Identifying the Effect of Truth Commissions on the Quality of Democracy

Milena Ang, Genevieve Bates, and Monika Nalepa
The University of Chicago

This version February 6, 2021

Abstract

Does transitional justice hinder or help democracy? This simple question poses a challenge because countries choosing to embark on transitional justice may be exactly the ones that would have had a successful pathway to democratization in the first place. We resolve it by measuring transitional justice as a series of events rather than one-shot instances of dealing with the past. This method of coding not only allows us to measure their impact on democratic quality, but is also more faithful to the operation of truth commissions on the ground. Among transitional justice mechanisms, we focus on truth commissions, one of the most popular institutions for dealing with authoritarian pasts in the last fifty years. We examine one possible mechanism through which truth commissions can enhance democratic quality, arguing that the transparency of truth commissions can prevent blackmailers from to extracting policy concessions from politicians compromised with "skeletons in their closet." According to this logic, transparency regimes result in less corrupt politics and better quality of democratic representation. To corroborate this claim, we use a difference-in-difference design to identify the causal link between truth commissions and political corruption on a panel dataset of 81 countries that transitioned to democracy since 1946.

1 Introduction

Countries that transitioned from authoritarian rule or that are recovering from civil conflict are confronted with the dilemma of whether and how to reckon with their violent past. Typically the first and most popular calls are for criminal prosecutions against those who bent or betrayed the constitution or who committed egregious human rights violations on the rulers' behalf. Such calls have been met with scepticism from the classical literature, which argues that "the prosecution of perpetrators of atrocities according to universal standards risks causing more atrocities than it would prevent, because it pays insufficient attention to political realities" (Snyder & Vinjamuri 2004, 5) and that sometimes "amnesty [...] is necessary to establish a new democracy on a solid basis" (Huntington 1993, 214). Criminal prosecutions against agents of repression are fraught also because of the legal challenge from non-retroactivity—Nullum crimen sine lege—the principle that one cannot prosecute perpetrators of acts if they were legal at the time they were committed.

In response to these and other challenges, including underdeveloped judiciaries, new democracies have sought creative non-punitive forms of reckoning with their pasts. The totality of measures for dealing with authoritarian elites, their collaborators, or perpetrators of human rights violations as well as their victims is studied under the heading of transitional justice (TJ) (Bakiner 2016, Loyle & Appel 2017, Zvobgo 2019*a*, Balcells, Palanza & Voytas 2020, Capoccia & Pop-Eleches 2020).

Generally, mechanisms of transitional justice fall into one of five categories. Next to trials of former perpetrators, there are lustrations, reparations, purges, and truth commissions. Lustrations deal with secret informers of the former authoritarian regime and can take the form of opening archives of the former secret police in order to uncover who worked as a secret collaborator or informer.[1] Reparations range from monetary compensation and the return of expropriated wealth in kind to purely symbolic public apologies.

[1]Lustration is also called TJ vetting. See Bates, Cinar & Nalepa (2020) for a detailed explanation of TJ vetting.

Purges, often confused with lustrations, denote the firing of members of an ancien régime institution whose collaboration with it is known, such as the dismissal of 13 Supreme Court justices in post-Noriega Panama (Human Rights Watch 1991). Truth commissions, the focus of this article, are "bodies set up to investigate a past history of human rights abuses in a particular country, which can include violations by the military or other government forces or armed opposition forces" (Hayner 1994).

The most famous example of a truth commission is the South African Truth and Reconciliation Commission (TRC) (Hayner 2010, Gibson 2006), formed in 1995 to investigate crimes against the South African people during the apartheid regime (1960-1994).[2] The TRC could investigate human rights violations committed by the state and various liberation movements, and it had the prerogative to offer amnesty to those who fully participated in the process and truthfully recounted their crimes, thus incentivizing perpetrators to come forward to share their story. In 1998, after touring the country for a couple of years, conducting hearings that were nationally broadcast on public radio, and collecting testimonies, the Commission released a five-volume final report. The report detailed the abuses committed by the Apartheid-era National Party government, the African National Congress (ANC)—the opposition-turned-ruling party—and other "leading political figures on both sides of the anti-apartheid struggle." (Keesing's Record of World Events 1998, 42536)

The TRC exemplifies an important feature of truth revelation. Although the Commission held hearings for a number of known offenders, many of their crimes implicated unknown perpetrators and in many instances the abuses that were revealed had been previously undisclosed. By making the history of abuses transparent, the TRC prevented agents with insider knowledge of who committed what crimes and where from from blackmailing the perpetrators by threatening to reveal to the public what they had done. Absent the transparency produced by the Commission, threats to denounce perpetrators of egre-

[2]The TRC was established via the Promotion of National Union and Reconciliation Act, passed by the South African parliament in July 1995 (Gibson & Gouws 1999).

gious and violent acts unless they gave in to their economic or policy demands may have derailed political representation in South Africa. In this sense, truth commissions remove *kompromat* or "skeletons in the closet" from the domain of political representation, decreasing the corruptibility of the new polity.

Truth commissions, of course, perform a host of other important functions, which we discuss at length in the next section. Yet we believe that this democracy-enhancing aspect underscores an often overlooked function of truth commissions. As noted before, existing literature has tended to focus on the punitive aspect of transitional justice; only recently have scholars begun to theorize and empirically illustrate the importance of disclosure and transparency for democracy. Horne & Levi (2004), for example, argue that by naming those who participated in human rights violations, lustrations can restore a sense of justice and increase victims' trust in democratic institutions. Ang & Nalepa (2019) show that because former dissident-turned-politicians can be blackmailed by former authoritarian elites, distorting political representation and worsening the quality of democracy, lustration can render this threat null by disclosing embarrassing information. Our work suggests that these democracy enhancing properties should be even more pronounced in the case of truth commissions. In order to be subject to lustration, one has to actually run for office or occupy a public position. Truth commissions, meanwhile, cast a wider net both in terms of extending to a larger set of potential collaborators, and in terms of covering a greater range of crimes committed on behalf of the regime. As such, truth commissions should be exceptionally poised to deliver democracy-enhancing results.

This paper contributes to the evolving literature on the effects of transparency regimes by focusing on truth commissions. We begin by arguing that truth commissions improve the quality of democracy in two ways. First, by exposing "skeletons in the closet," truth commissions effectively eliminate the leverage that authoritarian networks have over elites in the democratic period. Second, truth commissions can dismantle authoritarian networks by weakening the clandestine ties that underpin them.

Our contribution is also empirical. We use a panel dataset to leverage variation across countries and within countries across time to identify the causal effect of truth commissions on quality of democracy. This dataset enables us to record the advancement, or delay, of truth commissions over time. In doing so, our data acknowledges that truth commissions are the result of a politically contested process, and allows a more accurate metric of truth commissions. The dataset also offers us to estimate a modified difference-in-differences model, which helps us provide better evidence of the causal effect of truth commissions on the quality of democracy.

The paper begins by discussing our theoretical argument in detail, and from this discussion we develop a series of detailed predictions and theoretical expectations about the effects of truth commissions on democratic consolidation. In Section 3, we review existing research on identifying the causal relationship between transitional justice and dependent variables related to peace and democratic stability. We then present our own research design and explain how we operationalize the dependent and independent variables. Section 4 is devoted to data analysis and interpreting the results. We find that truth commissions are across all specifications robust in improving the quality of democracy. Section 5 concludes.

2 The transparency-enhancing function of truth commissions

We build on an understanding of truth commissions outlined in Hayner (1994), seeing them as state-sanctioned "bodies set up to investigate a past history of human rights abuses in a particular country, which can include violations by the military or other government forces or armed opposition forces" (14).[3] Onur Bakiner offers more nuance to this defi-

[3]Truth commissions should (1) not focus on ongoing human rights abuses as a human rights ombudsman might; (2) examine a pattern of human rights abuses over time rather than a specific event; (3) be temporary; and (4) have an official sanction from the state to carry out its operations (Hayner 1994, 14).

nition by distinguishing truth commissions from "similar investigatory, judicial, or commemorative practices and institutions, such as parliamentary human rights commissions, courts, monitoring institutions and NGO's truth finding efforts" (Bakiner 2016, 11). Departing from Bakiner's distinctions, we do not exclude commissions of inquiry that examine human rights violations committed in more specific events than an entire period of authoritarian rule or civil war. We consider this inclusion justified as we also include commissions that only partially completed their mandate.

This all-encompassing view of truth commissions brings our work close to that of Zvobgo (2019*b*) whose Varieties of Truth Commissions database sheds light on just how much diversity exists in the functions performed by truth commissions. With varying mandates, truth commissions can name individual perpetrators,[4] as the Commission on the Truth for El Salvador did, provide amnesty for participation, as the South African Truth and Reconciliation Commission did, and even provide direct reparations to victims, as the Truth, Justice, and Reconciliation Commission of Kenya's mandate suggests. Many, if not all, truth commissions make specific policy recommendations geared towards ensuring that such abuses do not occur in the future. Some even go on to recommend prosecution against certain perpetrators, as occurred with the Truth and Reconciliation Commission of Sierra Leone.[5]

Among the things that all truth commissions and commissions of inquiry have in common is that there is always more transparency after the commencement of a truth commission process than before it. Moreover, in the years when a given truth commission is particularly active, the boosts of transparency should be especially prominent. Finally, the longer a truth commission is in operation, the higher the likelihood that information about

[4]While the naming of individual perpetrators of human rights abuses is relatively uncommon in truth commissions, in outlining patterns of abuses, these commissions nonetheless often identify particular state institutions or groups of actors who played a role in either perpetrating abuses directly or supporting those who did so.

[5]It is important to note that the extent of perpetrator and/or victim participation in truth commission activities can play an important role in what these commissions achieve (Zvobgo 2019*a*, Zvobgo 2019*b*).

individual perpetrators may eventually come to light.[6] Since transparency pertaining to individual perpetrators is the mechanism that our theory focuses on, we build on these three aspects of truth commissions when we develop our measures in section four.

Authoritarian regimes sustain themselves thanks to agents who repress citizens of the authoritarian state either publicly or secretly. Following the democratic transition, both those who openly violated human rights and those who did so in secret may be held accountable for their abuses and secret collaborators of the regime might be revealed. We argue that truth commissions, to the extent that they disclose atrocities committed in secret on behalf of the authoritarian regime, improve the quality of the newly inaugurated democratic regime.

This positive impact is achieved through two related mechanisms. First, unearthing secret collaborations prevents the blackmail of former secret agents of repression who should find themselves in positions of power. Note that *unknown* or *secret* collaborators of former regimes might want to take part in the new regime. Yet, absent transparency mechanisms, they are susceptible to blackmail with the threat of revealing their 'skeletons in the closet', i.e. revealing that they participated in atrocious acts. These threats, especially when they come from persons with credible access to such evidence, can distort policy and representational outcomes. The threat of revealing skeletons holds sway over former elites, because the public likely cares about not being represented by past abusers. Thus, revealing such 'skeletons' could chastise a public official, and end his or her career. In return for not denouncing them, blackmailers can demand rents or policy concessions, either of which corrupts the new democracy.

Nalepa & Sonin (2020) who investigate this blackmail mechanism with a formal model show that lack of transparency can distort representation even when only "innocent" candidates are running for office. The authors argue that the mere presence of secret au-

[6]Even if a commission does not identify perpetrators by name, we think it is reasonable to believe that the passage of time and ongoing truth commission activities can encourage those who are more likely to individuals by name, rank, and/or affiliation.

thoritarian legacies enables blackmailers to pressure candidates into running on a non-representative platform (that is, a platform that diverges from the general voting population). The mere possibility of a collaborator of the previous regime running for office can tarnish the reputation of other candidates, leading to the reelection of extreme incumbents. Transitional justice procedures that reveal past atrocities, such as truth commissions, benefit voters regardless of the voters' preferences for normative transitional justice. In other words, the beneficial effects of transitional justice, in particular the effects of revealing skeletons in the closet, take hold regardless of–and sometimes despite–the public's demand for retribution. This is also a good opportunity to note that the public's sentiments towards former secret collaborators—those who would be exposed by a truth commission—need not be correlated to with their reluctance to reelect successor autocrats into office. The stress here is on the honesty of public persons; not what they did but whether or not they continue to lie about their past. As the experience of Post-Communist Europe indicates, the reelection of successor communist parties into office is perfectly compatible with the stigmatization of former secret collaborators (Horne 2017).

