UNIVERSITY OF CAMBRIDGE
Faculty of Economics

JANEWAY INSTITUTE CAMBRIDGE

CAMBRIDGE WORKING PAPERS IN ECONOMICS
JANEWAY INSTITUTE WORKING PAPERS

Building Bridges to Peace: A Quantitative Evaluation of Power-Sharing Agreements

Hannes Mueller
IAE-CSIC, BSE, CEPR

Christopher Rauh
University of Cambridge, CEPR, HCEO

Abstract

Power-sharing agreements are used as a tool to reduce political violence in regions of conflict, but agreements are often followed by violence. This is due to the fact that such agreements are introduced during periods of political violence when a country is inside the conflict trap, which makes it difficult to distinguish the effect of the agreement from the political context that generates persistent political violence. In this study we match on pre-agreement conflict risk to estimate the effects of power-sharing agreements on violence using a difference-in-difference method. The results show that violence falls immediately after an agreement, with the effects strengthening over time. Comprehensive agreements tend to be particularly successful. We show that broader institutional changes that have their nucleus in the agreements are crucial elements explaining these large changes.

Reference Details

2261 Cambridge Working Papers in Economics
2227 Janeway Institute Working Paper Series

Published 27 October 2022
Revised 04 August 2023

Websites www.econ.cam.ac.uk/cwpe
www.janeway.econ.cam.ac.uk/working-papers

Building Bridges to Peace: A Quantitative Evaluation of Power-Sharing Agreements

Hannes Mueller and Christopher Rauh

IAE-CSIC, BSE, CEPR, PRIO and University of Cambridge, CEPR, HCEO, IZA, PRIO

Abstract

Power-sharing agreements are used as a tool to reduce political violence in regions of conflict, but agreements are often followed by violence. This is due to the fact that such agreements are introduced during periods of political violence when a country is inside the conflict trap, which makes it difficult to distinguish the effect of the agreement from the political context that generates persistent political violence. In this study we match on pre-agreement conflict risk to estimate the effects of power-sharing agreements on violence using a difference-in-difference method. The results show that violence falls immediately after an agreement, with the effects strengthening over time. Comprehensive agreements tend to be particularly

We are grateful to Melina Ailayi for excellent research assistant work for this study and Kilian Kamkar for his great help. We also thank Björn Gehrmann and Anne-Kristin Treiber for enlightening discussions and detailed feedback on earlier drafts, and the Group of Friends (GoF) of Mediation at the United Nations for fruitful discussions. We thank two anonymous referees for their great reviews, Tommaso Frattini for his excellent discussion and participants for their comments at the Economic Policy conference, John Wallis and Toke Aidt for their feedback, and Mathias Thoenig for his guidance on the revision. The authors acknowledge and thank the Keynes Fund for support of this project. The views shared in this paper do not necessarily reflect the official policy position of the German Federal Foreign Office. The German Federal Foreign Office (GFFO) provided funding for the study through the Stabilisation Platform (SPF)/GIZ.

successful. We show that broader institutional changes that have their nucleus in the agreements are crucial elements explaining these large changes.

1. Introduction

Political violence is associated with tremendous human suffering of the directly exposed individuals, population displacements and long-term scarring of the affected economies. It is therefore no surprise that substantial effort goes into avoiding armed conflict or negotiating an end to it. A main policy carried into peace processes all around the world are power-sharing agreements. Power sharing refers to sharing of political power by different groups in the same government, like, for example, in a coalition government. Power sharing is a central pillar of armed conflict mediation and de-escalation attempts worldwide and is thus employed as a tool to reduce political violence. A large majority of peace agreements include power-sharing provisions, the most recent ones include Libya, the Mindanao agreement and Colombia.

But does power sharing work in practice to reduce political violence? At face value, the news is not good. Almost 90% of power-sharing agreements do not result in a complete halt of political violence, and there is some controversy regarding their role in key cases such as Afghanistan and Iraq, where conflict parties lack political legitimacy amongst the international community.

Quantitative studies should be able to provide answers but have been hampered by the fact that power-sharing agreements are not agreed upon in a political vacuum. They are the result of the specific national, regional and geopolitical configurations in which they are agreed upon. Power-sharing agreements are introduced during periods of intense political violence in an explicit attempt to formalize a (re-)distribution of power - one that is more congruent with the actual distribution of power and resources in a given country - and thereby reduces the continuation of violence.

It is a known fact in the conflict literature that countries can fall into the so-called conflict trap, which is very difficult to escape (Collier and Sambanis 2002, Rohner and Thoenig 2021, Mueller and Rauh 2022a). The trap is characterized by repeated cycles of political violence.[1] Most power-sharing agreements are agreed while the country is inside the trap, with the intent of breaking it. Figure 1 shows the share of countries experiencing violence in the months leading up to an agreement on the left and the average level of violence measured in battle deaths on the right. Both measures of violence drop in the months following the agreement,

[1] Throughout we use the definition of Uppsala Conflict Data Program (UCDP) for armed political violence. The basic unit of analysis for the UCDP's Georeferenced Event Dataset (GED) dataset is the "event". UCDP defines an event as: "An incident where armed force was by an organised actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date". See Sundberg et al (2013) and Croicu and Sundberg (2016) for more details.

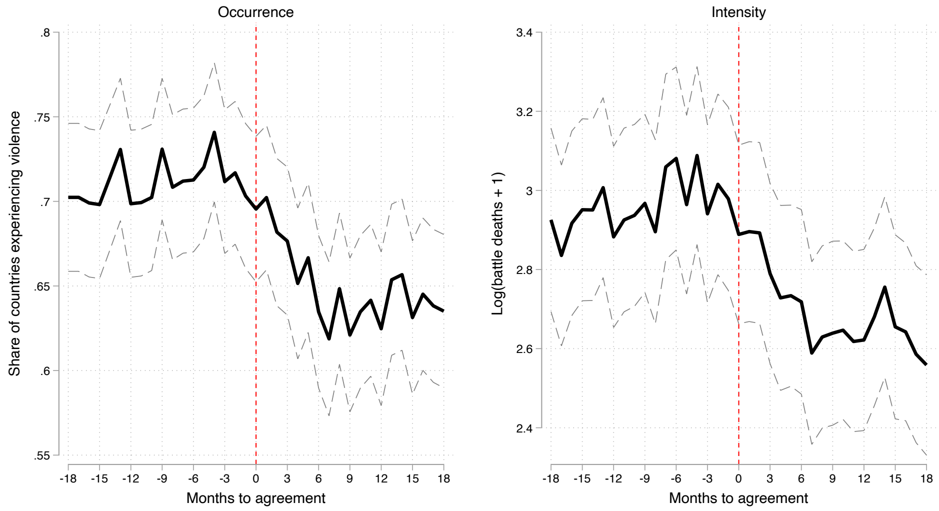

Figure 1: Violence occurrence and intensity before and after the introduction of agreements

but are still much higher than in a random group of countries that do not adopt agreements. Therefore, it is hard to distinguish the effect of the power-sharing agreement from the general political context that generates persistent political violence, i.e., the conflict trap. Without explicit handling of this endogeneity problem, the fact that high political risk is followed by both violence and an agreement, any attempt to measure effects will lead to a biased estimate of the effect of power sharing on political violence.

We study power-sharing agreements using the PAX dataset which provides quantitative analysis of the text of agreements. A primary distinction in the PAX coding of agreements is whether an agreement is *comprehensive*. We provide an overview of how these comprehensive agreements are different. Our study then uses a matched difference-in-difference method to estimate the effects of power-sharing agreements on violence in civil wars. The staggered and repeated treatment of these agreements presents a challenge to standard methods of difference-in-difference estimates.

We therefore propose a difference-in-difference method that focuses on situations around the adoption date and matches these situations with a control group using a risk forecast before the adoption. Our method first extracts event windows with 6, 12 and 18 months before and after treatment without overlap. We then construct a set of non-overlapping control windows of the same size using a sampling method based on the distribution of violence intensity forecasts in the months before power-sharing agreements. The control group is sampled to generate placebo event windows which have the same distribution of violence risk before the adoption date. In this way we match situations holding constant the distribution of violence risk before the (placebo) adoption date. We then use one of the new standard methods for difference-in-difference estimators developed by Callaway and Sant'Anna (2021) to show how violence evolves before and after the adoption date compared to the matched control group.

The results show a clear pattern in which violence falls immediately after an agreement. The effects strengthen over time and the point estimates suggest large treatment effects towards the end of the event windows. For comprehensive agreements we find, for example, that after 12 months the occurrence of violence falls by 20 percentage points whereas violence intensity falls by 60%. Importantly, we find no clear pre-trends before the adoption of power-sharing agreements. Also, we do not find that our own forecasts based on Mueller and Rauh (2022a) that track the news environment of countries are able to anticipate the effect of power-sharing agreements. Overall, we find no evidence that conflict escalations in anticipation of power-sharing agreements could be driving our results. The study confirms the robustness of these results using various aspects of agreements, 6, 12 and 18-month windows, alternative ways of defining event windows, and different ways of constructing the control group. Importantly, our results are also robust to controlling for the presence of peacekeepers. Instead, our results strongly suggest that studies of peacekeeping need to pay attention to the context of political agreements preceding the sending of troops.

We then zoom out of the immediate aftermath of power-sharing agreements to get an understanding of the broader institutional features associated with violence reductions. We use simple country fixed effect regressions to show that changes in institutional features of democracy are associated with reductions in political violence. Specifically, we analyze the Variants of Democracy (V-Dem) dataset, which distinguishes five different components of democracy (deliberative, participatory, liberal, electoral, and egalitarian). Our analysis suggests that all components of democracy measured by V-Dem are associated with reductions in violence, but the strongest associations are found in the liberal, electoral, and egalitarian components.

Digging deeper we find that the elements that are most closely related to violence reductions are the absence of exclusion across political, social, socio-economic, gender or geographic dimensions. These variables capture access to power, public services, justice, and civil liberties and whether these are restricted for specific groups. Improvements in the strength and neutrality of the legal system and public administration, fair access to public sector jobs and business opportunities, and strong and equal access to justice are specific institutional features that are strongly associated with reductions in violence.

We then turn back towards our difference-in-difference method to understand whether power-sharing agreements can be the starting point for these broader changes. Our analysis suggests that, in particular, comprehensive agreements are followed by institutional changes which are associated with reductions in political violence. This is significant in two ways. First, changes in the broader institutional set up might be a channel through which power-sharing agreements work, so that comprehensive agreements lead to larger violence reductions because they lead to larger knock-on effects on institutions. If this is true, then policymakers would need to keep these in mind when advising on power-sharing agreements. Secondly, these findings fit extremely well to the current view in policy circles and the academic literature that agreements are needed to be the result of a bargain that matches the underlying distribution of power. In this view, addressing exclusion through institutional changes is important because it brings de jure and de facto power more in line with each other.[2]

One of the key contributions of this article is the use of forecasts for the purpose of causal identification. The conflict risk forecast we use, follows the methodology of our webpage conflictforecast.org closely. Risk is estimated using a rolling forecast methodology which uses news text and violence dynamics in a random forest model to predict conflict outbreaks and intensity for the following 12 months into the future. The forecast takes the information set of a given month, say May 2015, and forecasts forward, then takes the next time-step to June 2015 and repeats these step for all months in the period 2000-2020. This gives us the ability to capture the risk of future outbreaks under the information set available at the time. When we match our treated group to our control group, we, therefore, explicitly control for the expected level of violence. Put differently, the method we deploy not only allows us to look at pre-trends but explicitly captures the degree to which power-sharing agreements are positive surprises. To further deal with endogeneity concerns we also conduct an exercise in which we predict power-sharing agreements using machine learning and match observations using the derived risk score.

