equal to one if the party of the mayor belongs to the coalition of the governor in charge in that state. *pre-election* is a dummy whether the city has been audited before 2006 gubernatorial election (October 2006). *corruption* is a measure of the irregularities found in the audit –version *Narrow corruption*–. *Fixed effects* include state fixed effects. *Controls* includes characteristics of the political background (number of voters and turnout of the last municipal election), of the social background (average income level, share of employed, average education and population of the city in 2006) and the dummies measuring the distance (in months) of the elections from the audit release, differential across states, for cities in the treatment group. Robust standard errors are clustered at the state level.

Table 9: Descriptive statistics

	Average value	SD
Panel A: Corruption measures		
Executive corruption	1.988	1.477
Legislative corruption	1.999	.9926
Panel B: Conflict measures		
Ethnic conflict	.0987	.2983
Number of fighters	.2817	.9142
Number of fatalities	.1593	.5716
N	8,823	

Notes: The indices *Executive/Legislative* vary in the range 0-4. *Ethnic war* is dummy indicating the insurgence of an ethnic conflict. *Number of fighters/fatalities* vary in the range 0-4. These numbers correspond to following figures: 0 less than 100, 1 between 100 and 1,000, 2 between 1,000 and 5,000, 3 between 5,000 and 10,000, 4 more than 10,000.

Table 10: Impact of corruption on ethnic conflict events

	Ethnic conflict		Number of **rebel fighters**		Number of **fatalities**	
Type of corruption:	Executive corruption	Legislative corruption	Executive corruption	Legislative corruption	Executive corruption	Legislative corruption
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: without controls						
Corruption	0.0215**	0.0323*	0.0722**	0.0993**	0.0399**	0.0573*
	(0.0102)	(0.0164)	(0.0309)	(0.0461)	(0.0170)	(0.0343)
N	8519	7636	8509	7627	8507	7627
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No
Panel B: with controls						
Corruption	0.0187*	0.0242*	0.0618*	0.0808*	0.0312*	0.0436
	(0.0110)	(0.0143)	(0.0355)	(0.0449)	(0.0183)	(0.0330)
N	7996	7238	7986	7229	7984	7229
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dep. var. mean	0.098	0.098	0.28	0.28	0.15	0.15

Notes: Every line and column represents a distinct analysis. The dependent variable is a dummy equal to one when an ethnic conflict emerges in a specific country/year (columns 1-2), the number of fighting rebels in case an ethnic conflict emerges in a specific country/year (columns 3-4) and the number of fatalities in case an ethnic conflict emerges in a specific country/year (columns 5-6). The indices of corruption are discrete measures in the range 0-4. The regression includes country and continent-year fixed effects, moreover the following controls are included in Panel B: population, GDP, area, elevation (average value and standard deviation), total past conflict onsets and total past years of peace. Robust standard errors are clustered at the country level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.