We build on these already identified dynamics to hypothesize a second mechanism connecting uncovering secret collaborators and democratic quality: transparency regimes can upend existing authoritarian networks that linger on even after the transition to democracy. Former secret collaborators, if undisclosed, can be manipulated by former authoritarian elites. Consider again the South African example, where even prior to the transition, the ANC would hold so-called "People's Courts" against suspected Apartheid collaborators. Those found guilty would be subjected to the excruciating torture of having a burning tire placed over their head (Price 1991). This example illustrates the stigma associated with the suspicion of being a collaborator. Even following the transition collaborators were not subjected to the horrific "doughnut death" as the burning tire sentence was called (Price 1991), they would arguable go to far lengths to avoid being revealed. But the TRCs act publishing names of those working for the Apartheid also disarmed the

power of authoritarian networks that may well have survived the transition to democracy otherwise.

Without a transparency mechanism in place, former authoritarian elites can threaten to expose secrets as a way of extracting policy concessions. Voters, meanwhile, are kept in the dark about whose interests politicians are really representing. By disclosing secret authoritarian legacies, truth commissions fulfill democratic goals by informing voters about candidates and allowing them to choose "honest" politicians. Politicians that have been named and shamed as collaborators by truth commissions can continue to try to represent voters, but such politicians will be more transparent to voters as their actions in office are less likely to be driven by the interests of blackmailers threatening to reveal skeletons in their closet. Truth commissions also undermine authoritarian networks and render them obsolete.[7]

In sum, truth commissions remove opportunities for blackmail–which may threaten the quality of democratic representation–and provide voters with information necessary for removing 'dishonest' officials–who were plausibly connected to former authoritarian elites. Importantly, these effects are brought about by the transparency aspect of the transitional justice mechanism, an effect that we argue is theoretically and empirically distinct from the criminal punishment of former members and collaborators of the authoritarian regime who abused human rights publicly. Therefore, truth commissions will improve the quality of democracy in two directly observable ways. First, by revealing abuses committed by secret agents of repression truth commissions will undercut the blackmail mechanism described above and reduce political corruption. Second, by revealing this compromising information to voters about the true characters of persons who may potentially continue occupying public office, truth commissions will reduce the power of authoritarian networks.

[7]We note above that truth commissions can also enable the punishment of former perpetrators of abuses by, for example, creating a file that can be used in a criminal justice trial. But since our argument focuses on the effect of uncovering the truth, the sanction itself is neither necessary nor sufficient for the mechanism that we have highlighted here.

3 What we know about the effects of transitional justice

Our claim that truth commissions can reduce the corruptability of a new democracy requires empirical evidence beyond a simple correlational analysis. A positive relationship between truth commissions and measures of democratic quality alone is insufficient evidence for our theory because the same factors that cause states to embark on transitional justice may lead them to have higher levels of democratic indicators later on.

Previous scholarship on the effects of justice institutions in transitional and post-conflict settings has used several techniques to address these endogeneity concerns. In their work on the relationship between post-conflict justice processes and civil war recurrence, Loyle & Appel (2017) take an instrumental variable approach, using the presence of transitional justice institutions in the region as an instrument. Assuming that neighbors with and without such institutions are assigned to any given post-conflict country at random and assuming that transitional justice institutions diffuse, these authors find that truth commissions, reparations, amnesties, and comprehensive trials—which they call "motivation post-conflict justice"— decrease the likelihood of conflict recurrence. As an additional robustness check to identify the determinants of post-conflict justice implementation, these authors also employ a strategy of matching on observables.[8] The results are consistent with their instrumental variable analysis.

Prorok (2017) also uses an instrumental variable approach, identifying three plausible instruments to estimate the causal effects of International Criminal Court (ICC) investigations on conflict duration in civil wars. Using as instruments (1) a state's affinity with the permanent five members of the UN Security Council, (2) a state's affinity with its neighbors, and (3) the number of neighboring states that have ratified the Rome Statute, Prorok finds that ICC involvement in a conflict setting significantly reduces the probability of conflict termination when government and rebel groups have committed similar levels

[8]This robustness check is an important improvement because an additional potential problem with an identification strategy that uses neighboring units as instruments is it's failure to satisfy the exclusion restriction, as demonstrated by Betz, Cook & Hollenbach (2019).

of atrocities. The Court's counterproductive impact decreases as the number of atrocities committed by one party to a conflict increases relative to that committed by other parties.

Loyle & Appel (2017) and Prorok (2017) have made important contributions to understanding the causal effects of national and international-level transitional justice processes on conflict termination, an important prerequisite to having a healthy robust democracy. Their research, however, stops short of telling us how these mechanism affect the long term quality of democracy.

Balcells, Palanza & Voytas (2020) investigate the long-term effect of memorialization efforts, such as the erection of museums devoted to victims of authoritarian repression, on political behavior. Concentrating on Chile, Balcells and her coauthors focus on detecting effects even among generations not directly affected by the former authoritarian regime. They use a field experiment with persons randomly assigned to visit a museum devoted to victims of military rule in Chile and an alternative art museum completely unrelated to politics. The authors find that the museum visit compared to the control has a significant effect on evaluations of the former regime, as well on political behavior more broadly.

While field experiments are not feasible for the assessment of the effects of personnel transitional justice, the robustness of democratic institutions is also one of the questions motivating Capoccia & Pop-Eleches (2020). These authors make use of a natural experiment—the division of Allied-occupied Germany into four zones—to estimate the effects of transitional justice policies on democratic consolidation in post-WW II Germany. Treating the assignment to different denazification policies in each of the four occupation zones as exogenous, they estimate the effects of transitional justice policies on Germans' readiness to uphold democratic values. One key dependent variable probes public support for a one-party political system; a second dependent variable is constructed from a 1957 election survey. Capoccia & Pop-Eleches (2020) ultimately find that differences in the scope and severity of transitional justice implementation have a variety of differing effects on these two democratization indicators. This fascinating and innovative work

is, however, unable to address questions about the broader effect that transitional justice mechanisms may have for countries that differ from postwar Germany.

Our research builds on these initial approaches, recognizing the importance of isolating the effects of transitional justice mechanisms from factors that contributed to the implementation of transitional justice in the first place, as the latter might be highly correlated with quality of democracy indicators.

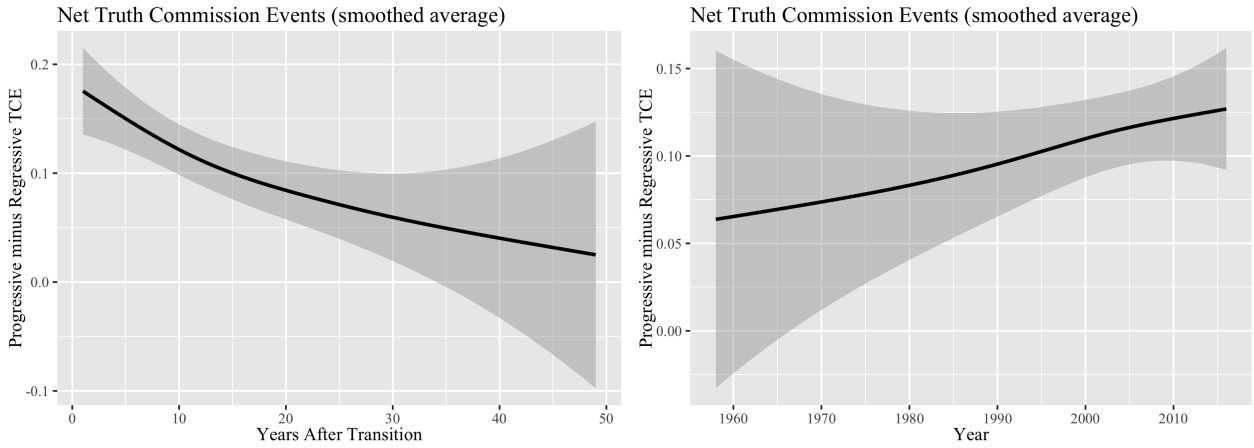
4 Research design: Exploiting time and intensity of truth commissions

In order to test our predictions, we make use of the Global Transitional Justice Dataset (GTJD) by Bates, Cinar & Nalepa (2020), which includes a panel of truth commission events for every country that transitioned to democracy between 1946 and 2016 (there are 81 such countries). The dataset goes above and beyond of merely recording the presence or absence of a truth commission, as it records all events that at an institutional level pertain to truth commissions. The institutional level refers to the fact that the actor initiating the event must be a national government actor. This excludes members of the international community, but also NGOs and local authorities. Second, the event is classified as advancing the work of a truth commission (a positive event) or as inhibiting it (a negative event). For example, a proposal of a bill made to the floor of the legislature to create a truth commission is considered a positive event, while a presidential veto of such a bill constitutes a negative event. Moreover, we assign positive or negative status to events associated with the operation of truth commissions. For example, the circulation of a report by a truth commission is classified as a positive event, while the stepping down of a commission's chairperson prior to the term's expiration or the de-funding of a commission are classified as negative events. Finally, the GTJD team aggregated the positive and negative events annually, creating an unbalanced panel (unbalanced because states transitioned

into democracies at different times and some transitioned out of democracy). Figure 1 presents a snapshot of the dataset as a trend of net (positive minus negative) truth commission events. The panel on the left shows net events in all 81 countries as a function of years lapsed since the transition and the panel on the right shows net events in the same 81 countries but as a function of the year.

Both trends comport with intuitions, showing more events in the first years following the transitions to democracy (on the left) and an increasing popularity of truth commission events over time (on the right).

Figure 1: Disaggregated Truth Commission Data over Time



The GTJD team used numbers of positive and negative events as building blocks for the development of the measure of truth commission severity.[9] In order to focus on the the effects of truth commissions over time we modify slightly the measure of transitional justice severity by assigning to each year a dummy variable that is defined as a function of severity up to that year.

More specifically, we create three such alternative measures. The first, is a "minimalist"

[9]Severity is defined as the total number of positive transitional justice events that occurred in country i over the total number of events plus one (one is added to the denominator to avoid dividing by zero)

measure:

$$TJ^1_{i,t} = \begin{cases} 1 & \text{if } \sum_1^{t-1} P_{i,t} > 0; \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where P_{it} represents the number of positive truth commission events in country i in year $t-1$. This measure corresponds to the transparency inducing effect of having any positive events at all, no matter how isolated. It codes the year with the first positive truth commission event as 1 and continues to code as 1 every subsequent year, regardless of the ensuing trajectory. Theoretically, this measure suggests that although a single positive TC event typically implicates more events that will follow, even this small and isolated event would result in more transparency than no positive events at all. An additional virtue of this measure is its simplicity. We use it in all regressions below as a first model.

Our second measure is more sensitive to the volume of transparency produced by each year's truth commissions effect and corresponds to the idea than more transparency is better for democracy than less transparency. The measure counts net events in country i in year $t-1$:

$$TJ^2_{i,t} = \begin{cases} 1 & \text{if } P_{i,t-1} - N_{i,t-1} > 0; \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Here, additionally, N_{it} represents the number of negative truth commission events in country i in year $t-1$. Our third measure acknowledges the added effect of an accumulation of transparency events from the past. A truth commission issuing a report will have a greater impact if in the preceding years its powers were expanded or its mandate extended. It counts the cumulative net events in year $t-1$ (measured as the total negative events in

year t subtracted from the total positive events in year $t-1$):

$$TJ^3_{i,t} = \begin{cases} 1 & \text{if } \sum_1^{t-1} P_{it} - \sum_1^{t-1} N_{it} > 0; \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

All three measures correspond to the transparency inducing aspects of truth commissions discussed in section 2. They also resemble the original severity measures developed by Bates, Cinar and Nalepa (2020), because they continuously capture the intensity of transitional justice. However, in contrast to their static measure, ours make use of the unfolding of truth commission events over time. In doing so, our measure better leverages the time series aspect of the Global Transitional Justice Dataset.

Figures 2,3, and 4 illustrate how the three measure stack up against one another. Countries are stacked on the vertical dimension, while the horizontal dimension represents years since the democratic transition. Using the example of the third measure, dark blue cells illustrate years in which the cumulative net transitional justice events were positive ($\sum_1^{t-1} P_{it} - \sum_1^{t-1} N_{it} > 0$); light blue cells illustrate years in which the cumulative net transitional justice events were negative or equal to zero ($\sum_1^{t-1} P_{it} - \sum_1^{t-1} N_{it} \leq 0$); and white cells represent censored data stemming from the fact that the country in the corresponding row has not been democratic long enough. The panel corresponding to the minimalist measure indicates that countries vary in how long they have to wait for their first positive event. On the one hand there is Argentina, South Africa, and Tunisia with positive truth commission events occurring within the first year or two after the transition. And on the other, there is Spain with the first positive events occurring 39 years after the transition. Somewhere halfway between these two extremes are Panama and South Korea.

Yet, recall, the minimalist measure only captures whether a first event occurred, ignoring completely everything that follows. Hence, it is not surprising that the net truth commission events measure shows a somewhat different picture. Now we can appreciate the difference between the Tunisian and Argentine truth commission processes, for instance.