In the following section we first turn towards these discussions in the academic and policy literature before discussing the empirical challenge and presenting data, our empirical methodology and results.

2. Related literature

In this section we will discuss the existing literature on the role of power-sharing in ending armed violence. In Section 4 we will turn towards the literature that discusses the conditions under which power-sharing occurs.

Both amongst practitioners and academics there is some doubt that peace settlements can bring peace.[3] Yet, there are at least two channels through which power sharing can help bring

[2] De jure power is political power allocated by political institutions (such as constitutions or electoral terms) whereas de facto power emerges from informal sources of power like the ability to engage in collective action, wealth, or the ability to wield coercive power (Acemoglu and Robinson 2006).

[3] German government contacts in the mediation unit stressed this point several times. An early literature in political science also postulates that agreements are likely to be worse than victories (Licklider 1995). Along similar lines Srinivasan (2021), using the case of the Sudans, argues that "idealised constitutional texts" are resisted by local actors and can fuel violence. For broader context see the discussion of Krasner (2002) on the role of international law in the realism school of international relations.

peace: providing a bargaining solution for a given distribution of power and providing a commitment device for intractable bargaining situations.

In the first view, the key to the success of power sharing is that institutional arrangements need to track a specific balance of power in the country. Providing an institutional arrangement which shares de jure power according to de facto power ensures peace. The leading group accepts the legally binding version of the agreement only because the threats of other groups in the coalition to organize a violent uprising is credible and very strong. This is the core mechanism described in Francois et al (2015) who demonstrate that power-sharing coalitions in Africa follow the strength of groups in the country. The need for institutional adaptation is supported by the finding that ceasefires alone do not appear to reduce violence substantially (Armand et al 2023). The influential report by Cheng et al (2018) makes the same point by modelling power sharing within a framework of *elite bargaining* in a *limited access order*,[4] i.e. they stress the role played by the rent distribution underpinning peace agreements. In their view, institutions also need to reflect the underlying configuration of power and resources. Where this is not satisfied, the incentives for violence increase.

There is also a more subtle role for institutions during ongoing bargaining. Fearon (1995) posits that one of the reasons for political violence is that dynamic shifts in power do not allow for a bargaining solution. For the government it is easier to repress groups in society that are gaining strength rather than negotiating and sharing power. The problem in these circumstances is that the weaker group today will want to renegotiate tomorrow from a more powerful position and cannot credibly commit not to do this. Without a way to commit, violence can break out. Acemoglu and Robinson (2001, 2005) apply this logic to explain the adoption of democratic institutions more generally. They see democratizations as a reaction of the elite to a temporary threat of violence by the majority. The uprising population knows that if it disarms, the government can repress it again, and therefore has an incentive to engage in violence. In this situation, institutional changes can provide a commitment device for the elite. This commitment can solve the dynamic power problem posited by Fearon and avoid violence. This implies that power-sharing agreements, perhaps with outside involvement, can help bring down violence if they provide a commitment device.

One source of variation in power-sharing agreements is the involvement of outsiders. Hörner et al. (2015) study the role of negotiation explicitly by applying the theory of mechanism design to the study of international conflict resolution. They show that, despite only being capable of making unenforceable recommendations, mediators can be effective as arbitrators. These encouraging findings contrast with recent work by Canidio and Esteban (2022) who show that conflict parties can have incentives to arm themselves more with mediation. As discussed in Blattman (2022), this already indicates that, depending on the underlying reason for conflict, we can be more or less optimistic about the role played by mediated negotiations.

[4] In limited access orders, the state does not have a secure monopoly on violence, and society organizes itself to control violence among the elite factions. (North et al, 2007).

In their studies of power sharing, Gates et al. (2016) and Strom et al. (2017) typify three categories of mechanisms through which power sharing works: 1) pooling of power: representatives of designated parties or groups hold particular offices or participate in particular decision-making processes, 2) dispersion of power: distribution of authority among groups or regions in a well-defined pattern, 3) constraining of power: limiting agent's power (a party or social group) to protect vulnerable groups, increasing the cost of repressing. According to their findings, the third category is most strongly associated with reductions in violence. Gates et al (2016) postulate that power-sharing institutions work best if they constrain governments from abusing less powerful groups and individuals, thereby solving the commitment problem. This provides less incentives for ordinary citizens to join potential insurgents, making conflict less likely.

There are variants of this argument. Besley and Persson (2011), for example, model the role of cohesive institutions as a constraint on rent extraction for the group in power. It is this commitment to an even distribution of public resources that reduces violence. The institutional commitment means that shifts in executive power do not shift the resource allocation, and this means that the incentives to fight for executive power are reduced. Strong institutions mean that the de jure and de facto power can fluctuate without triggering violence. Besley and Persson (2011) test their ideas using a measure of cohesive institutions and find that increasing cohesiveness indeed stops natural disasters or aid shocks from spilling into violence. Fetzer and Kyburz (2022) test this mechanism in the context of natural resource rents in Nigeria and find strong support. Regions that elect their local government are insulated from revenue shocks. Cheng et al (2018) also attribute institutions some degree of exogenous power over elite behavior, by determining the context in which they will make decisions. Elites play a role in shaping the pathway to conflict resolution given the patterns of development, the global/regional contexts, and pre-existing social structures. What is important here is how strong these institutions are. In fluid situations, like in Iraq or Afghanistan, elites will not feel bounded by dismantled or collapsing institutions. In other situations, as in Northern Ireland or Indonesia, where state institutions are more durable and some function in consistent ways regardless of who is in charge, elites will feel compelled to act inside the framework of these institutions.

There is a striking disconnect, however, between the academic literature on power sharing and the policy world. In policy circles the dispersion or constraining of power is not referred to as 'power sharing'. In diplomacy, 'power sharing' mainly refers to two or more conflict parties sharing executive power, e.g. in a "government of national unity" (similar to a coalition government). This view on power sharing is much more in line with the idea that de jure executive power needs to match de facto power in order to avoid violence. We will follow the academic literature by focusing on a data-driven definition based on "comprehensive" agreements defined in the next section.

3. Data description

This section describes the various datasets used in our quantitative analysis of power-sharing agreements. Over the past years, we have seen a significant improvement in the possibilities of examining this topic, due to the development of four datasets summarized in Appendix Table A1: fine-grained data on armed political violence, data on power-sharing agreements, data on political institutions, and forecasts of monthly conflict risk. We discuss these in turn.

Our goal is to conduct a study of monthly data for as many countries as possible reaching as long back in time as possible. As a result of this ambition, we restrict our analysis to a combination of the Uppsala Conflict Data Program (UCDP) (Sundberg and Melander 2013, Davies et al 2022) to measure armed political violence. We aggregate the Georeferenced Event Dataset (GED) at the country/month level summing over all types of fatalities and always take the *best* estimate. This gives us a dataset from 1989-2021 for over 170 countries. In some exercises, we normalize the violence data by population from the World Bank (2022).

We combine the resulting data with the PA-X Peace Agreement Database (Bell et al 2021) and Bell and Badanjak (2019) to capture power-sharing agreements. The PA-X dataset codes all peace agreements in the period between 1989 and 2020. A peace agreement is defined as a formal, publicly available document, produced after discussion with conflict protagonists and agreed to by some or all of them, addressing violence with an aim to end it.

Given the centrality of the PA-X data we discuss it in detail. The PA-X dataset codes different types of power sharing (political, territorial, economic, military). Power sharing refers to the specific divisions and amalgamations of power that ensure groups enjoy some form of equal 'participation' in the state's structures, and/or shared 'ownership' of resources. *Political* power sharing is defined using Lijphart's criteria, focusing on the establishment of, for instance, an executive grand coalition, the introduction of proportional representation in legislatures, mutual veto (or weighted majorities) in areas of groups' 'vital interests', and segmental (by concept, e.g. 'sport', 'education') autonomy. Given the specific interest of policymakers regarding this definition, we will analyze it separately from other definitions which tend to disperse power instead of sharing it. Territorial power sharing in PA-X is defined as divisions of power on a territorial basis. Economic power sharing is defined as joint participation in economic institutions, or territorial fiscal federalism. Military power sharing refers to provisions which share power in the institutions of the police, army, or security ministries.

In part of our empirical analysis, we do not distinguish between different types of power sharing, but we analyze them jointly. As a result, we have more than 440 power-sharing agreements spread across 68 countries in monthly data from 1989-2020 for over 170 countries - more than 70,000 country/month observations. We will put a particular focus on so-called *comprehensive* agreements which play a special role in changing conflict dynamics. These are defined by PA-X as agreements between parties that are engaged in an ongoing discussion, manage to agree on substantive issues in a comprehensive attempt to re-solve the respective conflict. When we focus on comprehensive agreements, we have 73 agreements in the data.

Figure 2 shows the composition of all and comprehensive power-sharing agreements along a small subset of the dimensions tracked by PA-X. The dotted, blue line shows the share of elements present in all power-sharing agreements. The solid, orange line shows the share present in comprehensive agreements. The main take-away is that comprehensive agreements have a lot more elements. The orange line runs outside the blue line on all categories. All comprehensive agreements mention the security sector, close to 90% human rights and equality, over 80% political power sharing, and close to 70% mention justice sector reforms. Justice sector reform is also a big outlier in terms of absolute increase in mentions from all to comprehensive agreements.

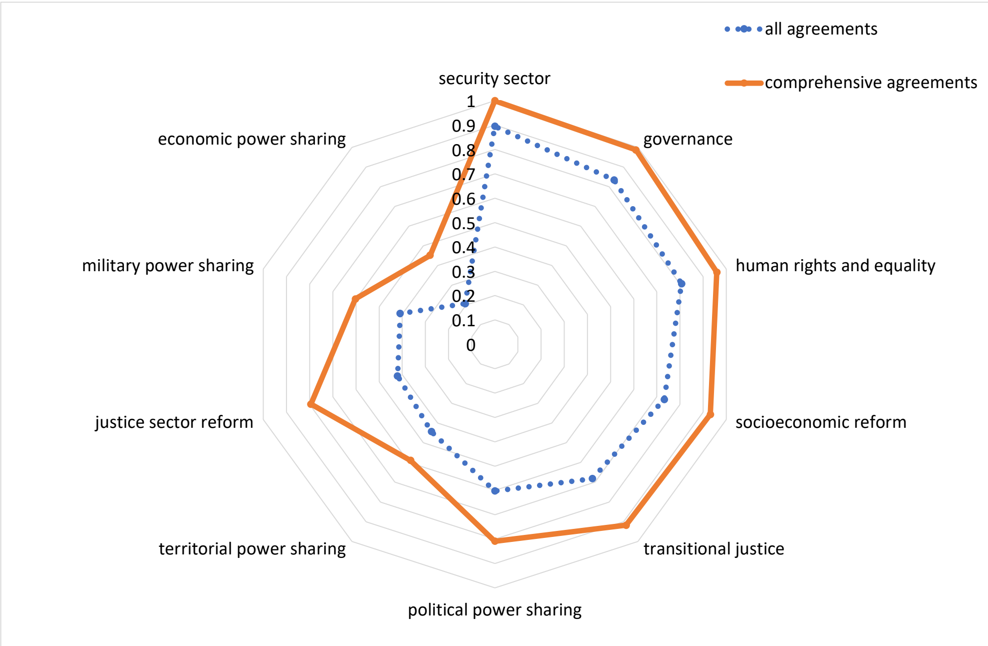

Figure 2: Comparison of all vs comprehensive power-sharing agreements

Notes: The figure displays a radar plot the components of all power sharing agreements (dotted line) and comprehensive agreements (solid line) as indicated by the V-Dem dataset.

To analyze long-term institutional changes, we add the Varieties of Democracy (V-Dem) data (Coppedge et al 2021) to capture political institutions. V-Dem is one of the standard datasets in the political science literature on political institutions and it tracks many aspects of these institutions for countries worldwide.