In Argentina, positive truth commission events happened early after the transition and ended quickly – reflecting the short (9-month) mandate of the National Commission on the Disappeared (United States Institute of Peace 2011). Tunisia on, the other hand, had more positive than negative events in almost all years following the first, suggesting a more drawn-out transitional justice process. The last panel reveals an even deeper story, showing the events accumulated over time. Now we see that in addition to Tunisia and South Africa, also South Korea and Panama, the two countries that arrived to transparency midway through the transition, actually held steady: once they experienced a positive event, they continued to have more positive than negative events.

Although the net and cumulative measures are sensitive to event volume, the variables themselves are dichotomous severity measures because such indicator variables are easier to interpret in a diff-in-diff framework. Our regressions use one year lags of these measures. However, figures 14 and 15 in Appendix A.5 summarize the coefficients obtained from running the regressions on ten different lags. We motivate the use of lags noting that any plausible model of the effect of transitional justice on the quality of democracy should allow for some lapse of time between the occurrence of truth commission events and the outcome variables, which in our case are political corruption and a measure of the quality of democratic representation.

4.1 The identification strategy

Finding dependent variables for a classical observational study of the relationship between transitional justice and the quality of democracy is very challenging because the implementation of transitional justice is endogenous to phenomena that are used as building blocks of so many democratic indicators, ranging from rule law to freedom from discrimination to freedom to run for office. We address this problem by implementing the difference-in-difference research design.

The classic explanation of the diff-in-diff method relies on comparisons of time trends

Figure 2: Truth commissions at a glance according to minimalist measure

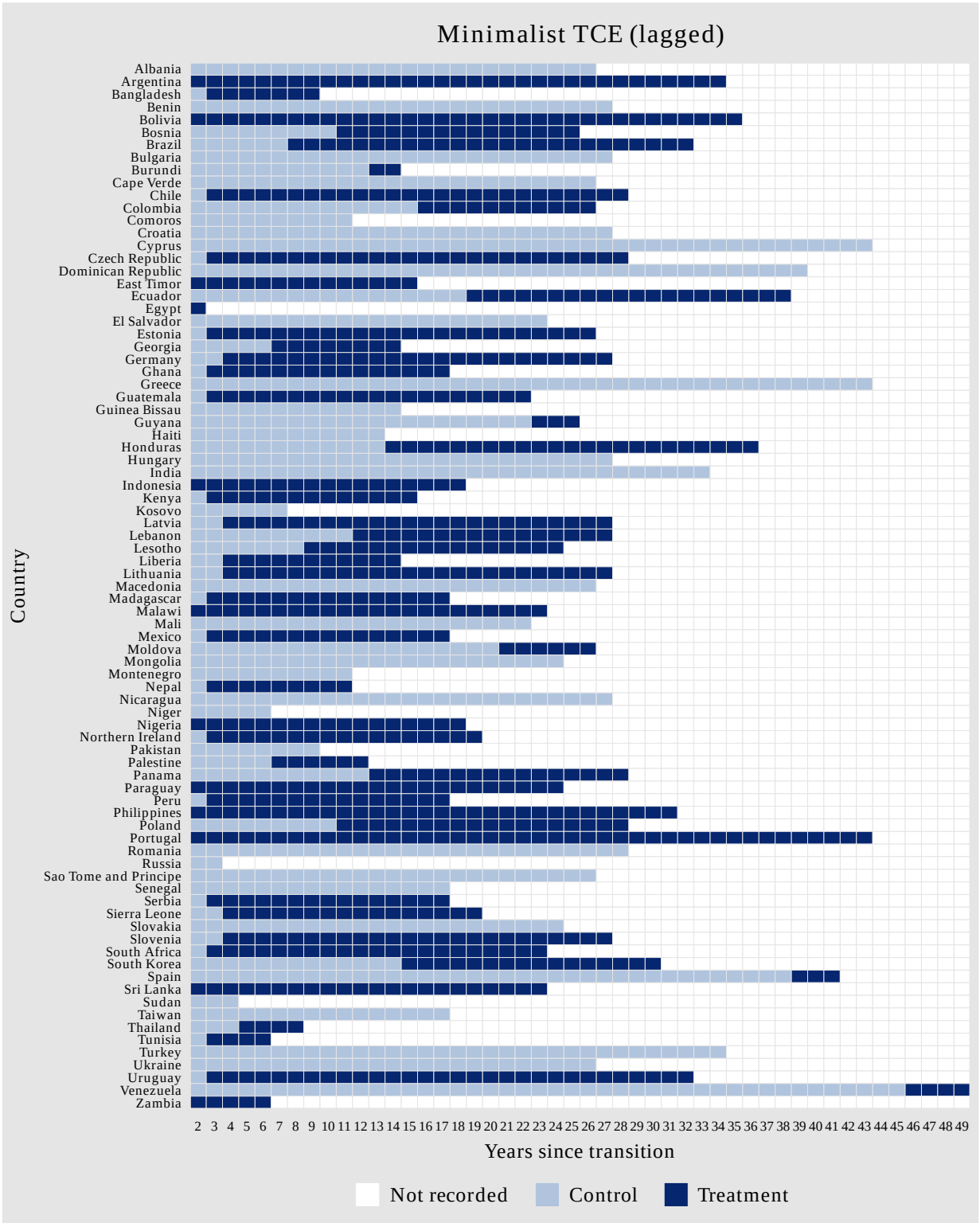
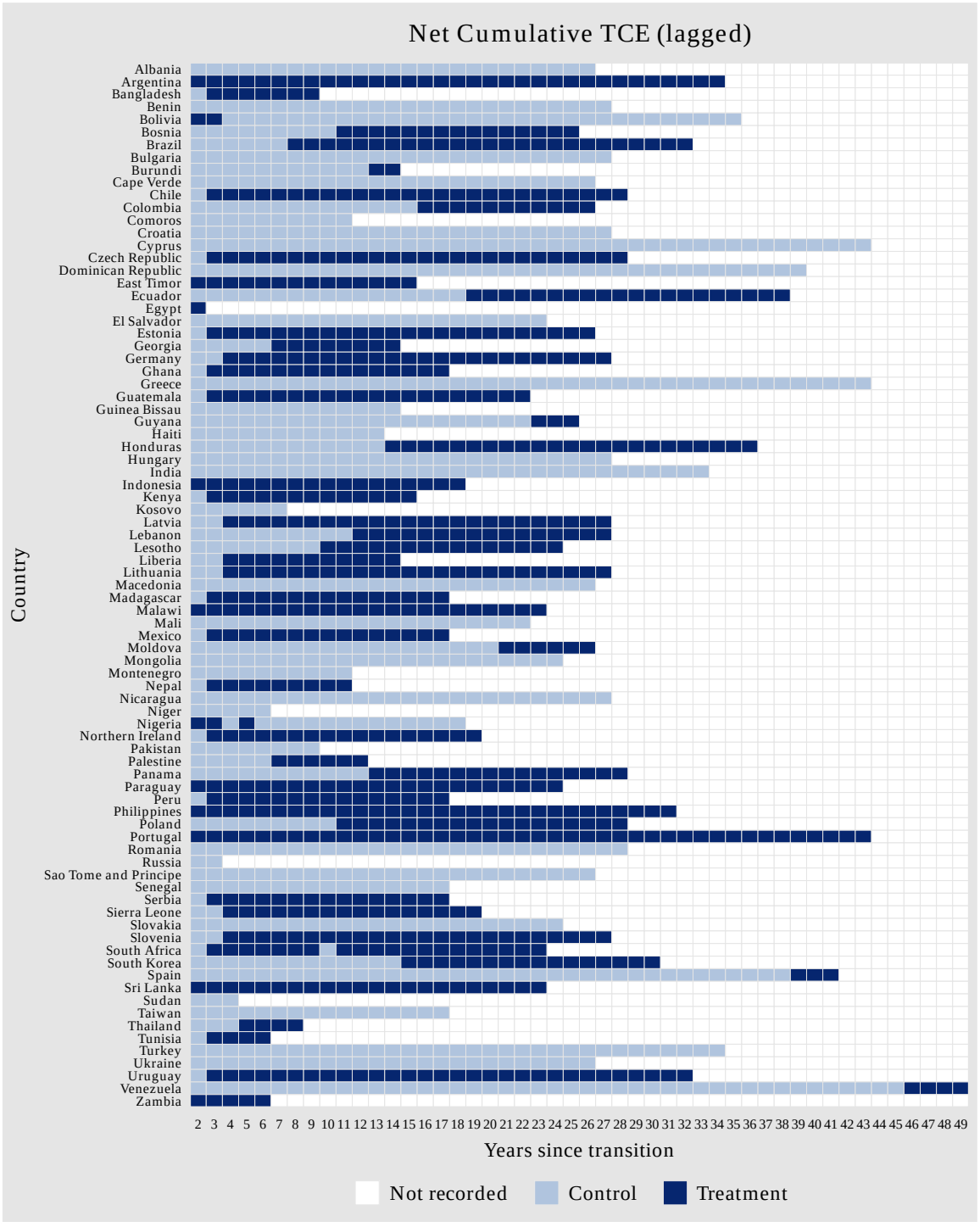


Figure 3: Truth commissions at a glance according to net measure

Figure 4: Truth commissions at a glance according to net cumulative measure



Dark blue cells represent years in which net cumulative truth commission events were positive; light blue cells represent years in which net cumulative truth commission events were negative; white cells represent censored data.

in countries that have been treated with the independent variable of interest—in our case, truth commissions—and those that have not received treatment. Such comparisons are warranted as long as the "treated" and "control" countries are matched to ensure that their pre-treatment trends on the dependent variable are similar enough (satisfying the parallel trends assumption). The case without the treatment serves as a counterfactual to the case with the treatment. Given appropriate data, this method can be generalized to panel data and incorporated into a regression framework. The data requirements, however, are quite stringent, as they require both the dependent and independent variables to vary across countries and over time.

For our dependent variables we use indicators of political corruption and measures of authoritarian turnover from the Varieties of Democracy (V-Dem) project. Both variables match closely the aspects of quality of democracy shaped by truth commissions according to our argument. The first variable—political corruption—captures the transparency effect of preventing blackmail with skeletons closet, as reducing the ability of blackmail to extract policy concessions is exactly what should decrease political corruption. The second variable, which measures the association of economic wealth with political wealth is a classical indicator of the strength of authoritarian networks. Recall that the severing of clandestine ties between members of the authoritarian regime was according to our argument in section 2, the second contribution of truth commissions.

V-Dem is a dataset based on interviews with approximately 5 independent experts per country and codes experts' subjective answers to questions on a host of regime characteristics over time. The characteristics experts are asked about are not directly observable, hence experts' subjective answers cannot be validated by relying on observational data. Yet V-Dem is exceptional among expert surveys in that it corrects for how differences of opinion or mistakes cause experts to diverge in their evaluations. Traditionally, datasets report expert-coded data with means and standard deviations, ignoring the fact that expert reliability and the way in which experts apply ordinal scales to ratings may systematically

vary. V-Dem, however, uses Item Response Theory to model and adjust for differences in how experts apply scales (Pemstein, Marquardt, Tzelgov, Wang & Miri 2015). In order to allow for scaling the independent coding by country experts, V-Dem scholars also encouraged experts to "bridge code" a second or third country. Although experts have less expertise in evaluating these second and third countries than they have at evaluating the countries in which they have primary expertise, this effort allowed V-Dem methodologists to compare the use of the ordinal scales across coders and correct for systematic differences.

V-Dem researchers asked about three thousand country experts hundreds of questions to arrive at 5 general indexes - electoral, liberal, participatory, deliberative, and egalitarian. Since these indices are somewhat broad, we focused on two variables that measure distortions in the quality of representation and the extent to which authoritarian elites have been able to reproduce. The first V-Dem variable that we use for this purpose is the Political Corruption Index (*v2x_corr*); the second is Power Distributed by Socioeconomic Status (*v2pepwrses*). We discuss each in turn below.

4.1.1 Political Corruption Index

Building on our theoretical discussion, we note that we cannot directly observe blackmail with skeletons in the closet. What we can detect, however, are the consequences of blackmail—policy distortions that result from blackmail—and these distortions can be observed as political corruption.

The variable of political corruption is composed by a series of indices that capture the extent of corruption on different bureaucratic offices and branches of the government, specifically executive corruption, public sector corruption, judicial corruption, and legislative corruption.[10]. We have transformed this political corruption index such that it takes values between 0 and 1, ensuring that higher values correspond to higher quality of

[10]See Empirical Appendix, especially section A.4, for a detailed discussion on the specific indices used

democracy–thus lower levels of corruption. While the index in question does not get directly at the the extent to which politicians are influenced by blackmail with secret police files, we argue that it does measure how much they succumb to pressures that impede their ability to represent voters.