Finally, we use data generated using our methodology at conflictforecast.org (Mueller and Rauh 2021) to generate forecasts of future violence outbreaks and intensity at the monthly level for the period 2000-2020.[5] Our methodology uses news topics gained from summarizing around 5 million news articles using unsupervised machine learning combined with features that capture conflict dynamics in a forecasting framework to predict the occurrence and intensity of conflict. Conflict dynamics are captured the time since the last conflict and the intensity of recent and ongoing violence. The prediction algorithm we use is a random forest. This is a method which combines variables using many decision trees in order to discriminate between outcomes. The random forest trained on past data can then be used to forecast conflict in situations which are mostly driven by ongoing conflict dynamics as well as situations with subtle conflict risk captured by news stories (Mueller and Rauh 2022a,b). We forecast the outbreak of any violence in the next 12 months using a random forest classifier and the intensity of violence per capita in the next 12 months using a random forest regression. We use the same hyperparameters as our work for conflictforecast.org.

The forecast data we generate is based on rolling out-of-sample forecasts 12 months into the future. At each point in time T in the period 01/2000 to 12/2020 we take the information available up until time T and forecast violence 12 months into the future. We save these forecasts as the value at time T and then take one time step to T+1 and repeat. This procedure yields forecasts which are based on past information environments in the periods 01/2000 to 12/2020.

Throughout the article we use two sets of forecasts derived in this way. The first approach uses only past violence as predictor of future conflict. We use this to match treatment and control groups in our event studies as we explain in the following section. When we look at the relation between power-sharing agreements and the forecast error of our prediction models, we also include news text as a predictor. This forecast integrates discussions about economics, international diplomacy, and other topics. On the webpage these news topics are shown as bubbles on the respective country page. Reports on ongoing negotiations might be taken into account in this forecast.[6] This allows us to check whether power-sharing agreements are reflected in the forecasts relying on information sets before the adoption date.

4. Power sharing as an endogenous treatment

Identifying the impact of power-sharing agreements is complicated by the fact that these agreements are specifically targeted at addressing violence or situations with a lot of future potential for violence. In fact, there is a large literature in political science that discusses why

[5] We follow our method very closely. The only exception is that we need to use a single topic model (LDA) fit of the data to the data to be able to cover the entire period 2000 to 2020. Our original LDA model starts with a vintage of 2010.
[6] Figure A2 shows the news landscape leading up to a power-sharing agreement in Mozambique in August 2014, for example. Topics related to negotiations are clearly visible.

peace agreements occur and why they last.[7] In this section, we first provide an overview over this problem and then discuss our approach.

4.1. Overview

Power-sharing agreements select into situations with a high propensity to armed violence. Figure 3 shows the average propensity of adoption of all power-sharing agreements (top) and comprehensive agreements (bottom) from PA-X with increasing forecasted risk of an outbreak of violence (left) and the forecast intensity (right). To produce these figures, we bin all our observations in percentiles of forecast values – lowest to highest. We then show the mean value of power-sharing agreements in the three following months for each of these bins. Clearly, the adoption of a power-sharing agreement is strongly associated with violence risk. For example, in the top figure we show that in the three months following observations at the median of our outbreak risk and intensity forecast the likelihood of a power-sharing agreement was close to 0. In the three months following the highest intensity forecast, the likelihood of a power-sharing agreement is around 6% and following the highest outbreak risk it is 3%. For comprehensive agreements the pattern is very similar.

Importantly, our forecasts use a rolling forecast method which means that we use only the information available up until time T to produce an outbreak forecast for the period T+1 to T+12. In Figure 3 we furthermore show the forecasts derived from a model that only uses conflict dynamics to forecast conflict. This means that the risk percentiles used in Figure 3 are not directly affected by the power-sharing agreement. We will return to this point in the robustness checks.

[7] Fearon (2004) discussed five factors for the duration of civil wars empirically and theoretically. For an overview of bargaining problems like uncertainty and commitment problems see Walter (2009).

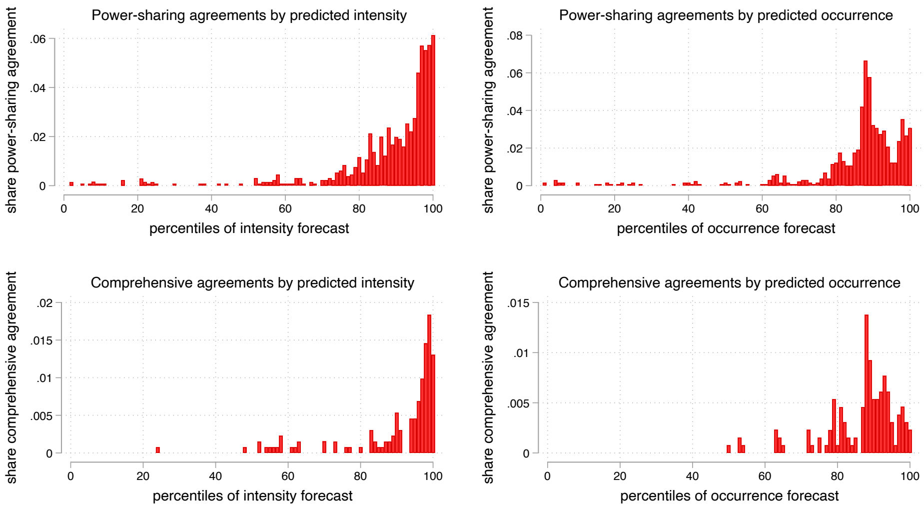

Figure 3: Likelihood of power-sharing agreement by conflict occurrence and intensity forecast percentiles

Notes: The panels show the share of country-months of all (top) and comprehensive (bottom) power-sharing agreements across percentiles of predicted intensity of violence (left) and likelihood of violence (right) in the three months leading up to an agreement.

The pattern shown in Figure 3 matches the narratives of mediation specialists we were in contact with at the German and UK government. Their actions are motivated, in part, by the prevention of future armed violence. This risk is evaluated by regional experts in the ministry headquarters in collaboration with staff in the local embassies. It is therefore entirely plausible that policymakers target situations with threatening violence dynamics, i.e. situations where future violence is most likely.

Why is this a problem for identifying the effect of power-sharing agreements? In Figure 4 we illustrate the typical context of peace agreements in directed acyclic graph (DAG). Circles indicate variables and arrows indicate causal relationships. In Figure 4 we are interested in identifying the marked arrow – the effect of peace agreements with power-sharing provisions and future violence. However, peace agreements are introduced as a reaction to a specific country context. Often this context is characterized by an active armed conflict which is itself driven by competition over resources or executive power. This competition will independently affect conflict risk, i.e. the risk of armed violence continuing or re-emerging, with or without an agreement in place. But because the peace agreement is, in part, itself a reaction to these factors, it becomes impossible to distinguish the effect of these problems and the effect of the peace agreement. In the jargon of causal inference, the backdoor criterion is violated.

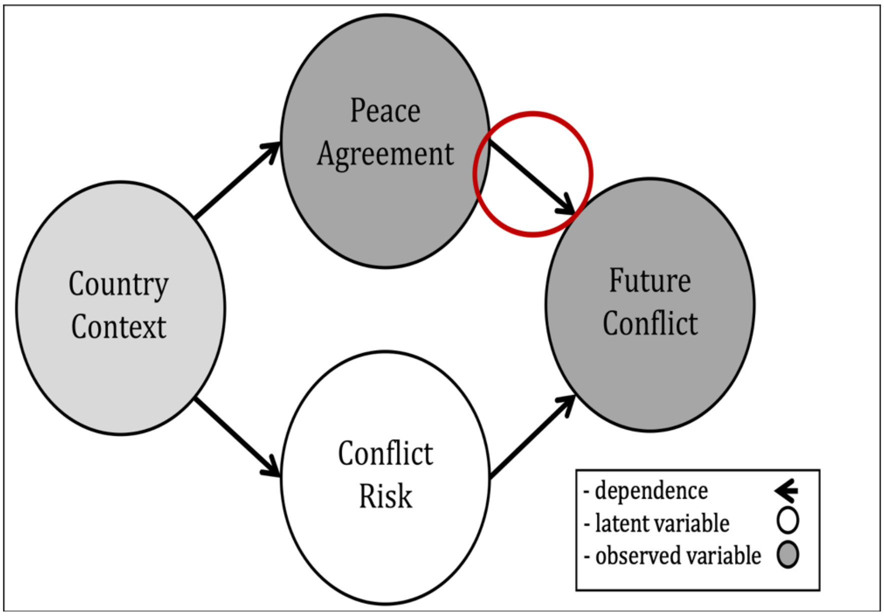

Figure 4: The identification problem

This violation imposes a potential bias of the effect of any study that tries to analyze the effect of power-sharing agreements and violence. If agreements work imperfectly, we will find that agreements are associated with increased violence compared to situations without peace agreements. Blaming peace agreements for violence is then akin to a situation in which a medical treatment to a severe illness is blamed for the following poor health. It is necessary to consider the conditions under which the treatment was administered.

4.2. Our approach

We combine a standard event study approach with an attempt to measure conflict risk to get around this problem. Our method compares monthly violence data in event study windows before and after the adoption of power-sharing agreements. This controls for factors at the country level that stay constant during this relatively short time window – in our case 6, 12, and 18 months before and after adoption.

However, the availability of an archive of rolling forecasts described above also allows us to control for conflict risk. The idea here is to match the conflict risk at the time of adoption in the truly treated event windows with other countries in similar situations with comparable conflict risk but no adoption of power sharing in the following months. This closes the backchannel shown in Figure 4 as we are making conflict risk observable and then construct a control group with the same characteristics.

This method is most similar in spirit to a synthetic control method in which the weights are constructed such that the resulting control group matches the treated group as closely as possible before treatment (Abadie 2021). Our idea is that conflict forecasts are an extremely useful way to construct these weights as they are meant to capture the outlook of the dependent variable. As a result, they explicitly match the treated units to a control group with a similar violence outlook in the absence of treatment. Note, that we match on the average forecast 1 to 3 months before treatment. The fact that we find no significant treatment effects in all periods before treatment, even the unmatched ones, speaks to the fact that this is a remarkably effective way to construct a control group. What is special about our way of generating the synthetic control is that the matching method is explicitly constructed using the violence forecast, i.e. the expected future realizations of the dependent variable.[8] The effect we identify this way needs to come from the treatment effect being unanticipated. We return to this point in Section 5.3 by looking at forecast errors explicitly.

Another version of the matching we conduct as a robustness check is to forecast the adoption of a power-sharing agreement and match on this prediction instead. This way of matching tries to explicitly compensate for the endogenous placement of agreements and, instead of finding a good control for conflict risk, tries to develop a control group that matches in the propensity of receiving the treatment. Our matching method allows us to look at dynamic effects for units matched on the likelihood of receiving the treatment. This type of matching will change the estimate if the presence of a power-sharing agreement signals a special circumstance that has a dynamic effect on conflict, which begins after the treatment and cannot be detected through pre-treatment matching on conflict risk.

Note, that our approach has limitations which need to be taken into account when interpreting the results. It is likely that other policies, like mediation, foreign aid, or external security controls, are implemented to support the peace agreement we study. If these other policies have an effect and their timing coincides with the month of the power-sharing agreement, then our method will capture the overall effect of these policies. The matching on the likelihood of

[8] Our method also has some connection to double ML methods like Chernozhukov et al (2018) in this regard but allows us to observe pre-trends and dynamic effects through the combination of matching and difference-in-difference.

receiving a power-sharing agreement should take care of some of this effect. However, given that our intention is to study the overall effect of peace agreements, our results should be regarded as the evaluation of a policy instruments with its supporting policies. In robustness checks, we control for one measure that is often introduced coincidently with power-sharing agreements: peacekeeping.

An alternative approach would be to try and find exogenous variation in the policy instrument. Such an approach has, for example, been implemented to study the effect of foreign food aid on armed conflict by Nunn and Qian (2014). The problem with this approach, is that exogenous variation in the policy instrument means that the policy is not endogenous, i.e. it is not demand-driven. But foreign policies which are not driven by local requirements and a demand for intervention by local actors might not be the most effective type of foreign interventions. Estimates are then causally identified but the treatment is a very specific one so that results do not generalize.

In light of these challenges, we implemented a method which combines difference-in-difference estimates in time-windows around the adoption of a power-sharing agreements with a matching of treated time/country windows with comparable non-treated time/country windows. Before we describe the method in Section 5.2. we discuss some case studies.

5. Results

5.1. Case studies

We first explore the link between violence and power sharing along the lines of three case studies. These case studies are selected in a completely subjective way and serve to illustrate the aspects discussed above and to motivate our empirical estimation strategy. We will show average, and therefore generalizable, effects in the next section.

Our first case study, shown in the top panel of Figure 5, is Mozambique. The country experienced high levels of political violence in the beginning of the sample with around 150 fatalities per month (5 on the log scale). These levels of violence were the result of the Frelimo-Renamo conflict, which lasted fifteen years (1977-1992) and ended with the 1992 Rome peace agreement. The comprehensive agreement led to the establishment of multiparty elections in 1994. We mark the agreement with a red line - the dramatic drop in violence after the agreement is clearly visible.

Frelimo has won every election since, amidst widespread allegations of fraud and suppression of the opposition. Renamo has maintained an armed guerrilla force, and violence has occasionally erupted between them and the government, such as in 2013 and 2016, although it has never reached the level of intensity previously seen. In reaction to a particularly bloody outbreak in 2014, a new peace agreement was signed. Again, we see a decrease in violence following this agreement. The recent violence in the north of Mozambique involves a violent extremist group (IS) who was not part of the initial power sharing deal.

The middle panel of Figure 5 demonstrates the case of Angola. Here the effect of power sharing is less clear with violence levels being higher despite a cycle of seven consecutive power-sharing agreements, two of them comprehensive. In several cases, the number of fatalities decreased dramatically after the adoption of an agreement. For example, in 1991 the UNITA and the MPLA government signed the Bicesse Peace Agreement. The agreement provided for the establishment of a multi-party system, which allowed presidential elections to be held the following year. In the aftermath there was a brief episode of peace, but violence broke out again in 1992. A slightly longer stabilization can be observed after the comprehensive agreement in 1994, however, violence is again only de-escalating for a few years.

The case of Angola illustrates clearly how local and external actors use power sharing repeatedly to decrease violence, with actors re-negotiating the distribution of power repeatedly. Sometimes this seems to reduce violence temporarily and at other times there is no effect, which means the content of the power-sharing deal did not provide sufficient incentives to cease violent competition. The case of Angola also illustrates the feedback of violence to agreements well: Once violence recedes, so does the frequency of power-sharing attempts.

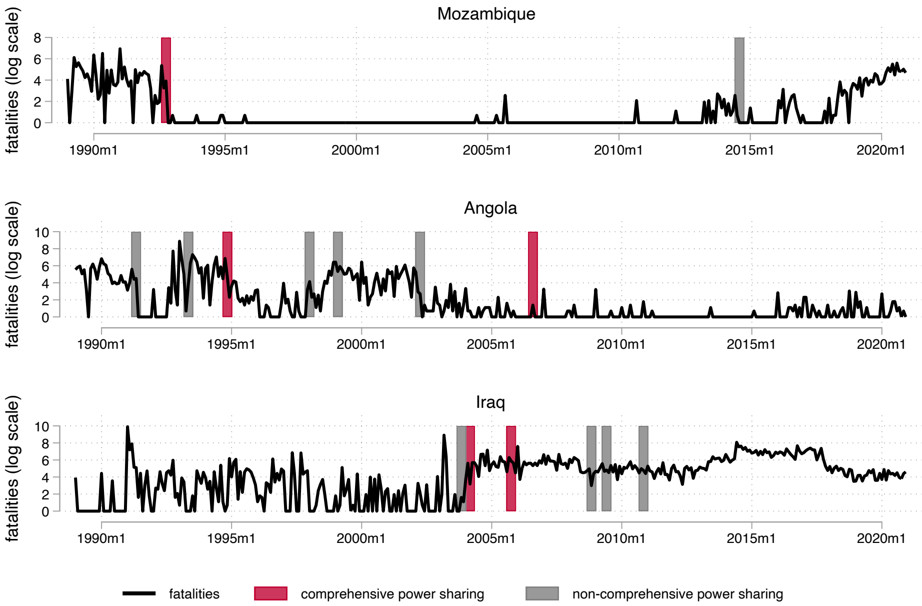

Figure 5: Power sharing case studies

The history of Angola is therefore aligned with the theoretical model proposed in Figure 4 and the selection analysis shown in Figure 3. Power-sharing agreements are a reaction to a dire situation. The aftermath of power sharing is, therefore, on average, still characterized by violence, albeit at a lower level. What is important to note, however, is that often, the months following directly after a power-sharing agreement has been agreed, are less violent than the months preceding it. We will return to testing this proposition statistically in the next section.

It is important to bear in mind that amongst recorded power-sharing agreements there are quite a few that seem to have had no or even a negative effect. The bottom panel of Figure 5 shows the case of Iraq where the first power-sharing agreement was concluded after the US-led Iraq invasion. The agreement preceded a dramatic escalation in violence which the following agreements could not appease. These agreements could also not prevent an insurgency by the Sunni tribes, who were not part of the deal, and who later formed the Islamic State, pushing violent deaths to unprecedented level. We will return to this example, but it should be kept in mind that what we document is the average quantitative effect of peace agreements which includes failures like Iraq.

5.2. Matched difference-in-difference method

The staggered and repeated treatment of power-sharing agreements shown in Figure 5 poses a challenge to modern methods of difference-in-difference estimates for the correct identification of the Average Treatment Effects on Treated (ATT) (Callaway and Sant'Anna 2021).[9] In the standard adoption of the method we would have pre-treatment months of one power-sharing agreement coincide with the post-treatment months of another agreement.
In addition, we have a selection problem in which the adoption of an agreement takes place in circumstances in which violence is ongoing, has escalated, or is in danger of escalating. This means that country fixed effects alone do not help controlling for risk.

As discussed in detail in Section 4.2 we, therefore, combine the difference-in-difference method with a matching method which approximates synthetic controls and, as a robustness check, employ propensity score matching around pre-defined treatment windows. We focus our analysis on the months before and after the adoption of power-sharing agreements and do our best to build a good control group of country windows from the remaining, untreated data. In a first step, we select windows around the adoption of power-sharing agreements while ensuring that these are non-overlapping. We take a window range between 6, 12 and 18 months centered around the adoption month which we call month 0.[10]

We then sample from the remaining untreated data through a sampling method to construct a control group. We start by constructing a control group that is sampled using the distribution of violence intensity forecasts. To understand this sampling method, we direct the reader back to Figure 3. On the top left of Figure 3 we show the adoption likelihood by conflict intensity percentile. We can use these distributions to sample placebo treatments in the untreated data. Specifically, we use the distribution shown on the top left of Figure 3 to draw random treatments of power-sharing agreements across our entire dataset with the same likelihood shown in the figure.[11] Most of the percentiles have a likelihood of 0 and therefore do not receive a placebo treatment. This ensures that countries like Sweden or Germany are not part of our control group. Even for the percentiles with the highest violence forecast the adoption likelihood is never higher than 6% which means our placebo treatments are attributed in a very sparse way across the dataset. However, the distribution of conflict intensity forecasts in the resulting control group will approach the distribution of the actually treated units as can be seen for an example in Appendix Figure A3.

After we attributed random placebo treatments, we ensure that the windows around these treatments are intact so that we can track violence in the entire window before date 0 and after

[9] We use the csdid package for STATA developed by Fernando Rios-Avila (Rios-Avila et al 2021) based on Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). We implement the doubly robust DiD estimator based on stabilized inverse probability weighting and ordinary least squares.
[10] To increase sample size, we accept windows that have 10% of their observations missing, i.e. this is between 1 and 3 observations of a given window depending on the window size.
[11] Specifically, we draw placebo treatments 2 months after these observations as Figure 3 shows the average adoption likelihood 1-3 months before adoption.

date 0. This means we can now compare a treated group with some distribution of conflict risk before the adoption date 0 with a control group that has the same distribution of conflict risk in the months before treatment at 0. The only difference is that the treatment group indeed receives a treatment at 0 whereas the control group does not.

Having constructed our estimation sample of country windows *i* in group *g* in month *t*, we estimate the impact of power-sharing agreements on outcome *y* using the regression equation

$$y_{igt} = \alpha_i + \gamma_t + \beta_k \times \sum_{t=-(w-1)}^{w} D_{kgt} + x_{it} + \epsilon_{igt}$$

where *w* is the window size of either 6, 12, or 18 months, x_{it} are control variables, and D_{kgt} is a set of indicators that take the value one if, for country *i* in month *t*, the introduction of a power-sharing agreement was *k* months away. For the untreated countries $D_{kgt} = 0$ throughout. In all specifications we include calendar month dummies to control for seasonality. Further, in some specifications we also control for the presence of peacekeeping troops.

As discussed in a wide range of papers (e.g., De Chaisemartin and d'Haultfoeuille 2020, Callaway and Sant'Anna 2021, Borusyak et al 2022), the two-way-fixed effect estimation of the model provides biased estimates if a stringent set of assumptions does not hold. We, therefore, rely on the estimator developed by Callaway and Sant'Anna (2021) which aggregates the ATT by averaging the β_k coefficients for $k \geq 0$ using inverse probability weighting to compute a group-time average treatment effect. In our main specifications we look at the effect of agreements on a binary indicator for the occurrence of any battle death and log of the number of battle deaths + 1 as outcomes *y*. In other words, we estimate the effect on the probability of any violence and the intensity of violence. Later we also study the impact on peacekeeping and institutional features.

A visual representation of the estimated β coefficients is shown in Figure 6 for a window of size of $w=12$ months and a control group matched through the intensity forecast percentiles in months -3 to -1. The default omitted category is the first time period, that is $-w$ if the agreement is introduced at 0. The first observation is that parallel pre-trends appear to hold as the estimates before the introduction of the agreement at $t=0$ fluctuate around zero and remain statistically insignificant. There is a clear pattern in which point estimates for violence fall immediately the month after the power-sharing agreement. For violence occurrence the treatment effect here is around 10 percentage points whereas for intensity it is around 0.4 log points or 30%. The point estimates fall one month after the adoption data and stay relatively stable, without a noticeable rebound back to baseline. The overall effects are statistically significant at 10% confidence.

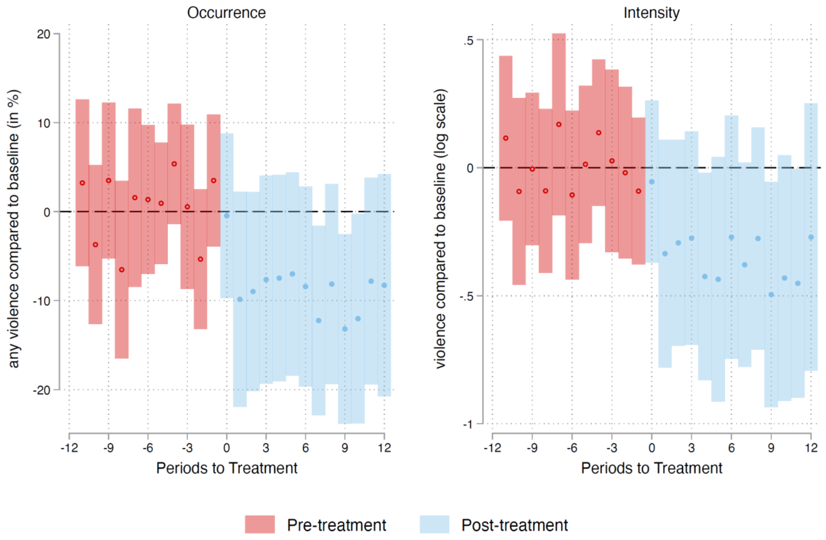

Figure 6: Violence occurrence and intensity after power sharing

Notes: This figure shows the estimated impact of all power-sharing agreements on the occurrence (left) and the intensity of violence (right) in the twelve months before and after the agreement. The control group is assigned placebo agreements based on the distribution of predicted intensity of violence in the three months leading up to agreements.