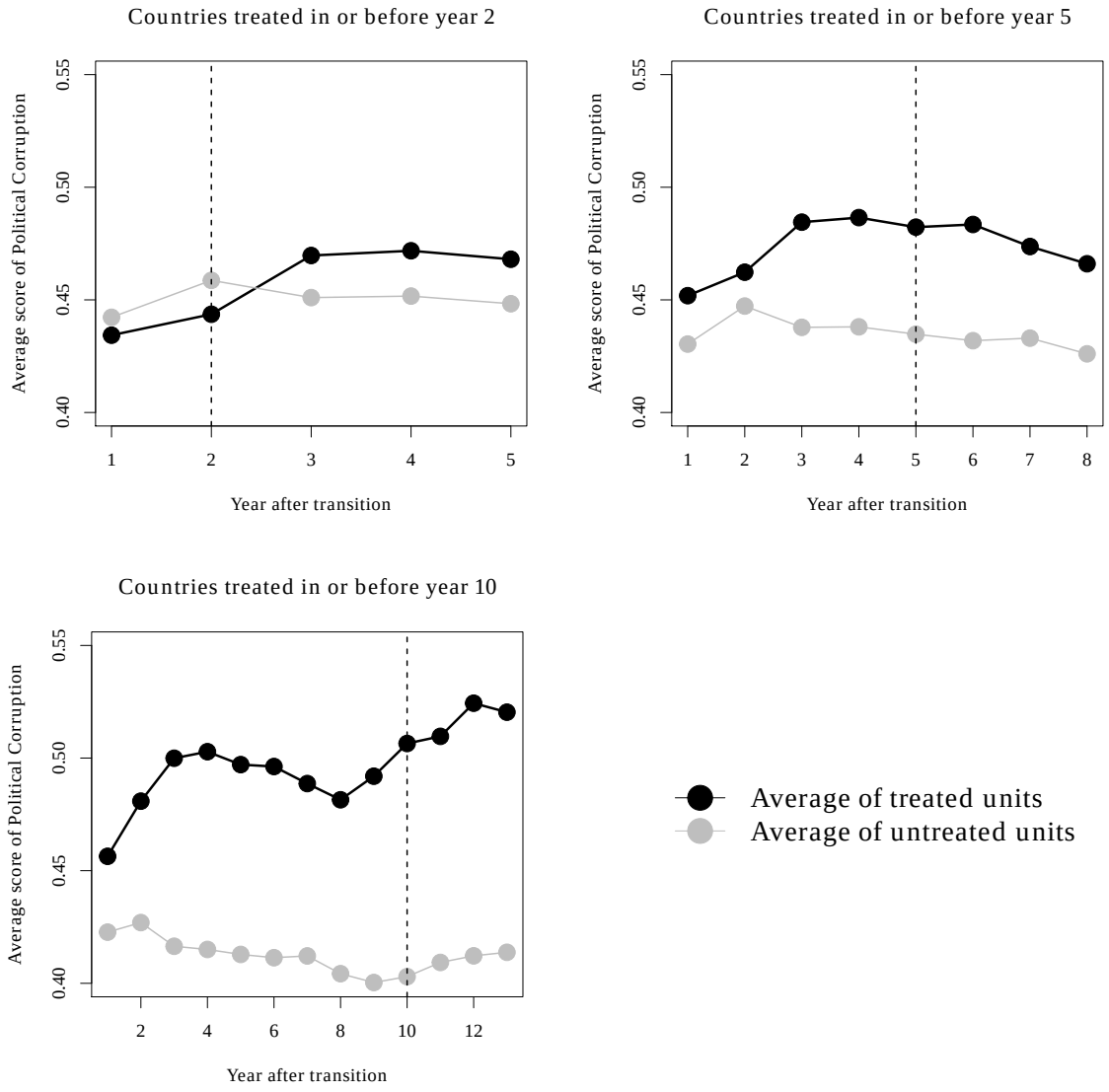
An assumption of the diff-in-diff framework is that, absent treatment, political corruption trends would develop in countries that were never treated with truth commission events according to a similar pattern as countries that were treated. This assumption is commonly referred to as the "parallel trends assumption." In our case, since treatment can occur at any time following the transition to democracy, we need to be specific about defining treatment when demonstrating parallel trends.

Figure 5 compares the average political corruption index of countries for which we have not observed treatment with average values of the political corruption index of countries that were treated within 2, 5, and 10 years following their democratic transition. In all these graphs, we interpret being "treated" as taking the value of 1 in the net truth commission events measure described in equation 3. The control group is constant for all these interpretations of treatment and includes countries that have not had a single TCE, and thus, have never treated.

We first observe that countries that adopted "late" truth commissions (after year 5 or 10), score more democratic on the political corruption measure than countries that were not treated–recall that in our transformation, higher levels mean less corrupt, or more democratic. This corroborates the intuition for engaging this research design in the first place: a pooled analysis across time and space would find a correlation between truth commission events and political corruption, but it would be spurious.[11] Second, the gap between the treated and untreated widens as we extend the period of time during which the treatment is allowed to occur. More concretely, the gap between untreated countries and those treated by year 2 is .015, but about .05 by year 5 and .1 by year 10. The trends

[11] In the Empirical Appendix, we supply such an OLS regression and indeed shows such a significant relationship

Figure 5: Parallel Trends: Political Corruption Index



are mostly parallel prior to the treatment, which is marked with a vertical dashed line.

4.1.2 Political Power Distributed by Socioeconomic Status

The second dependent variable we use taps directly into the ability of authoritarian elites to survive the transition and for their networks to flourish. To understand its suitability, it is useful to consider why the power of former authoritarian elites may extend beyond the life span of an authoritarian regime. Autocrats may be well positioned to cap-

ture state resources at the time of democratic transition, which they can then use to entrench themselves in power (Brun & Diamond 2014, Haggard & Kaufman 2016, Albertus & Menaldo 2014).

The outgoing autocrats' access to resources can be cut off if they or their successors are voted out of office following the transition to democracy. Many autocrats, however, amass their fortunes in secret (The Economist 2021). Even when the kleptocracy is uncovered, the time lapsed before autocrats are actually removed from office grants them opportunities to stash away their fortunes. Various cases from around the world demonstrate, further, that this removal may only be temporary (Kitschelt 1999). Grzymala-Busse (2002), for instance, attributes the revival of successor authoritarian parties to the organizational advantage authoritarian parties hold over parties that are new to the party system. This organizational advantage allows them to make better use of state resources when they eventually do find themselves in government. Transitional justice could be the last resort to curb the privileged position of members or parties of the former autocrats, their collaborators, or their enforcement apparatuses (Stan et al. 2009, David 2011, Vinjamuri & Snyder 2004, Escriba-Folch & Wright 2015).

In light of this argument, truth commissions may plausibly be interpreted as mechanisms preventing former authoritarian elites from holding on to such economic resources. Therefore, a variable measuring the association between economic wealth and political power is an ideal candidate for a dependent variable operationalizing the strength of authoritarian networks. Additionally, given the temporal nature of our data, an ideally suited dependent variable also measures this association over time. Fortunately, the V-Dem Expert Survey contains such a measure.

Called "Power Distributed by Socio-economic Status" (*v2pepwrses* or PdSES), the variable is based on the following question posed to V-Dem experts: "is political power distributed according to socio-economic position?"(Coppedge, Gerring, Lindberg, Skaaning,

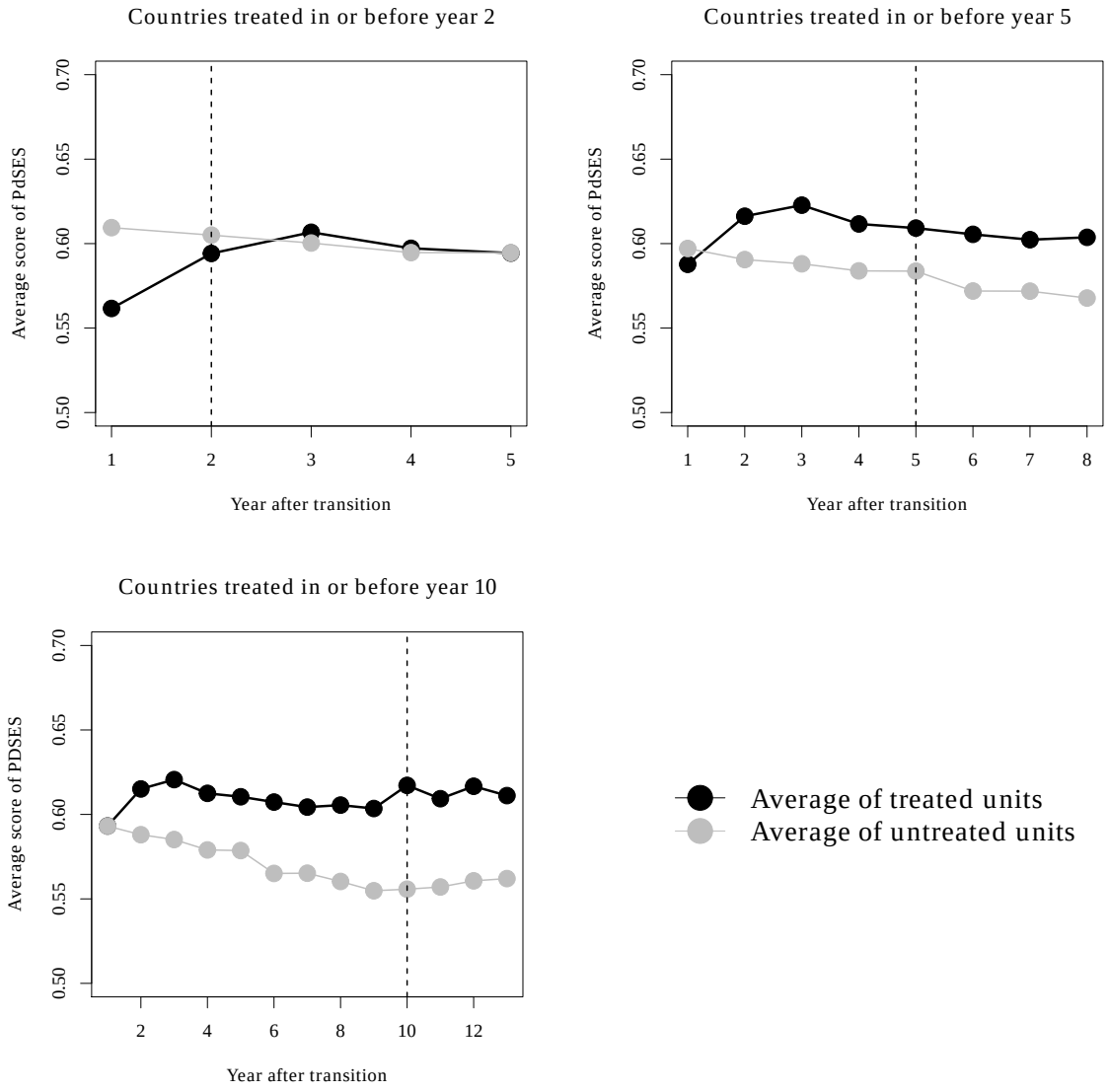
Teorell, Ciobanu & Saxer 2017).[12] In his clarification note, John Gerring elaborates that the measure was designed to gauge the extent to which economic inequalities translate into political power (Coppedge et.al. 2017b). Other scholars of non-democratic regimes have noted that the ability of economic elites to lock in political power is one of the markers of high capacity authoritarian states. In the case of some non-democracies, the link between political and economic power can be fully institutionalized. For instance, Hollenbach (2019), in his article on elite interests and public spending in 19th century Prussia, explains how in the empire's cities, a portion of the male population contributing a third of the tax revenue had the same voting power as the portion of the male population contributing the second third of the revenue and as the portion of the male population contributing the bottom third of the tax revenue, even though, obviously, these groups were numerically unequal. Power Distributed by Socioeconomic Status is also a particularly reasonable measure of quality of democracy for our purposes because while it measures an important aspect of democracy, it is unlikely to be correlated with rule of law, which could also affect the implementation of transitional justice. If the goal of transitional justice is to undermine the privileged position of authoritarian elites, this score should increase with the severity of the transitional justice mechanism in question. Similarly than with political corruption, we have transformed the original variables to range between 0 and 1, and for higher values to indicate higher quality of democracy (lower levels of association between political and socioeconomic power).

We again start with investigating the parallel trends assumption using PdSES as the outcome variable. Truth commissions have been operationalized as a treatment in exactly

[12] Answers to the question were distributed along a 5-point scale. The possible answers included (0)"Wealthy people enjoy a virtual monopoly on political power. Average and poorer people have almost no influence"; (1)"Wealthy people enjoy a dominant hold on political power. People of average income have little say. Poorer people have essentially no influence"; (2)"Wealthy people have a very strong hold on political power. People of average or poorer income have some degree of influence but only on issues that matter less for wealthy people"; (3)"Wealthy people have more political power than others. But people of average income have almost as much influence and poor people also have a significant degree of political power"; and (4)"Wealthy people have no more political power than those whose economic status is average or poor. Political power is more or less equally distributed across economic groups" (Coppedge et.al. 2017b).

the same way as in figure 5, so using cumulative net events.

Figure 6: Parallel Trends: PdSES



Comparing the PdSES trends of countries treated and untreated with truth commissions reveals that for all but very early adopters, the trends before treatment are largely parallel. These trends clearly indicate that a diff-in-diff framework is an appropriate approach to investigating the effect of truth commissions on the quality of democracy measured by the association of political and economic wealth captured by PdSES. They also suggest that truth commissions may help weaken authoritarian networks to the extent that

these are captured by PdSES.

It is worth keeping in mind while interpreting these figures that there are many factors (economic wealth key among them) in addition to transitional justice that affect either political corruption trends or the extent to which political and economic power go together. Since the diff-in-diff design calls for time varying covariates, many factors—such as historical legacies—are impossible to control for. However, we can control in all regressions for GDP per capita, implying that the parallel trends assumption must only hold conditional on GDP per capita.[13] At the same time, because presenting parallel trends graphs conditional on continuously-measured variables is very difficult to do, the graphs of parallel trends presented above do not take GDP per capita into account.

4.2 The statistical model

A traditional difference-in-difference framework with just one pre-treatment and one post-treatment period would estimate:

$$Y_{i,t} = \mu_i M_i + \lambda T + \gamma D_{i,t} + \beta X_{i,t} + \epsilon_i$$

where $D_i = 1$ if country i experienced transitional justice of a given type, and $D_i = 0$ otherwise. M_i is the country dummy and assumes 1 when the TJ event is associated with country i.[14] In the classic difference-in-difference set-up, there are only two periods: $T = 0$ for the pre-treatment period and $T = 1$ for the post-treatment period. Consequently, μ_i can be interpreted as the country intercept and λ as the post-treatment intercept. $X_{i,t}$ represents the set of covariates upon which the treatment effect is conditioned. Note that a consequence of the above notation is that $D_{i,t} = M_i * T$.

The panel structure of our data calls for two important adjustments to the model. First,

[13] Although there are other economic variables we could collect data on, such as economic inequality, existing measures—for instance, the Gini coefficient—do not vary much with time and so would not be technically useful in our research design

[14] Note that there are one fewer dummies than there are countries. Albania is our baseline country

each country receives the treatment at a different time t (that is, each country has a different year that marks the pre- and post-treatment period). Second, each country appears in the dataset for multiple periods. To correctly specify the diff-in-diff model with multiple time periods, we build on Angrist & Pischke (2008) and Besley & Burgess (2004), and propose to estimate the following model:

$$Y_{i,t} = \mu_i M_i + \lambda_t T_t + \gamma D_{i,t-1} + \beta X_{i,t} + \epsilon_i \tag{4}$$

This is essentially a two-way fixed effect regression where μ_i represents country-specific fixed effects and λ_t represents year-since-transition specific fixed effects. T is the time period dummy and is equal to 0 in every year lapsed since the transition except for T in row $Y_{i,T}$. Note that in the regression framework above, we use $D_{i,t-1}$, the treatment from the year preceding the year in which the dependent variable was recorded. $D_j, t-1$ can also be written as $\mu_i * \lambda_{t-1}$; it takes the value of 1 when country i was treated in period $t-1$. $X_{i,t}$ represents the set of covariates conditioning the treatment. Here, we use a single covariate: GDP per capita.

In a paper from 2018, Andrew Goodman-Bacon shows that this general difference-in-difference estimator is a weighted average of all possible two-group and two-period diff-in-diff estimators in the data, with the greatest weights given to treatments that occurred close to the "middle of the panel" (Goodman-Bacon 2018). This is because the weights are directly proportional to the treatment variances.[15]

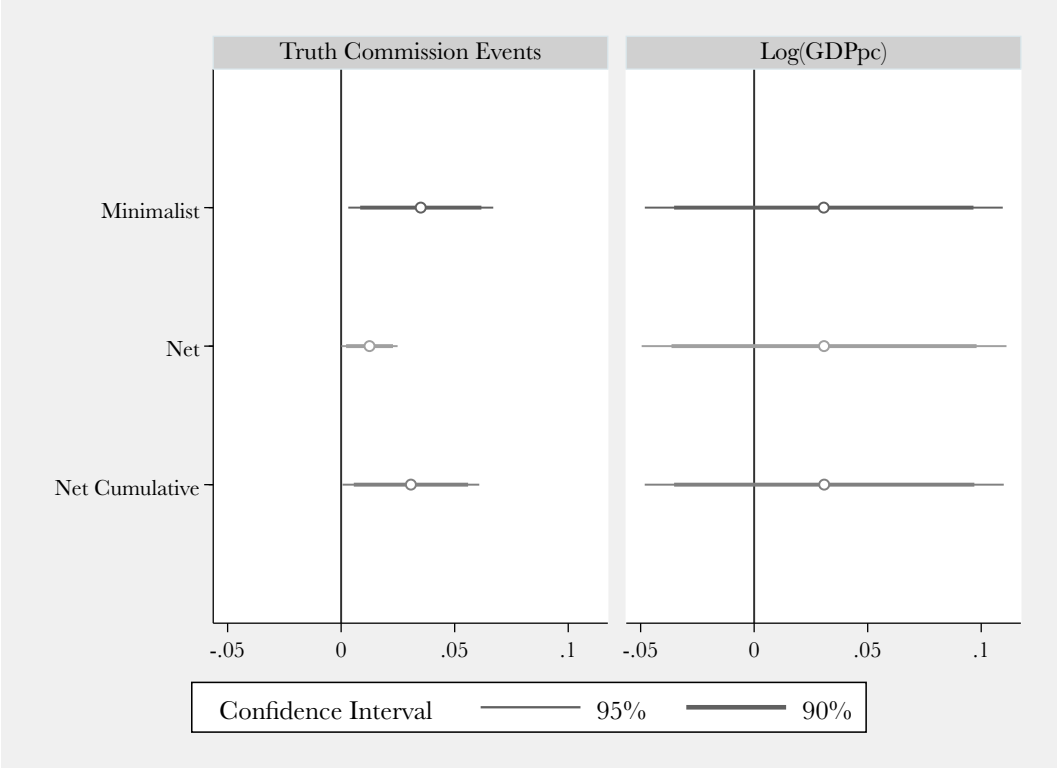
4.3 Results and Discussion

We begin by estimating the effect of truth commissions on the political corruption index. Figure 7 summarizes three models, each one of them pertaining to one of the three lagged measures of truth commission severity and controlling for GDP per capita, logged (re-

[15] In our case, these would be the cases treated within the first five years after the transition compared to those that were never treated.

sults in table form are included in the Empirical Appendix, table 2). The figure shows that the average effect of truth commissions on political corruption varies according to the operationalization of TCE, but it is a positive effect in all cases. Relative to countries with no positive events of truth commissions at all (minimalist), having one positive event decreases corruption by almost 3.5 percentage points. Countries with with more positive than negative events (net) see an smaller uptick (1.2 percentage points). When truth commission events are operationalized as the net cumulative number of events, the estimated effect is 3 percent points.

Figure 7: Effect of Truth Commission Events (TCE) on Political Corruption



Notes: Truth Commission Events are lagged. All models include country and year after transition intercepts, and standard errors clustered by country.
Thicker confidence intervals represent statistical significance at a higher level; thinner confidence intervals represent statistical significance at a lower level.

We are aware that the average effect estimated might be interpreted as substantively small–after all, the political corruption index can (theoretically) range from 0 to 1. However, these estimates already take into account country-specific (time-invariant) effects

and year-after-transition specific effects. Both of these are likely already capturing important variation in political corruption, so the effect associated with TCEs is modest but relevant in this context.

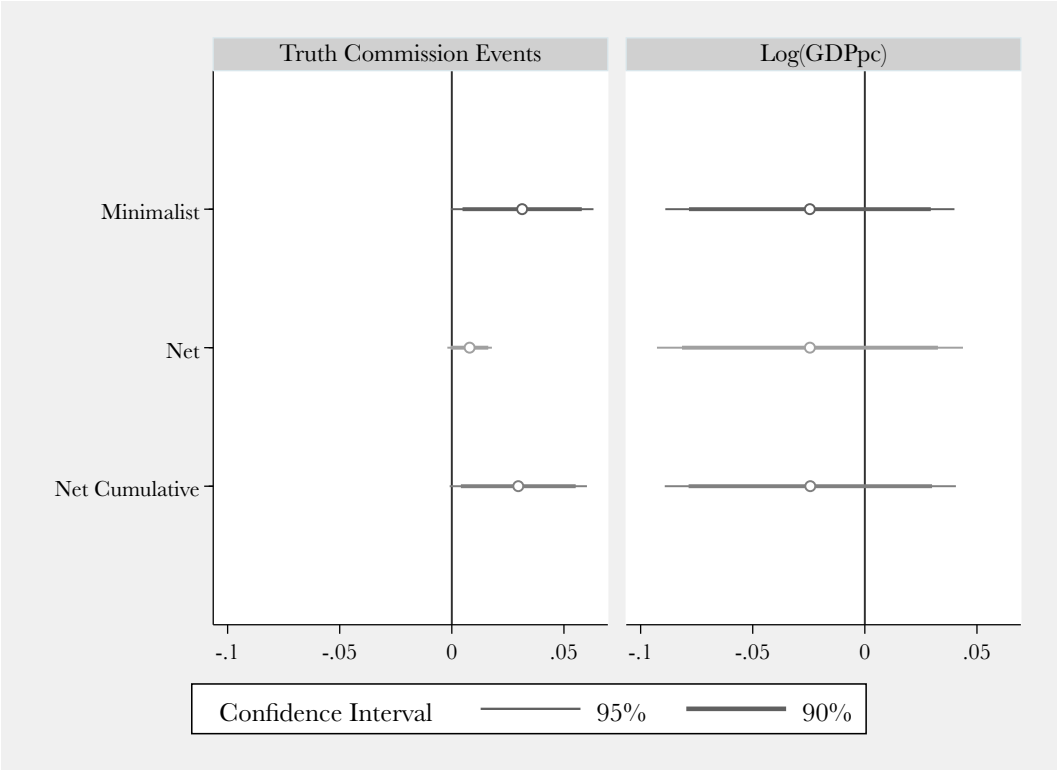
In addition, political corruption captures the extent to which public officials can be blackmailed with the threat of revealing secrets from their authoritarian past, but we are aware that this dependent variable is quite complex. Therefore, it is likely that the index reflects a number of other political dynamics besides the one we propose here. To account for this, we replicated this analysis using the sub-components of the corruption index: legislative corruption, judicial corruption, public sector corruption, and executive corruption.[16] Each of these sub-components reflect a narrower dimension of corruption and focus on specific bureaucrats and administrators of the state. These estimations, all of which are consistent with what we have reported here, can be consulted in the Appendix.

As discussed, we believe that the mechanism suggested–that truth commissions prevent blackmail of politicians, thus improving the quality of democracy–would also affect the association between economic wealth and political power. To account for this possibility, we use the V-Dem variable "Power Distributed by Socioeconomic Status," or PdSES described in detail in section 4.1.2. This variable captures whether former authoritarian elites are able to resist turnover and continue to accumulate political and economic wealth. This variable taps into a very different aspect of quality of democracy than corruption, namely the ability of authoritarian networks to survive the transition. According to our theory we would expect truth commissions to decrease the link between political and economic power by weakening or breaking these authoritarian networks.

Figure 8 summarizes results from the estimations of interest (results in table form are reported in the Empirical Appendix, table 3). The figure below focuses only on the coefficients corresponding to PdSES's effect our three different measures of truth commission severity: minimalist, net events per year, and cumulative net events per year.

[16] See Empirical Appendix for a detailed description of these sub-components.

Figure 8: Effect of Truth Commission Events (TCE) on PdSES



Notes: Truth Commission Events are lagged. All models include country and year after transition intercepts, and standard errors clustered by country
Thicker confidence intervals represent statistical significance at a higher level; thinner confidence intervals represent statistical significance at a lower level.

The results depicted here are consistent, although less conclusive–than the results of the regressions with political corruption as the dependent variable. In all operationalizations, truth commissions are associated with higher levels of democracy, although these results are only significant at the 90% confidence level. Again the substantive effects are modest, ranging from 3.1 percent (for the minimalist operationalization of truth commissions) to 0.7 for the effect of net events. Taken together, these results support the main expectation described previously: truth commissions are an effective tool for dismantling authoritarian networks.