We find stronger effects for comprehensive agreements which we show for 12 months windows in Figure 7. Again, we build the control group using the intensity forecast percentiles in the months -3 to -1 leading up to the agreement. We find that violence falls one month after the adoption of a comprehensive agreement. However, violence dynamics in the point estimates now trend slowly downwards. For violence occurrence we now find treatment effects that approach a 20-percentage point reduction and for intensity a reduction of 100 log points, which is a fall in the intensity of violence of close to 60%. The ATT is statistically significant at 1%.

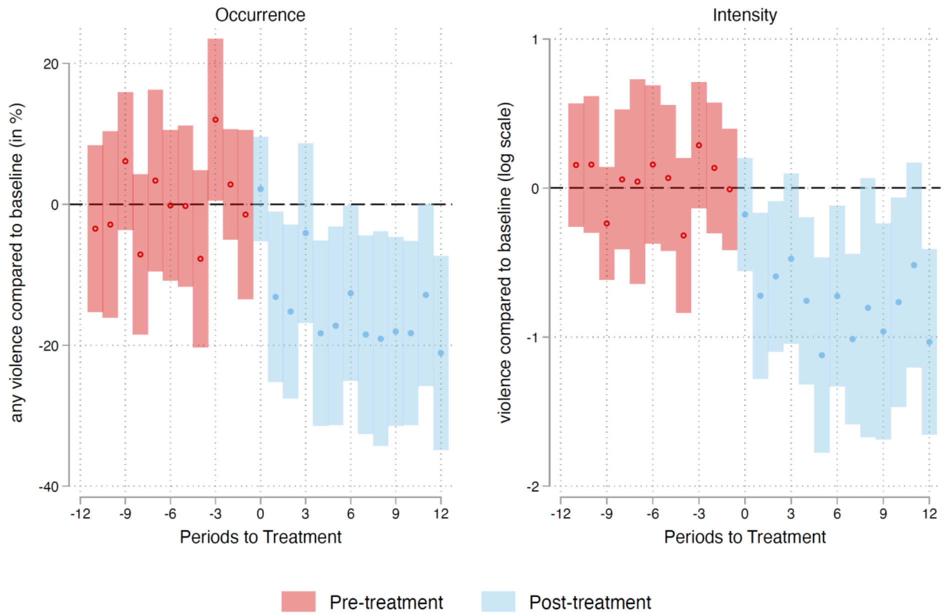

Figure 7: Violence occurrence and intensity after comprehensive power sharing

Notes: This figure shows the estimated impact of comprehensive power-sharing agreements on the occurrence (left) and the intensity of violence (right) in the twelve months before and after the agreement. The control group is assigned placebo agreements based on the distribution of predicted intensity of violence in the three months leading up to agreements.

In Table 1 we summarize our treatment effects for varying window sizes. Results suggest a stronger, more statistically significant effect with longer time windows both for all power-sharing agreements and for comprehensive agreements. The point estimate lies between 2 and 12 percentage point reductions in occurrence for all agreements and 7 to 14 percentage points for comprehensive agreements. Results are only weakly statistically significant for all agreements, but are more precisely estimated for comprehensive agreements.

These results and visual inspection of the monthly ATT point estimates in Figures 6 and 7 suggest that effects strengthen over time. We confirm this using the equivalent approach with 18-month windows in Appendix Figures A4 and A5. The point estimates for all power-sharing agreements and comprehensive agreements show a downward trend. After 18 months the point estimates for all power-sharing agreements suggest close to 18 percentage points reductions in occurrence and 45% reduction in violence intensity. For comprehensive agreements the effects are a 20-percentage point reduction in occurrence and a 70% reduction in violence intensity. However, due to the endogenous survival of agreements, we later also estimate a version of the model that includes all agreements, whether they feature follow-up agreements or not.

These results suggest that there is a fast response of violence to the adoption of power-sharing agreements. The point estimates we find towards the end of the event windows suggest very large treatment effects and the monthly estimates show a negative treatment effect setting in at the adoption date. We will demonstrate a remarkable robustness of these results in Section 5.4.

Table 1: The ATT of power-sharing agreements on violence occurrence and intensity

Panel A: All power-sharing agreements

	(1)	(2)	(3)	(4)	(5)	(6)
	6 months window		12 months window		18 months window	
	any	intensity	any	intensity	any	intensity
ATT power sharing	-2.323	-0.171	-8.984*	-0.373**	-12.201*	-0.445*
	(2.890)	(0.121)	(4.839)	(0.187)	(6.467)	(0.250)
N	231824	231824	361057	361057	459141	459141
treated	135	135	71	71	47	47

Panel B: Comprehensive power-sharing agreements

	(1)	(2)	(3)	(4)	(5)	(6)
	6 months window		12 months window		18 months window	
	any	intensity	any	intensity	any	intensity
ATT power sharing	-7.248*	-0.428**	-14.202***	-0.745***	-14.056***	-0.875***
	(3.841)	(0.198)	(4.657)	(0.252)	(4.804)	(0.267)
N	57239	57239	100397	100397	137764	137764
treated	59	59	47	47	44	44

Robust standard errors, clustered at the country-window level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Any violence (columns 'any') is equal to 100 if there is any fatality according to UCDP in that month and zero otherwise. Violence intensity (columns 'intensity') are log(fatalities +1). All regressions restrict the sample to a window around the adoption of power-sharing agreements for treated countries and control for month fixed effects. The control group is a random sample of countries without power-sharing agreements but with the same distribution of predicted conflict intensity before adoption.

5.3. Difference-in-difference results looking at forecasting errors

Our forecasts are based on a complex way of capturing conflict dynamics using machine learning which is geared to get the best possible estimate of future violence by including all possible information available at the time of the forecast. The full model also uses the information contained in millions of news articles capturing the political, economic and international context of each country. This reliance on news in the full model allows us to capture the information environment leading up to power-sharing agreements. In what follows we will use our conflict forecast to see whether the adoption of power-sharing agreements can be anticipated. We can use the forecast to see whether, viewed from an information environment before adoption, the adoption of power sharing leads to positive surprises in which the forecast becomes overly pessimistic right before the agreement.

We forecast the occurrence of violence 12 months into the future. This means that we can look at forecast errors at points in time at which the treatment was not yet implemented. As a first step, we now shift the definition of *treatment* so that time 0 is the moment a power-sharing agreement will be adopted in 12 months. We then ask, given the information available at time T, was the forecast for the next 12 months too optimistic or too pessimistic and how does this evolve as the adoption date approaches?

The forecast error here is defined as the true realization minus the forecast at time T – both measured as averages in the 12 consecutive months. Positive errors mean that, at the time of the forecast, we are too optimistic and underestimate future violence compared to the control group.[12] Negative forecast errors mean that we are too pessimistic in our forecast. In our analysis we let T go from -6 to +12 months. At month 0 the future adoption of a power-sharing agreement enters the forecast horizon. Importantly, the forecast in months 0 to 12 then spans a mix of treated and untreated months. Only in months 11 and 12 are the entire forecasting horizons inside the treated time interval because the prediction at 0 is 12 months into the future.

This means we expect a drop of forecasting errors in the periods 0 to 12 if power-sharing agreements are not anticipated. Anticipated adoptions would mean that future violence reductions due to power-sharing agreements are already "priced in" at time 0 to 12. If the effect of agreements were priced in at time 11, for example, we would not see a negative error when compared to the control group. But if agreements (or their effect) are a surprise, we expect the forecast error to become negative as the adoption date approaches.

We show two difference-in-difference estimates for all power-sharing agreements in Figure 8. In the left panel we show the results using errors for violence outbreaks before agreements and on the right for the errors for the predicted intensity of violence. As discussed above, the

[12] We now need to sample the control group from the entire a sample as matching by conflict risk would lead to selection problems.

treatment now refers to the entering of agreement into the *forecast horizon* at point 0, the actual adoption of the agreement happens at time 12. In the point estimates we see clear patterns in which there is a downward trend in the forecast error indicating a forecast that is becoming too pessimistic when using the information available before the adoption of an agreement. Our forecast model overpredicts violence. Regarding intensity, there is a slight increase in the error point estimates in months 2 to 6, followed by a clear decline closer to the agreement date. This is consistent with the idea that power-sharing agreements target situations with vicious conflict dynamics captured by an upward drift of the forecast error of our forecasting system. Agreements are then associated with a dramatic and systematic trend reversal in the forecasting error, i.e. even taking all possible information into account that is available in the months leading up to the agreement. Power-sharing agreements and their effect are positive surprises for our forecasting system even though it is able to capture the news environment.[13] Appendix Figure A6 demonstrates that the same patterns hold for comprehensive agreements.

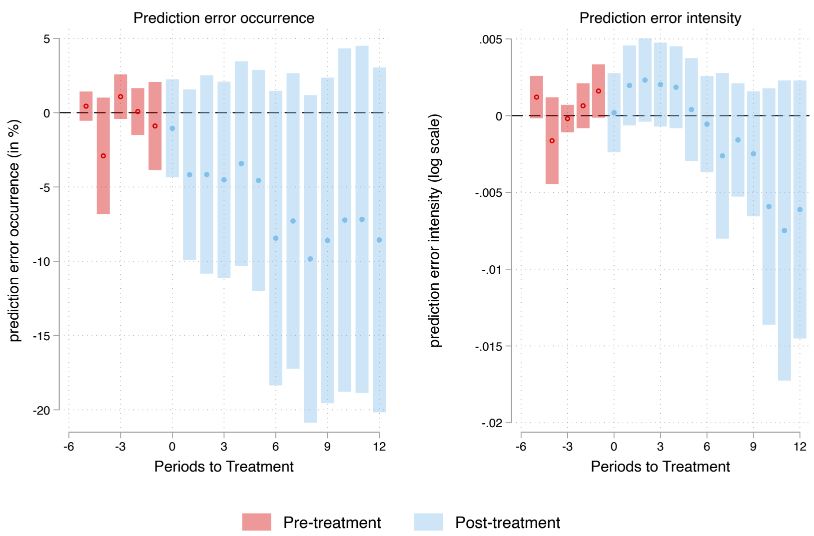

Figure 8: Forecast error for all power-sharing agreements

Notes: This figure shows the estimated impact of all power-sharing agreements on the forecast error of the occurrence (left) and the intensity of violence (right). In this figure the agreement takes place at month 12. The forecast is for the next twelve months so the impact of the agreement enters the forecast window at month 0. A negative forecast error indicates an overprediction of violence. The control group is random.

[13] We have also analyzed the forecasting errors after the adoption dates. This exercise suggests that having the PAX data available when conducting the forecast could bring down forecast errors. The system stays too pessimistic even after the adoption date and only slowly adjusts.