4.4 Robustness check: Synthetic matching

We conclude by presenting a series of synthetic matching estimations as a robustness check for our estimations above. The underlying logic of synthetic matching is to "borrow" information from a number of untreated units (that is, countries that have not established a truth commission) to create a synthetic counterfactual that can be compared to the trajectory of a treated unit or country. The observed untreated countries are compared on a number of covariates to the treated country during the pre-treatment period to create a series of weights. These weights will then be assigned to each of the untreated units to create the synthetic control, such that a weighted average of the outcome variable that spans both the pre- and post-treatment period. Following this process, we are able to compare the trajectory of the dependent variable of the treated unit to the synthetic match (Abadie, Diamond & Hainmueller 2015).

In its original presentation, synthetic matching requires a balanced panel and a single treated unit. In addition, since the method estimates the weights based on the pre-treatment trajectory, the analyst must collect data on the untreated units from a number of time periods before the treatment occurred. Since we observe more years of those countries that transitioned earlier, our data is unbalanced panel. In addition, we have multiple treated units, and some countries experience truth commission events immediately upon

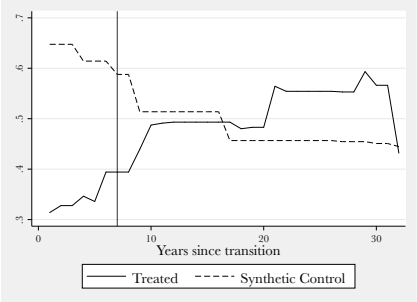
transition whereas some take years to implement them, meaning that the time of treatment varies by country. Despite these obstacles, we were able to process the data to conduct a conservative synthetic matching robustness check proceeding as follows. We use the minimalist measure, as the synthetic control method is not designed for units that "fall out of treatment". which our other measures allow for.

In the first step, we select countries that experienced their first positive truth commission event four or more years following the democratic transition (the rationale here is to have enough observations in the pre-treatment period). There were eight such countries in our dataset: Brazil, Burundi, Georgia, Honduras, Lebanon, Lesotho, South Korean, and Thailand. The second step involved creating synthetic control groups for each of these countries. In order to be part of the synthetic control a country must not only be untreated, but also it must have a democratic tenure that is at least the same as the treated country. In the third step, we use GDP per capita as the matching covariate. Assigning the appropriate weights results in a balanced panel, where we are able to estimate the synthetic matching parameters.

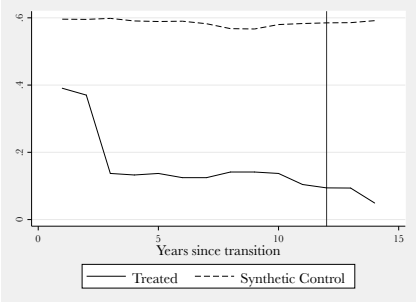
Figures 9 and 10 compare each of the eight treated countries with their corresponding synthetic control according to political corruption and PdSES, our two outcome variables. In each subfigure, a vertical line marks the year since transition when the country first experienced a positive truth commission event. In the case of political corruption, we see that in the pre-treatment period, all eight (eventually) treated countries are more politically corrupt that then their synthetic controls. However—as we argue in response to the treatment—some outperform the synthetic control over time (see in particular, Georgia, Brazil and South Korea). The remaining countries, meanwhile, close the gap in political corruption separating them from their synthetic controls. The sole exception to this pattern is Burundi. However, Burundi is an outlier in the sense that its first positive truth commission event occurred late—about 12 years following the democratic transition to democracy.

Figure 9: Synthetic matching for selected countries (outcome: Political Corruption)

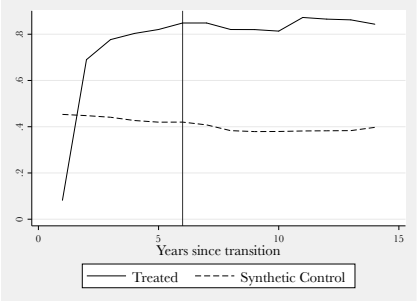
(a) Brazil



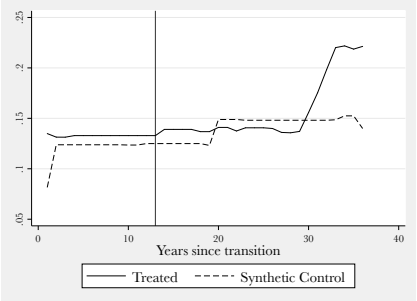
(b) Burundi



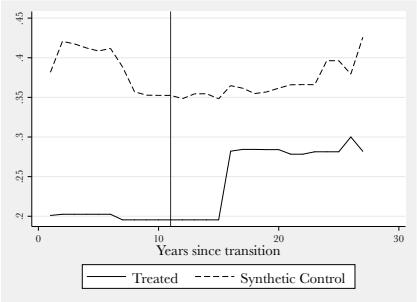
(c) Georgia



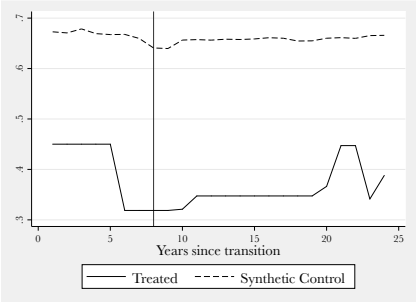
(d) Honduras



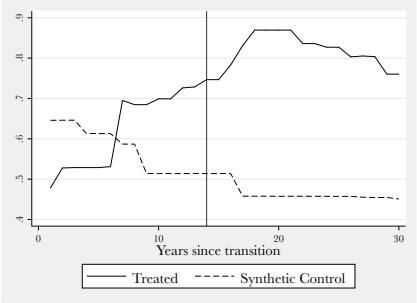
(e) Lebanon



(f) Lesotho



(g) South Korea



(h) Thailand

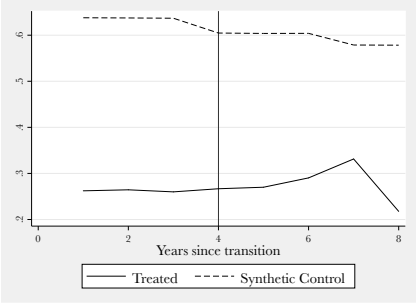
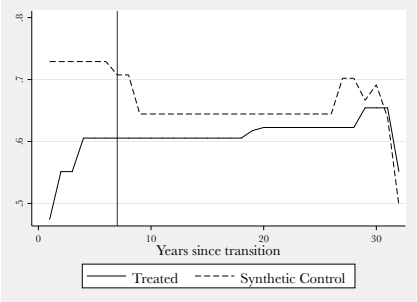
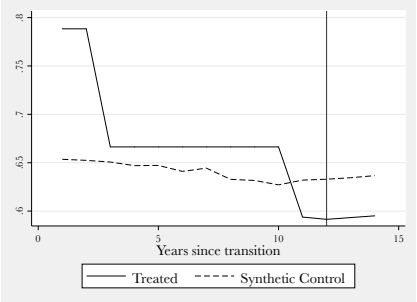


Figure 10: Synthetic matching for selected countries (outcome: PdSES)

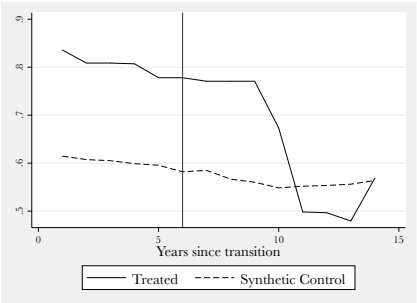
(a) Brazil



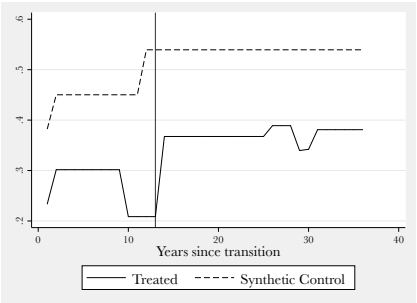
(b) Burundi



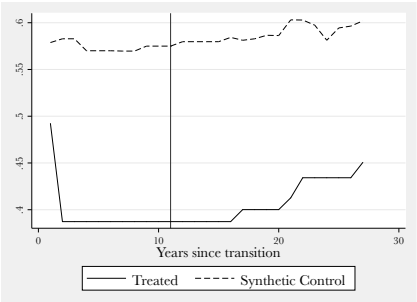
(c) Georgia



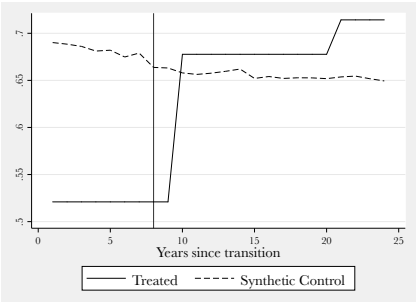
(d) Honduras



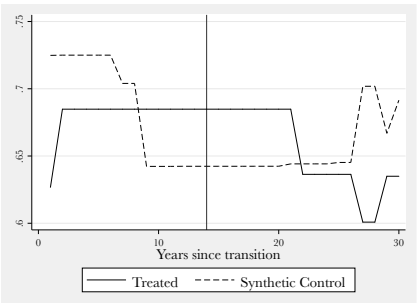
(e) Lebanon



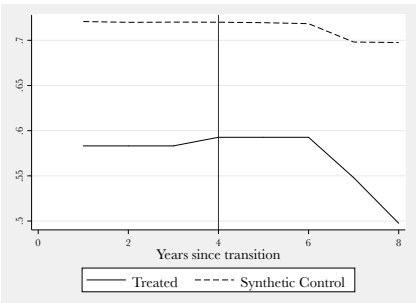
(f) Lesotho



(g) South Korea



(h) Thailand



Performing a similar analysis, but using PdSES as the outcome variable paints a slightly different picture. As Figure 10 indicates, the cases of Georgia, South Korea and Thailand suggest a worsening quality of democracy according to PdSES. According to estimations in panels c,g, and h, treated countries (represented by the solid line) exhibit steeper downward trends than countries in the control group (represented by dashed lines). In the cases of Brazil, Burundi and Honduras, although at the time of treatment, the difference between treated and untreated is vast, they draw tighter to one another. Finally, Lebanon and Lesotho are the only two countries that change trajectories in the way that our theory would expect. Consequently, the synthetic matching results using PdSES partially support our hypotheses, which is most likely due to the limited number of countries that allow for synthetic matching.

Our theory is ambivalent regarding how long it takes for a truth commission to reduce political corruption, but presenting complete trajectories of 8 separate treated countries against synthetic controls in Figures 9 and 10, suggests a time frame for when this effect of truth commissions might materialize, resulting in divergence between the trajectories of treated and untreated countries. For the specific eight countries portrayed here, the exact moment when an uptick by TCE is induced is particularly easy to observe, simply by eye-balling the charts. However, we caution readers in interpreting these charts as each of these treated countries is observed for different periods of time and is compared with a different feasible set of synthetic controls. In the Empirical Appendix A5, we further address the question of time that it takes for a truth commission to have an effect, by replicating the main dif-in-dif regressions for 10 alternative lags.

5 Conclusion

Transparency enhancing mechanisms, such as truth commissions, are notoriously left out of the classical critiques of transitional justice. When Huntington writes that "[e]ven if a

moral and legal argument could be made for prosecution, this would fall before the normative imperative of creating a stable democracy" he is assuming that transitional justice can only be punitive, and therefore, never democracy-enhancing. In this paper we show that truth commissions through their transparency effects improve the quality of democracy. The mechanism through which transparency accomplishes these goals, we argue, is first, by preventing blackmail of former members and collaborators of the ancien régime with 'skeletons in the closet' and reducing political corruption. Second, and bolstering our first argument, "doing nothing" although it need not produce immediate negative consequences, may, over time, strengthen the power of authoritarian networks, particularly the networks involving secret legacies of the authoritarian regime. Damaging information collected by the former authoritarian secret police for the benefit of authoritarian elites may turn elected politicians into clients of agents who threaten to reveal these politicians' "skeletons in the closet."

Our paper is also a first attempt at testing a causal theory of transitional justice with a global dataset that disaggregates truth commission events across time. This allows us to detect the possibility of delayed effects and also suggest a natural way of measuring intensity or severity of truth commissions.