These results confirm the assumption that the adoption of power-sharing agreements can be modelled as a positive surprise which was hard to anticipate with the information set available in the months before the adoption of the agreement. The larger share of the forecast window lies behind the adoption date, the more pessimistic is the forecast. Interestingly, the point estimate for occurrence suggests a substantial error, i.e. around 10 percentage points for both all and comprehensive agreements.[14]

5.4. Experimentation with the matching method

An important aspect of our identification strategy is the selection of the placebo control group. We therefore show three alternative ways of attributing placebo treatments to untreated country windows: 1) a completely random control group; and 2) a control group that is sampled using the same distribution of the predicted likelihood of the occurrence rather than the intensity of violence; 3) a control group that is selected by the likelihood of getting treated by a power-sharing agreement.

Results for the first two matching methods are shown in Appendix Table A2 columns (1) to (4). We find very similar or somewhat stronger results when sampling by conflict outbreak risk (Appendix Figure A7) and find weaker results when sampling randomly (Appendix Figure A8). This is consistent the idea that the targeting of escalating situations by efforts of the international community leads to a downward bias in the estimated treatment effects unless this escalation risk is explicitly taken into account through the matching on risk.

Finally, we match by the likelihood of an adoption of a power-sharing agreement. For this we forecast the likelihood of a power-sharing agreement through a model similar to our conflict forecast the months since the last agreement and last comprehensive agreement as additional predictors.[15] We find that, despite having a decent forecast model with an area under the receiver-operating curve (ROC-AUC) of 0.9, which is a common measure of the tradeoff between false-positive and the true-positive rates, the precision of these forecasts is relatively low. It is hard to predict power-sharing agreements, simply because they are so rare. Still, this way of matching explicitly compensates for the endogenous placement of agreements and, instead of finding a good control for conflict risk, tries to develop a control group that matches in terms of the propensity of receiving the treatment. Results are shown in Appendix Table A2 columns (5) and (6) and Appendix Figure A9. Again, results are robust but are now more similar in their point estimate to the random matching.

[14] We forecast violence per capita for intensity which yields a very different dimensionality here than in the main results using log fatalities.
[15] We use standard cross-validation instead of rolling forecasts. See Appendix A for a more detailed discussion.

5.5. Robustness and additional results

A difficult concern to address is reverse causality. Violence could be trending upwards before the adoption of power-sharing agreements because of the ongoing negotiations. Mediation attempts can increase violence because they increase the incentives to engage in violence to strengthen bargaining power (Canidio and Esteban 2022). This would make the months right before an agreement a bad control group and this would mean that our method of matching on risk in the three months before the agreement leads to an overestimate of the true effect. Appendix Figure A10 investigates this possibility by matching the control group using the months 4 to 6 before the adoption month. Results are robust to this. Note also that the estimation method we use would allow us to track deviations of violence in the treated group before the treatment but there is little evidence for this in Figures 6 and 7.

An additional channel of reverse causality is that increasing violence triggers a new agreement so that our sample is increasingly selected if we select intact event windows of lengths 12 and 18 months. We therefore run an alternative way of defining windows which ignores repeated treatments and simply cuts windows around all adoption dates. In other words, we do not select cases based on their duration. The control group is sampled as before. Note, that the sample is now selected negatively as we have repeated coverage of cases with repeated treatments. Treatments that do not work and are followed by another surge or ongoing violence then lead to another case which is also included in the sample.

Figures 8 and 9 show the results. There is significant decrease in violence following all agreements and comprehensive agreements. Importantly, the effect of agreements is again growing over time which means this pattern is not due to selection effects. The size of the point estimate is now somewhat lower for occurrence but very similar in size for intensity.

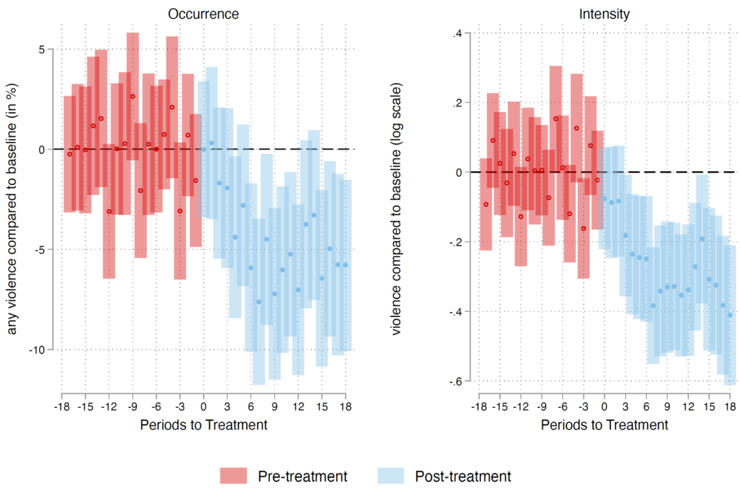

Figure 9: Violence occurrence and intensity after all power sharing (whether agreement is replaced or not)

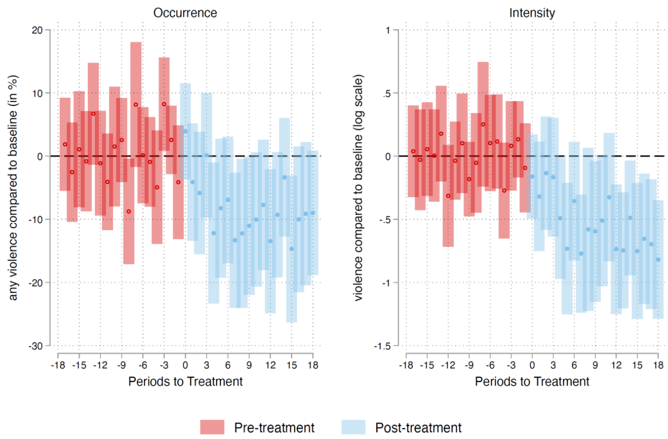

Figure 10: Violence occurrence and intensity after comprehensive power sharing (whether agreement is replaced or not)

Notes: These figures show the impact of power-sharing agreements on the occurrence (left) and the intensity of violence (right) in the 18 months before and after the agreement. The control group is assigned placebo agreements based on the distribution of predicted intensity of violence in the three months prior to agreements. The sample includes all agreements at month 0, whether they are replaced in the following months or not.

One remaining issue in interpreting the findings in the previous section as causal treatment effects of power sharing is that power-sharing agreements are accompanied by other policies, often put in place by the international community, which also contribute to the reduction in violence and coincide with the implementation of a power-sharing agreement. This would lead to an omitted variable bias in as far as these other policies are not part of the comprehensive peace agreement but are additional measures that happen to coincide with the agreement.

A policy which is closely linked to international attempts of pacification is peacekeeping missions. We therefore use two datasets from the UN webpages. First, the number of total peacekeeping troops present in a country and, second, the monthly budget spent on peacekeeping in a country. In the latter case we had to interpolate between quarterly or even yearly reports such that the timing is not precisely measured. However, presence of troops is relatively well recorded with even a handful of peacekeepers being tracked.

Figure 10 shows that peacekeeping is very clearly a policy that coincides with power-sharing agreements. Peacekeeping troops (left panel), peacekeeping budgets (middle panel) and a dummy indicating troop presence (right panel) shoot up in the aftermath of a comprehensive agreement (and slightly less in Appendix Figure A11 for all agreements). However, we find only weak evidence of an independent effect of peacekeeping on violence. Appendix Figure A12 reports the difference-in-difference estimate of the impact when peacekeeping missions are measured by the log of the number of troops + 1. This definition exhibits the strongest, yet often not significant, relationship in terms of reductions of violence.

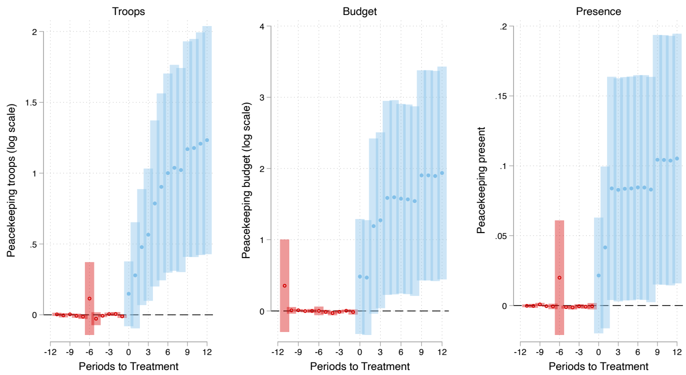

Figure 11: Peacekeeping activity after comprehensive power-sharing agreements

Notes: This figure shows the estimated impact of comprehensive power-sharing agreements on the log of peacekeeping troops + 1 (left), log of peacekeeping budget + 1 (middle), and presence of peacekeeping troops (right) in the twelve months before and after the agreement. The control group is assigned placebo agreements based on the distribution of predicted intensity of violence in the three months leading up to agreements.

Not surprisingly, our main results are robust to peacekeeping controls which we show in Table 2. Here we focus on comprehensive agreements and fix the window size to 12 months before and after the signature of an agreement, while controlling for peacekeeping activity in three different ways. The main finding is that the estimated ATT of comprehensive power-sharing agreements remains unchanged when compared to columns (3) and (4) in Table 1, Panel B. If anything, these results suggest that the studies of peacekeeping, especially work using yearly variation, need to pay attention to the context of political agreements that prepare the ground for peacekeepers. Further work using subnational data might help provide better identification of the independent effects of peacekeeping.

As already discussed, agreements are often followed by other agreements. We, therefore, estimate the effect of agreements that come into existence separately for countries that have had no agreement within the past two years and for those that have had an agreement within the past two years. In Appendix Figure A13 one can see for violence occurrence that the drop in the likelihood of violence is greater and more significant for countries that have had previous agreements. The same pattern emerges for the intensity of violence in Appendix Figure A14. This finding is consistent with the idea that a sequence of power-sharing agreements is leading up to a comprehensive agreement. The passing of an agreement out of nowhere might not be sufficient to halt violence, but it can form a steppingstone for further agreements and reductions in violence.

Table 2: The ATT of comprehensive power-sharing agreements on violence occurrence and intensity controlling for peacekeeping

	(1) any	(2) intensity	(3) any	(4) intensity	(5) any	(6) intensity
ATT power sharing	-14.189***	-0.762***	-14.119***	-0.766***	-14.198***	-0.768***
	(4.680)	(0.253)	(4.669)	(0.253)	(4.663)	(0.253)
Ln(troops+1)	✓	✓				
Ln(budget+1)			✓	✓		
Any peacekeeping					✓	✓
N	100397	100397	100397	100397	100397	100397
treated	47	47	47	47	47	47

Robust standard errors, clustered at the country-window level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Any violence is equal to 100 if there is any fatality according to UCDP in that month and zero otherwise. Violence intensity are log(fatalities +1). All regressions restrict the sample to a window of 12 months around the adoption of comprehensive power-sharing agreements for treated countries and control for month fixed effects. The control group is a random sample of countries without power-sharing agreements but with the same distribution of predicted conflict intensity before adoption. Columns (1)-(2) control for peacekeeping troops using ln(troops+1), columns (3)-(4) for the peacekeeping budget using ln(budget+1), and columns (5)-(6) using a dummy for whether peacekeeping troops are present.

Finally, we investigate the differential effects of the ten different components of agreements illustrated in Figure 2 by running one separate specification for each component. The average treatment effects across the 18-month windows are displayed in Figure 12 for all agreements, i.e. whether they endure or not. For agreements with clean treatment windows, i.e. those that are not superseded by other agreements during the treatment window, coefficients are displayed in Appendix Figure A15. The estimated treatment effects appear to be greatest for agreements featuring elements of military power sharing, governance, transitional justice, human rights and equality, and economic power sharing. Many of these features are rarely included in non-comprehensive agreements and might, therefore, be some of the key elements distinguishing comprehensive and non-comprehensive agreements. These patterns highlight two findings: first, involving military power-sharing could be a critical element and, second, elements that deal with the vertical distribution of power between elite and population, like governance and human rights and equality, also play a critical role. We will return to this point in the discussion section.