Our data analysis, carried out within a difference-in-difference framework, supports our main expectation that truth commissions improve the quality of democracy. These results complement existing research carried out by Bates, Cinar & Nalepa (2020) and Ang & Nalepa (2019) in an Hierarchical Modeling framework, which found that lustration has a positive effect on the quality of democracy. The HLM framework however, did not isolate the effect over time, nor did it identify the effects of the truth commissions as we do here, but focused on lustrations. While lustration and truth commissions are both transparency enhancing, there are important differences in the ways they operate. Compared to truth commissions, lustration requires large mass of positive events relative to negative events and thus, a measure picking up such intensity is called for. As indicated earlier, the

diff-in-diff framework compelled us to force a count measure of transitional justice events into an indicator variable. Future work could address theoretical expectations about truth commissions and lustration policies more specifically, exploring in detail the mechanisms by which each truth revelation procedure operates to increase or decrease the quality of democracy.

Finally, we underscore our paper's empirical contribution to the growing research agenda on transitional justice datasets and democratic outcomes. Recent years has seen a number of transitional justice datasets (Binningsbø, Loyle, Gates & Elster 2012, Van der Merwe, Baxter & Chapman 2009, Thoms, Ron & Paris 2010, Loyle & Binningsbø 2018). This literature has focused on identifying the effects of a variety of TJ measures on variables like trust in governmental institutions (Horne 2017), peace (Binningsbø et al. 2012), or democratic stability (Olsen, Payne & Reiter 2010). Our approach enriches this research by accounting for the dynamic nature of truth-commission implementation and by test a mechanisms that is notoriously hard to detect.

References

Abadie, Alberto, Alexis Diamond & Jens Hainmueller. 2015. "Comparative politics and the synthetic control method." *American Journal of Political Science* 59(2):495–510.

Albertus, Michael & Victor Menaldo. 2014. "Gaming democracy: elite dominance during transition and the prospects for redistribution." *British Journal of Political Science* 44(3):575–603.

Ang, Milena & Monika Nalepa. 2019. "Can transitional justice improve the quality of representation in new democracies?" *World Politics* 71(4):631–666.

Angrist, Joshua D & Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Bakiner, Onur. 2016. *Truth Commissions: Memory, Power, and Legitimacy*. University of Pennsylvania Press.

Balcells, Laia, Valeria Palanza & Elsa Voytas. 2020. "Do transitional justice museums persuade visitors? Evidence from a field experiment." *Journal of Politics* Forthcoming.

Bates, Genevieve, Ipek Cinar & Monika Nalepa. 2020. "Accountability by Numbers: A New Global Transitional Justice Dataset (1946–2016)." *Perspectives on Politics* 18(1):161–184.

Besley, Timothy & Robin Burgess. 2004. "Can labor regulation hinder economic performance? Evidence from India." *The Quarterly journal of economics* 119(1):91–134.

Betz, Timm, Scott J Cook & Florian M Hollenbach. 2019. "Spatial interdependence and instrumental variable models." *Political Science Research and Methods* .

Binningsbø, Helga Malmin, Cyanne E Loyle, Scott Gates & Jon Elster. 2012. "Armed conflict and post-conflict justice, 1946–2006: A dataset." *Journal of Peace Research* 49(5):731–740.

Brun, Diego Abente & Larry Diamond. 2014. *Clientelism, social policy, and the quality of democracy*. JHU Press.

Capoccia, Giovanni & Grigore Pop-Eleches. 2020. "Democracy and retribution: Transitional justice and regime support in postwar West Germany." *Comparative Political Studies* 53(3-4):399–433.

Coppedge, Michael, John Gerring, Staffan I Lindberg, Svend-Erik Skaaning, Jan Teorell, Vlad Ion Ciobanu & Laura Saxer. 2017. "V-Dem Country Coding Units V7.".

David, Roman. 2011. *Lustration and transitional justice: Personnel systems in the Czech Republic, Hungary, and Poland*. University of Pennsylvania Press.

Escriba-Folch, Abel & Joseph Wright. 2015. "Human rights prosecutions and autocratic survival." *International Organization* 69(02):343–373.

Gibson, James L. 2006. "Overcoming apartheid: Can truth reconcile a divided nation?" *The Annals of the American Academy of Political and Social Science* 603(1):82–110.

Gibson, James L & Amanda Gouws. 1999. "Truth and reconciliation in South Africa: Attributions of blame and the struggle over apartheid." *American Political Science Review* 93(3):501–517.

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with variation in treatment timing.".

Grzymala-Busse, Anna M. 2002. *Redeeming the communist past: The regeneration of communist parties in East Central Europe*. Cambridge University Press.

Haggard, Stephan & Robert R Kaufman. 2016. *Dictators and democrats: masses, elites, and regime change*. Princeton University Press.

Hayner, Priscilla B. 1994. "Fifteen truth commissions-1974 to 1994: a comparative study." *Hum. Rts. Q.* 16:597.

Hayner, Priscilla B. 2010. *Unspeakable Truths 2e: Transitional Justice and the Challenge of Truth Commissions*. Routledge.

Hollenbach, Florian M. 2019. "Elite interests and public spending: Evidence from Prussian cities." *The Review of International Organizations* pp. 1–23.

Horne, Cynthia. 2017. "Vetting, Purges, and Lustration: Measurement Choices and Empirical Implications.".

Horne, Cynthia M & Margaret Levi. 2004. Does lustration promote trustworthy governance? An exploration of the experience of Central and Eastern Europe. In *Building a trustworthy state in post-socialist transition*. Springer pp. 52–74.

Human Rights Watch. 1991. "Human Rights in Post-Invasion Panama: Justice Delayed is Justice Denied.".
URL: *https://www.refworld.org/docid/45d316002.html*

Huntington, Samuel P. 1993. *The third wave: Democratization in the late twentieth century*. Vol. 4 University of Oklahoma press.

Keesing's Record of World Events. 1998. "South Africa.".

Kitschelt, Herbert. 1999. *Post-communist party systems: competition, representation, and inter-party cooperation*. Cambridge University Press.

Loyle, Cyanne E & Benjamin J Appel. 2017. "Conflict Recurrence and Postconflict Justice: Addressing Motivations and Opportunities for Sustainable Peace." *International Studies Quarterly* 61(3):690–703.

Loyle, Cyanne E & Helga Malmin Binningsbø. 2018. "Justice during armed conflict: A new dataset on government and rebel strategies." *Journal of conflict resolution* 62(2):442–466.

Nalepa, Monika & Konstantin Sonin. 2020. "How Does Kompromat Affect Politics? A Model of Transparency Regimes." *A Model of Transparency Regimes (March 23, 2020). University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-29).

Olsen, Tricia D, Leigh A Payne & Andrew G Reiter. 2010. "The justice balance: When transitional justice improves human rights and democracy." *Human Rights Quarterly* 32(4):980–1007.

Pemstein, Daniel, Kyle L Marquardt, Eitan Tzelgov, Yi-ting Wang & Farhad Miri. 2015. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data.".

Price, Robert M. 1991. *The apartheid state in crisis: Political transformation in South Africa, 1975-1990*. Oxford University Press New York.

Prorok, Alyssa K. 2017. "The (in) compatibility of peace and justice? The International Criminal Court and civil conflict termination." *International organization* 71(2):213–243.

Snyder, Jack & Leslie Vinjamuri. 2004. "Trials and errors: Principle and pragmatism in strategies of international justice." *International security* 28(3):5–44.

Stan, Lavinia et al. 2009. *Transitional justice in Eastern Europe and the former Soviet Union: Reckoning with the communist past*. Routledge.

The Economist, journal=The Economist. 2021. "The world must not accept the jailing of Alexei Navalny." pp. 39–40.

Thoms, Oskar NT, James Ron & Roland Paris. 2010. "State-level effects of transitional justice: what do we know?" *International Journal of Transitional Justice* 4(3):329–354.

United States Institute of Peace. 2011. "Truth Commission: Argentina.".

Van der Merwe, Hugo, Victoria Baxter & Audrey R Chapman. 2009. *Assessing the impact of transitional justice: Challenges for empirical research*. US Institute of Peace Press.

Vinjamuri, Leslie & Jack Snyder. 2004. "Advocacy and scholarship in the study of international war crime tribunals and transitional justice." *Annual Review of Political Science* 7:345–362.

Zvobgo, Kelebogile. 2019*a*. "Demanding Truth: The Global Transitional Justice Network and the Creation of Truth Commissions." *Unpublished manuscript* .

Zvobgo, Kelebogile. 2019*b*. "Designing truth: Facilitating perpetrator testimony at truth commissions." *Journal of Human Rights* 18(1):92–110.

A Empirical Appendix

This appendix is composed of five sections. In the first, we present in full detail the regression tables used to construct the main results figures in the main text: Figure 7 and Figure 8. In the second, we conduct a simple OLS regression on the pooled data, just as a benchmark for our diff-in-diff analysis. In the third, we estimate the same models as in the main text, but removing one country at a time, to ensure that the robustness of our results. In the fourth, we run diff-in-diff analysis on all the subcomponents of the political corruption index. Finally, it the fifth section, we explore the first ten lags of truth commissions events (as opposed to just the first lag) used in the same diff-in-diff framework.

Table 1: Descriptive statistics

	Min.	1stQ.	Med.	Mean	3rdQ.	Max.	Obs.
Political Corruption	0.05	0.26	0.44	0.48	0.69	0.99	1923
PdSES+	0.04	0.51	0.62	0.60	0.68	0.98	1923
Minimatist TCE	0.00	0.00	0.00	0.48	1.00	1.00	1923
Net TCE	0.00	0.00	0.00	0.08	0.00	1.00	1923
Net Cumulative TCE	0.00	0.00	0.00	0.45	1.00	1.00	1923
log(GDPpc)	5.39	7.40	8.43	8.33	9.28	10.73	1923
Year since transition	1.00	6.00	13.00	14.15	21.00	49.00	1923
Public Sector Corruption	0.05	0.33	0.53	0.54	0.77	0.99	1923
Executive Corruption	0.04	0.31	0.51	0.54	0.76	0.99	1923
Legislative Corruptio	0.08	0.31	0.40	0.43	0.54	0.90	1923
Judicial Corruption	0.09	0.36	0.45	0.49	0.63	0.94	1923

A.1 Table of Results

Table 2: Effect of Truth Commission Events (TCE) on Political Corruption

	Minimalist	Net	Net Cumulative
TCE	0.0351*	0.0125*	0.0307*
	(0.016)	(0.006)	(0.015)
Log(GDPpc)	0.0307	0.0308	0.0309
	(0.040)	(0.040)	(0.040)
Constant	-0.229	-0.190	-0.226
	(0.325)	(0.318)	(0.326)
Observations	1747	1747	1747
R^2	0.960	0.960	0.960

Standard errors in parentheses

Truth Commission Events are lagged

All models include country and year specific intercepts

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Effect of Truth Commission Events (TCE) on PdSES

	Minimalist	Net	Net Cumulative
TCE	0.0314	0.00794	0.0296
	(0.016)	(0.005)	(0.015)
Log(GDPpc)	-0.0245	-0.0245	-0.0243
	(0.032)	(0.034)	(0.033)
Constant	0.774**	0.792**	0.775**
	(0.267)	(0.268)	(0.268)
Observations	1747	1747	1747
R^2	0.862	0.859	0.861

Standard errors in parentheses

Truth Commission Events are lagged

All models include country and year after transition specific intercepts

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.2 OLS with pooled data

We begin by presenting the table below, which shows the results of a pooled linear regression model. This table shows our estimates of the association between truth commission events (TCE) and the two outcomes of interest, including GDP logged as a covariate. The results below show a positive association between the minimalist measure of truth commission events and the net cumulative measure and political corruption and PdSES. When we use the measure of net TCE, however, the association is negative although statistically indistiguishable from 0. Of course, these results use a panel dataset without proper fixes to account for non-independence, and without controlling for selection into treatment. As such, these results are illustrative of a general trend but cannot be treated as conclusive.

Table 4: Effect of TCE on Democractic Quality - Pooled models

	PC	PC	PC	PdSES	PdSES	PdSES
Minimalist TCE	0.0509***			0.0206***		
	(5.35)			(3.63)		
Net TCE		0.0329			0.00529	
		(1.91)			(0.52)	
Net Cumulative TCE			0.0648***			0.0110
			(6.80)			(1.92)
Log(GDPpc)	0.136***	0.138***	0.135***	0.0396***	0.0403***	0.0397***
	(34.90)	(35.26)	(34.55)	(16.98)	(17.29)	(16.93)
Constant	-0.682***	-0.677***	-0.675***	0.259***	0.262***	0.263***
	(-20.77)	(-20.43)	(-20.68)	(13.20)	(13.31)	(13.37)
Observations	1747	1747	1747	1747	1747	1747

t statistics in parentheses

PC- Political Corruption

PdSES - Power Distributed by Socioeconomic Status

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.3 Removing one country at a time

We also estimated the same models but removing one country at a time to make sure that our results are robust.