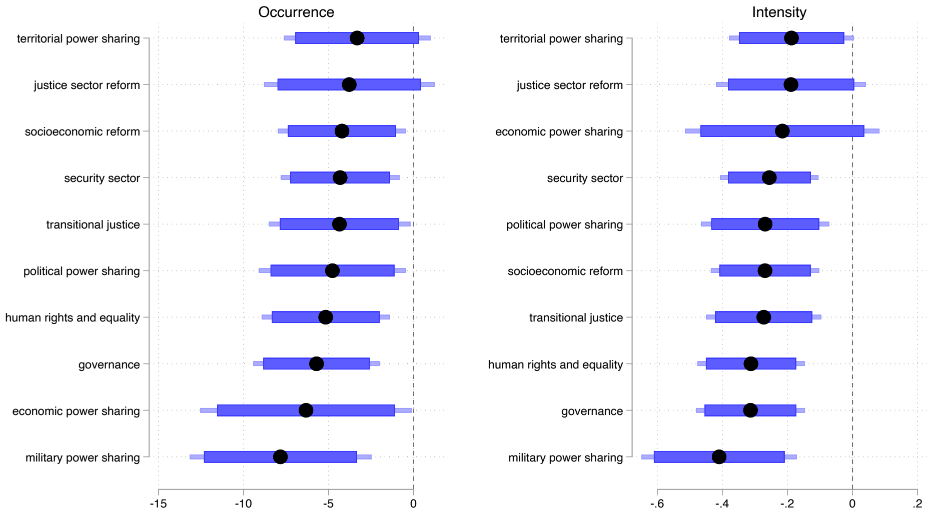

Figure 12: ATT for different components of agreements

Notes: This figure shows average treatment effects from separate difference-in-difference estimations for each of the ten components of power-sharing agreements in an 18-month window including all agreements, whether they were replaced within the event window or not. The thin bars represent the 95% and the thick bars the 90% confidence intervals.

6. Building bridges: The aftermath of power-sharing agreements

In this section we explore additional data on institutions to shed light on the broader institutional implications of power-sharing agreements. We will pay special attention to the role of comprehensive agreements. This will provide hints as to why they reduce political violence in a more robust way.

6.1. The duration of power-sharing agreements

We have already seen in the case studies, power-sharing agreements are no panacea. They are deployed repeatedly, sometimes within months. Even comprehensive agreements can be tried more than once in the history of a country and so the duration of an agreement will be endogenous to the country context. A complete analysis of agreement durations lies beyond the scope of this study.[16] Yet, we need to pay attention the mechanisms underlying heterogeneity in duration in the context of the identification problem shown in Figure 3.

[16] For a study on the duration of peace agreements after international wars see Fortna (2003).

In Figure 13 we show the share of comprehensive agreements which are replaced by new comprehensive agreements within the first 18 months. We split the sample into comprehensive agreements at the median in terms of predicted risk in the three months leading up to the agreement. We see that after 6 months 10% of the agreements have been replaced by a new comprehensive agreement, independent of the pre-agreement level of risk, suggesting that comprehensive agreements can also survive for enough time to reduce violence even when drawn up in high-risk situations.

Then the lines start to diverge and more than one third of the agreements introduced in high-risk situations are replaced within 18 months compared to less than 1 in 5 introduced in lower risk situations. This is an important finding as it suggests a selection problem in which more difficult situations lead to more violence and lower duration. We already showed in the previous robustness section that the violence-reducing effects hold when looking at the impact of agreements unconditional on duration. This means that even as we mix the effects of agreements which remain in place with the effect of new agreements, we get similar overall results as in our main findings.

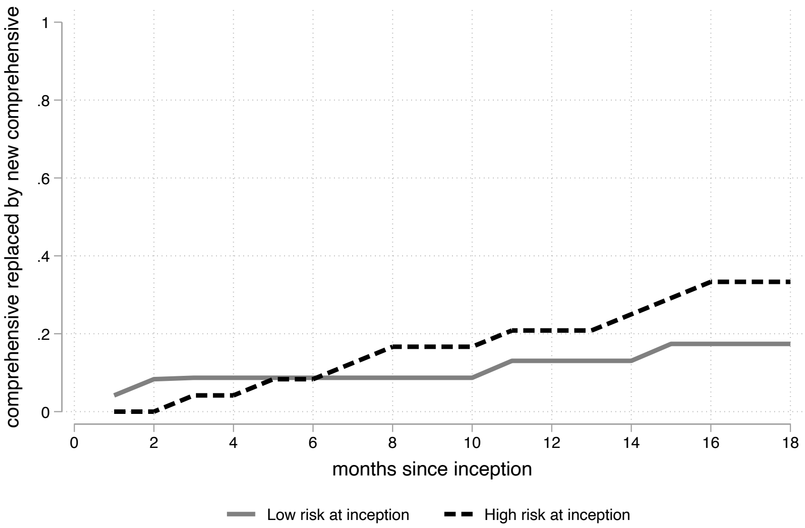

Figure 13: Share of comprehensive agreements replaced by new comprehensive agreements depending on level of risk at inception

Importantly, another way of reading Figure 13 is that the large majority of comprehensive agreements are still in place after 18 months. Does this mean that power-sharing agreements can be part of a broader shift towards more peaceful equilibria? Put differently, can countries escape the conflict trap with power-sharing agreements as part of the escape plan?

In this context it is worth asking whether there is any hope for long-term effects of power-sharing agreements, and through which channels it may act. One direct channel through which power-sharing agreements could affect long-term trajectories are political institutions. This is well-understood in policy circles. International organizations like the UN/DPPA stress that peace agreements have a clear link to "constitution making" (Berghof Foundation and UN/DPPA 2020) and this means they can have a profound impact on the development on political institutions through their role.

However, for policy purposes it is worth understanding what the institutional features are that are associated with such long-term changes in violence. For this purpose, we now turn towards the V-Dem dataset. Instead of trying to prove the causal effect of a specific mechanism our goal here is explore descriptive evidence of which cluster of institutional features relate to long-term falls of violence and which features change systematically when power-sharing agreements are introduced. Clean identification of effects is a lot harder here as we will look at changes of violence and institutions across decades. The following evidence should be regarded as providing possible channels instead of providing hard evidence for a causal effect of institutional features on violence.

6.2. Long-run institutional changes and violence

The Variants of Democracy (V-Dem) dataset allows us to analyze what type of institutional changes are associated with reductions in violence. We begin the analysis with the top layer of the dataset which captures the five different components of democracy that V-Dem measures. These are:

1. Deliberative: consultation and engaged society
2. Participatory: popular vote, elected local/regional government
3. Liberal: judicial & legislative constraints on executive
4. Electoral: clean elections, freedom of expression, suffrage
5. Egalitarian: equal protection and access

Figure 14 shows how strongly these elements are correlated with the extent of political violence occurrence and intensity. Throughout this section we control for country fixed effects and month fixed effects, i.e. the associations shown here control for fixed components like the country history and geography. Importantly, this implies that the results we show here are based on changes at the country level, i.e. our findings are based on realistic changes that have previously been observed at the country level.

The white dots in Figure 14 represent the average associations in our dataset. The bars indicate the uncertainty around this average experience. Broader bars indicate that specific histories can differ more from the average. In the top panel we find that all components of democracy measured by V-Dem are associated with reductions in violence – even when we control for country context and international context. We find the strongest associations in the liberal, electoral and egalitarian components and weaker associations with the deliberative and participatory components. The egalitarian component in the top panel of Figure 14 is particularly strongly associated with reductions in violence, which suggests that broadening horizontal and vertical inclusion can decrease violence. An increase of one standard deviation in this component is associated with a 10% reduction in violence.

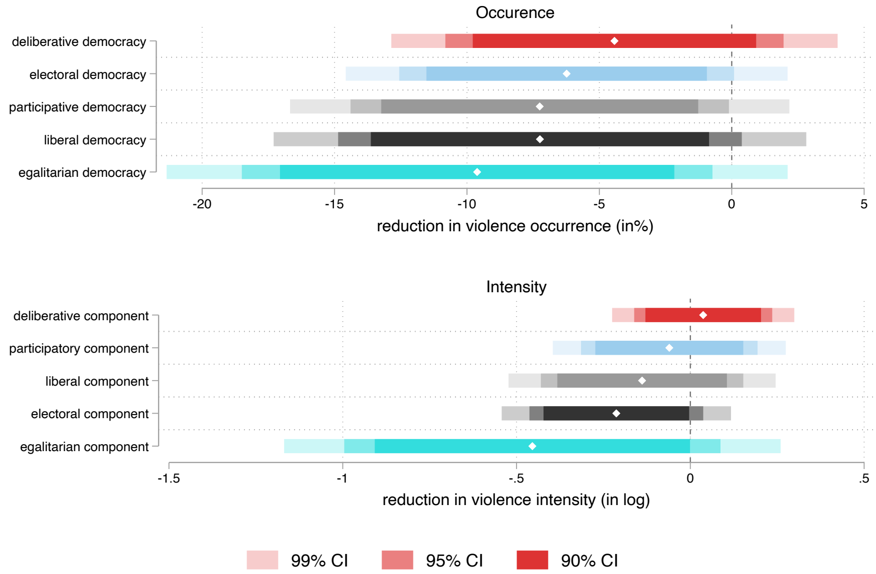

Figure 14: Facets of democracy and reductions in violence

Notes: The figure displays regression coefficients and their confidence intervals from regressing conflict occurrence (top) and intensity (bottom) on facets of democracy from V-Dem in a cross-country panel with fixed effects.

The top panel of Figure 15 looks at facets of the egalitarian component as measured by the V-Dem dataset. The elements that are most closely related to violence reductions are the absence of exclusion across political, social, socio-economic, gender or geographic dimensions. These variables capture access to power, public services, justice and civil liberties and whether these are restricted for specific groups.

The associations in the bottom panel of Figure 15 are large. Violence is reduced by between 15% and 25% when exclusion is reduced by one standard deviation. The association is even stronger if we look at violence intensity in the bottom panel of Figure 15. A coefficient size of slightly below -1 in Figure 15 implies that violence reduces by almost two thirds. This is a very strong and statistically significant association. Despite not knowing whether this is a causal effect, tracking exclusion indicators could be an important task for policymakers.

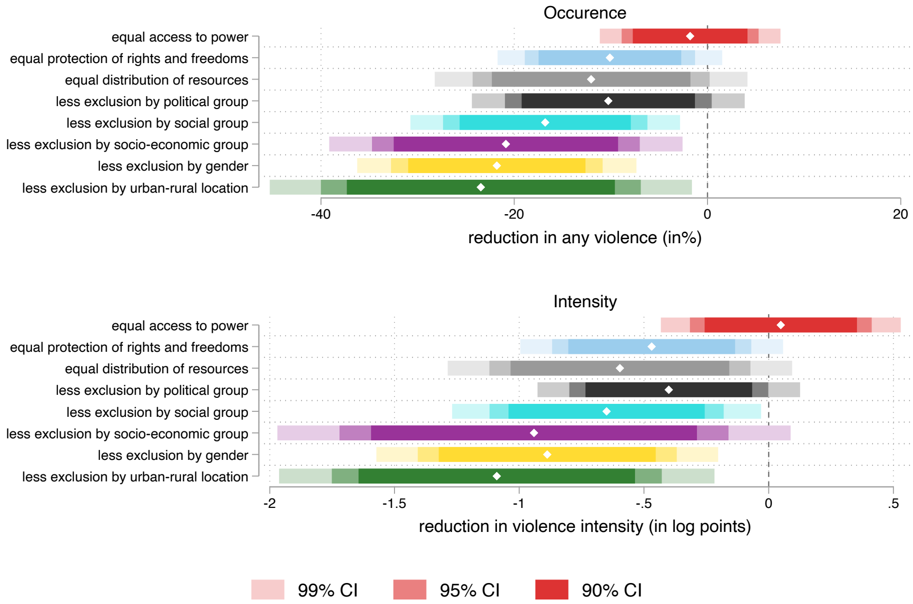

Figure 15: Egalitarian dimensions and reductions in violence

Notes: The figure displays regression coefficients and their confidence intervals from regressing conflict occurrence (top) and intensity (bottom) on egalitarian dimensions from V-Dem in a cross-country panel with fixed effects.