Figure 11: Removing one country at a time

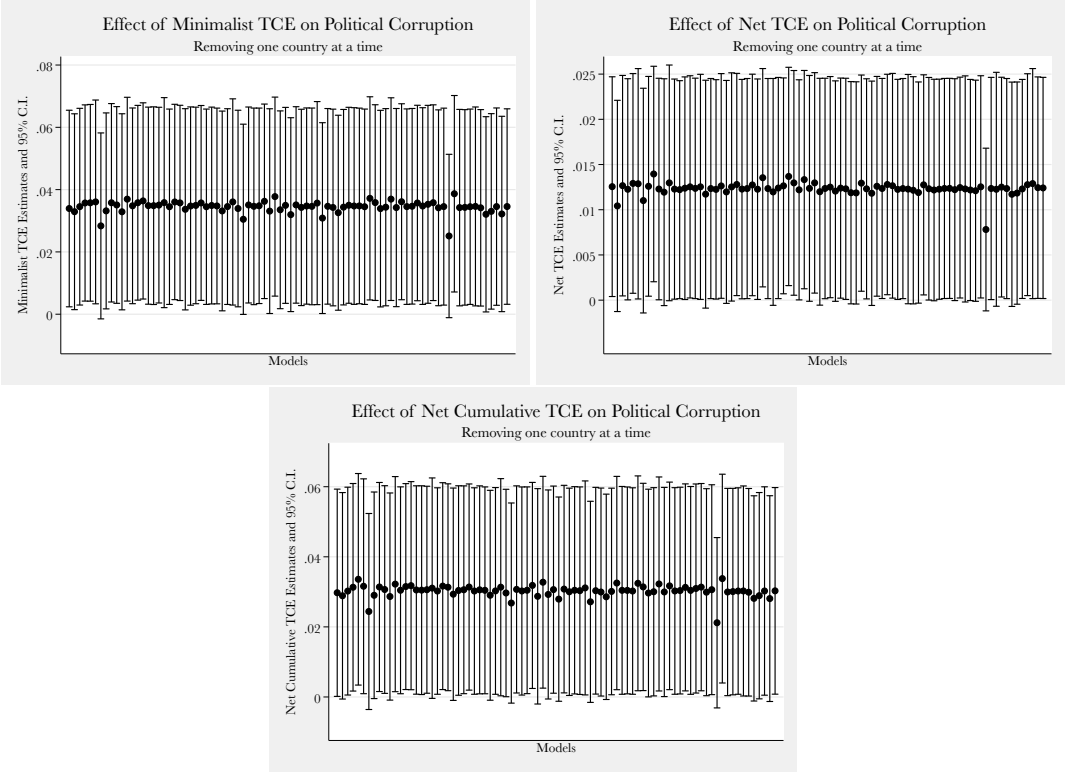
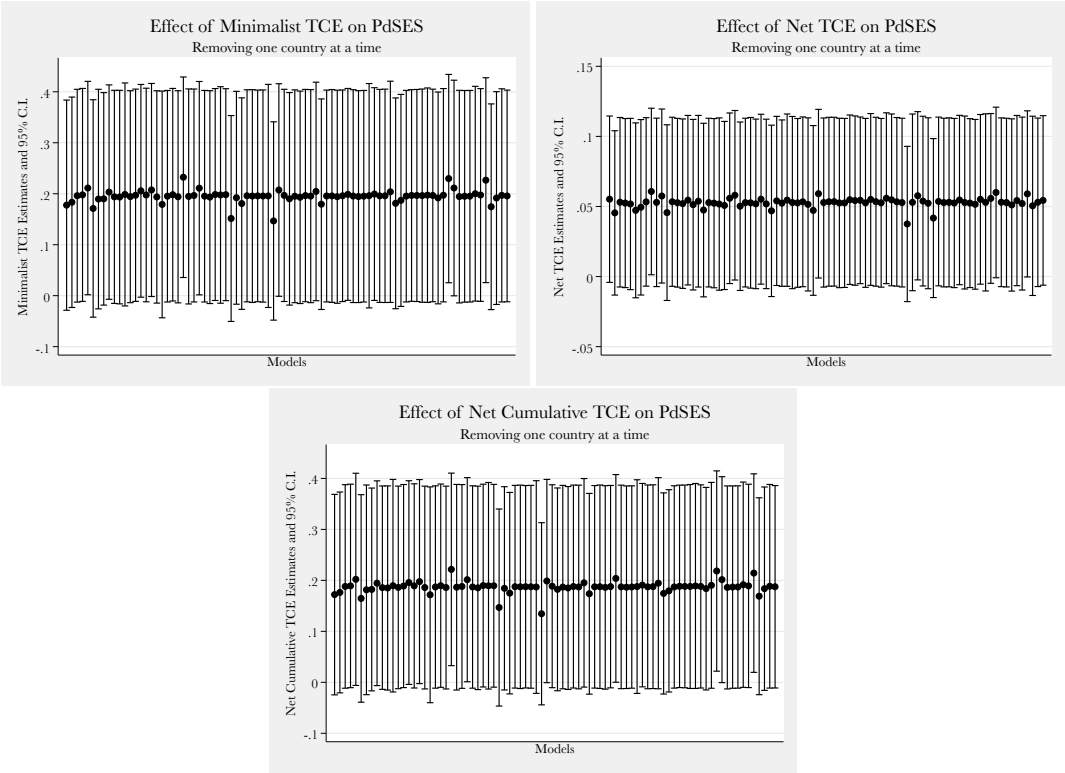


Figure 12: Removing one country at a time



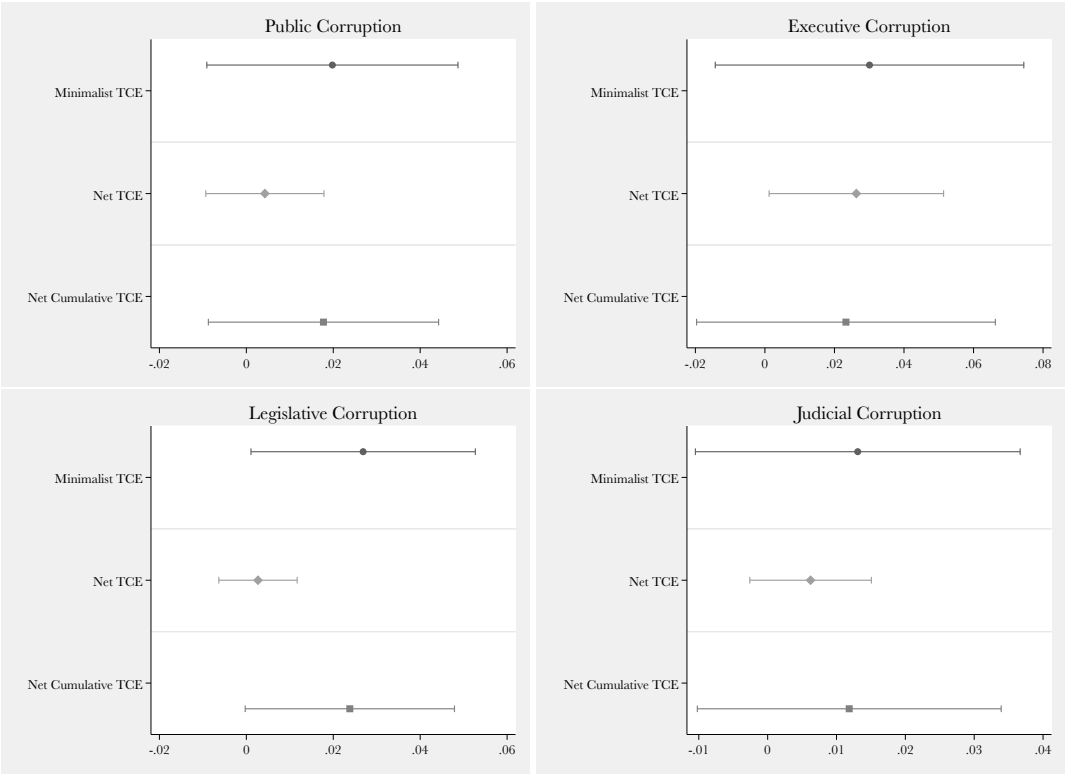
A.4 Decomposing the Political Corruption Index

Given that the political corruption index is composed by 4 sub-components–executive, public, judicial, and legislative corruption–we also present a similar set of estimations than in the paper (equation 2) but using each of their sub-components as outcomes of interest:

> Public sector corruption (`v2xpubcorr` in the V-Dem datasets): To what extent do public sector employees grant favors in exchange for bribes, kickbacks, or other material inducements, and how often do they steal, embezzle, or misappropriate public funds or other state resources for personal or family use?
>
> Executive corruption (`v2xexecorr`): How routinely do members of the executive, or their agents grant favors in exchange for bribes, kickbacks, or other material inducements, and how often do they steal, embezzle, or misappropriate public funds or other state resources for personal or family use?
>
> Legislative corruption (`v2xlgcrrpt`): Do members of the legislature abuse their position for financial gain?
>
> Judicial corruption (`v2jucorrdc`): How often do individuals or businesses make undocumented extra payments or bribes in order to speed up or delay the process or to obtain a favorable judicial decision?

Our measure of cumulative net truth commissions has a positive effect on reducing each of the components of corruption, although some—most notably Public Corruption and Executive corruption—have much narrower confidence intervals. Note, that even though the sub-components were originally measured on different scales, they have been normalized for comparability to a 0-to-1 scale.

Figure 13: Coefficients of TCE on the Four Components of the Political Corruption Index on Truth



All models include country and year after transition intercepts, and standard errors are clustered by country

A.5 Robustness check: Expanding the lags in the diff-in-diff

Last but not least, we consider, as a robustness check, models that include lags for different numbers of years. We do this because our previous analyses rest on the assumption that a one-year lag is sufficient to capture the start-up costs of truth commissions. It may be presumptuous to think that truth commissions start having an effect as early as one year after the first truth commission event occurs. Therefore, in this section, we report the coefficients on ten different lags used in our diff-in-diff framework. We use our measure of net cumulative truth commission events. Figure 14 presents the results for political corruption and Figure 15 illustrates the coefficients for power distributed by socioeconomic status. As can be seen in both figures, our decision to lag our truth commission variable by one year in all of the above analyses was reasonable.

Figure 14: Diff-in-Diff regression coefficients for 10 lags, Net Cumulative TCE (political corruption)

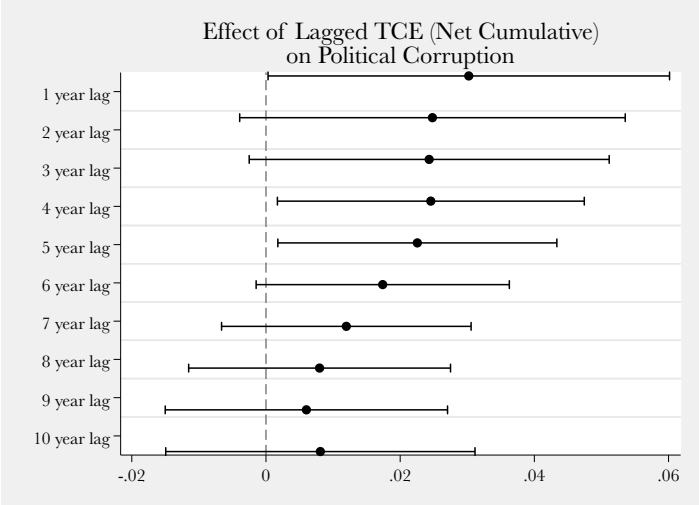
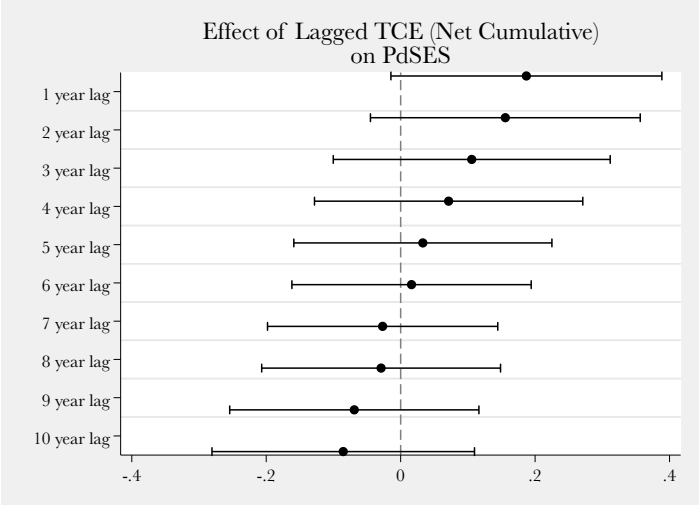



Figure 15: Diff-in-Diff regression coefficients for 10 lags, Net Cumulative TCE (PdSES)