However, even if we believed that a part of theses associations is causal, running from institutions to reductions in violence, it is hard to take away concrete policy advice. What are the concrete institutional features that are most strongly associated with reductions in violence? Our analysis, summarized in Figure 16, suggests that improvements in the strength and neutrality of the legal system and public administration might play a key role. Likewise, fair access to public sector jobs and business opportunities are strongly associated with reductions in violence. The most significant reduction in violence is observed with strong and equal access to justice. A darling of the international community, free and fair elections, seems to be associated with reductions in violence but this association is not pronounced.

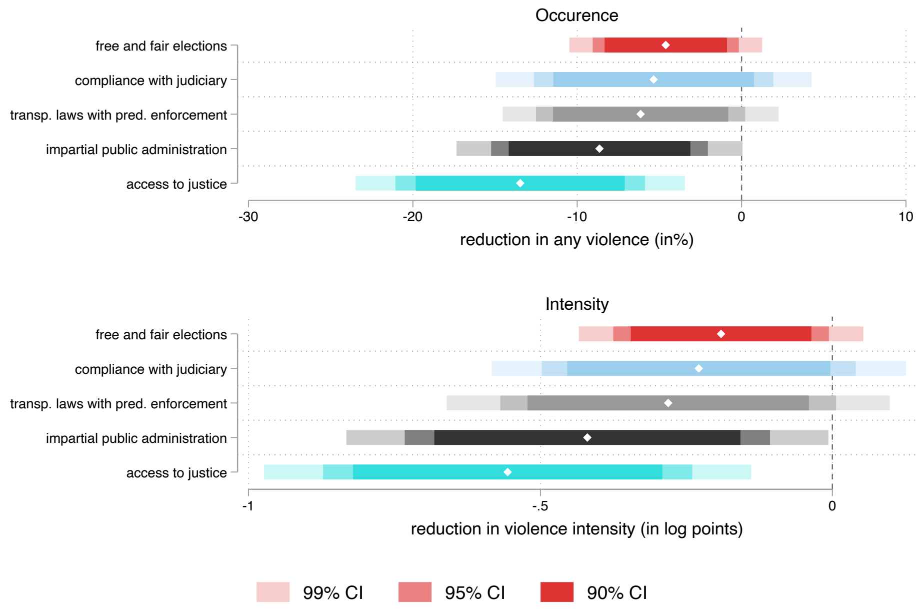

Figure 16: Institutional features and reductions in violence

Notes: The figure displays regression coefficients and their confidence intervals from regressing conflict occurrence (top) and intensity (bottom) on institutional features from V-Dem in a cross-country panel with fixed effects.

It is worth noting that these institutional changes do not only seem to capture a change of distributions of power inside the elite but suggest a shift in the rules and relationship between elite and population. The associations indicate that reductions in violence are associated with reduced exclusion along gender, geographic and socio-economic dimensions, stronger public service provision and access to justice.

6.3. Building the bridge with comprehensive agreements

Given these associations – what role does power sharing play in all of this? Is the intuition in policy circles correct that external help in peace agreements can facilitate institutions building? And why do comprehensive agreements appear to have a larger, amplifying effect over time?

To explore these questions, we exploit the monthly variation of V-Dem to use our difference-in-difference methodology. The control group to a treatment of a power-sharing agreement are again countries with a similar intensity forecast in the three months prior to the (placebo) adoption of an agreement. In our analysis, we will contrast the aftermath of all agreements and comprehensive agreements.

It is important to keep in mind that the different elements of power-sharing agreements shown in Figure 1 always coincide. At the same time, changes in dimensions in the V-Dem data also coincide. This makes it impossible to provide hard evidence on specific channels. The goal here is, therefore, merely to explore the changes in the aftermath of all and comprehensive agreements. The difference-in-difference methodology we rely on does, however, make us confident that the institutional changes we see are directly related to the adoption of power-sharing agreements.

In Figures 17 and 18 and Appendix Tables A3 and A4 we turn towards the broad V-Dem categories of democracy. These all increase to a similar extent. However, the point estimates are lower for the set of all power-sharing agreements than for the subset of comprehensive agreements. We, again, normalize each of the V-Dem variables by their respective standard deviation so that these results suggest that an increase in democracy scores following all power-sharing agreements of 5% of a standard deviation for all agreements (Figure 17) and 10% for comprehensive agreements (Figure 18).

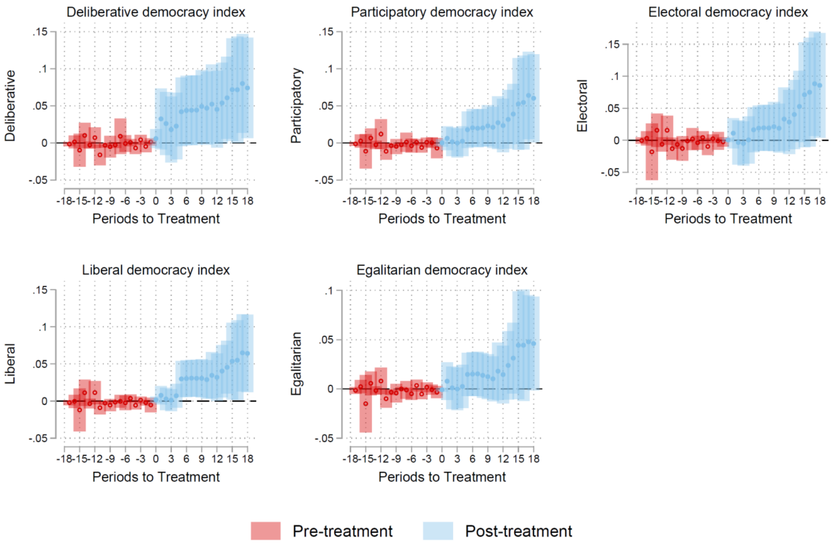

Figure 17: Democracy dimensions in the aftermath of all agreements

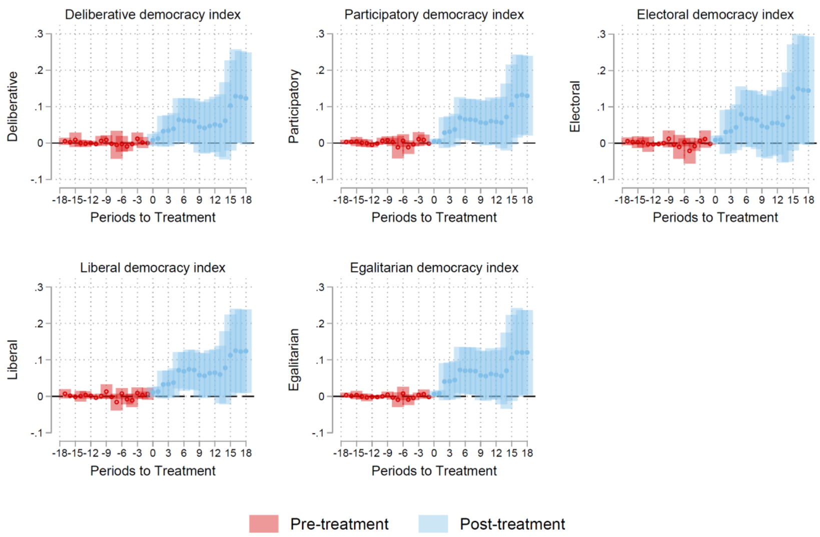

Figure 18: Democracy dimensions in the aftermath of comprehensive agreements

Notes: The figures show the estimated impact of agreements on dimensions of democracy in the V-Dem data in the 18 months before and after an agreement.

The figure does not really allow us to distinguish which institutional features improve. We therefore move one level lower in the aggregation by looking at the exclusion dimensions in Figures 19 and 20. Here we see some qualitative differences between all power-sharing agreements and the comprehensive agreements. In both cases we get the strongest decreases in exclusion along the lines of political groups and social groups. The changes are, again, slightly larger for comprehensive agreements. This makes a lot of sense given the goal of these agreements is to share power between political and social groups.

However, we see very little effects for socio-economic and urban-rural groups in all agreements and even some temporary worsening along the gender dimension. Here is where the comprehensive agreements are followed by much larger changes along all three categories. We know from the associations shown in Figure 15 that these are also the institutional dimensions that are most strongly associated with reductions in violence in the long run.[17]

In Appendix Figures A16 and A17 we also show that in the different dimensions of equality, comprehensive agreements are followed by much stronger changes and that the strongest difference seems to lie in the equal access dimension which captures power distributions by gender, social group and by socioeconomic position. Equal resource distributions change dramatically after comprehensive agreements. In Appendix Table A5 we show that the associations also hold when including all comprehensive agreements, no matter whether they are replaced or not.

Overall, these findings are in line with the view on power sharing as a solution to political bargaining. The distribution of power seems to change significantly in the aftermath of agreements. The larger and more diverse response to comprehensive agreements also suggests a possible explanation for why comprehensive agreements lead to stronger reductions in violence which are also building up over time. However, it is hard to pin it on a single factor with gender, rural-urban dimensions and the distribution of resources all playing important roles. In any case, it is hard to see these changes purely as a result of a changing distributions of power inside the elite.

[17] Our findings suggest an alternative interpretation to the mediation analysis conducted by Borman et al (2019) who show that de jure changes in the institutional set-up are mediated by behavioural changes. We show, using monthly instead of yearly data, that power-sharing agreements can spill over into broader institutional changes. In this view, institutional changes follow more informal arrangements and stabilize them.

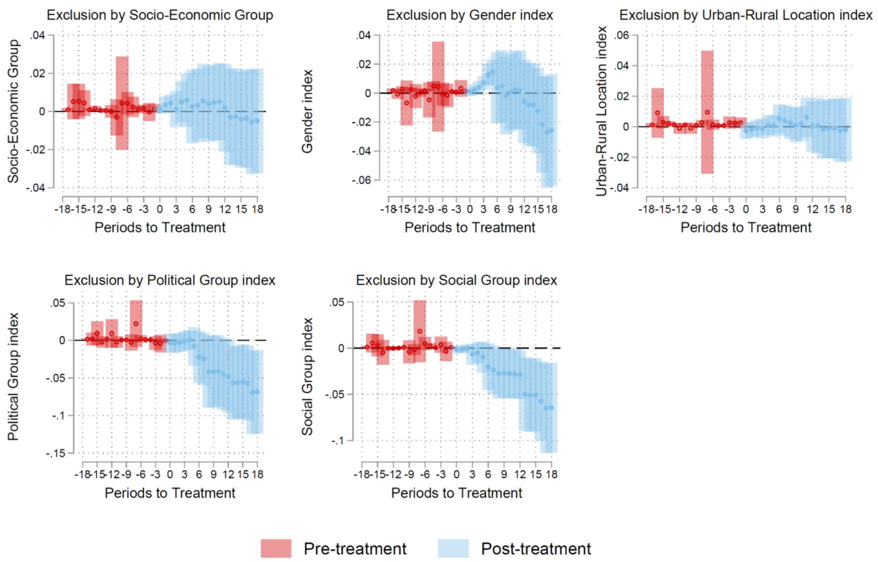

Figure 19: Exclusion dimensions in the aftermath of all power-sharing agreements

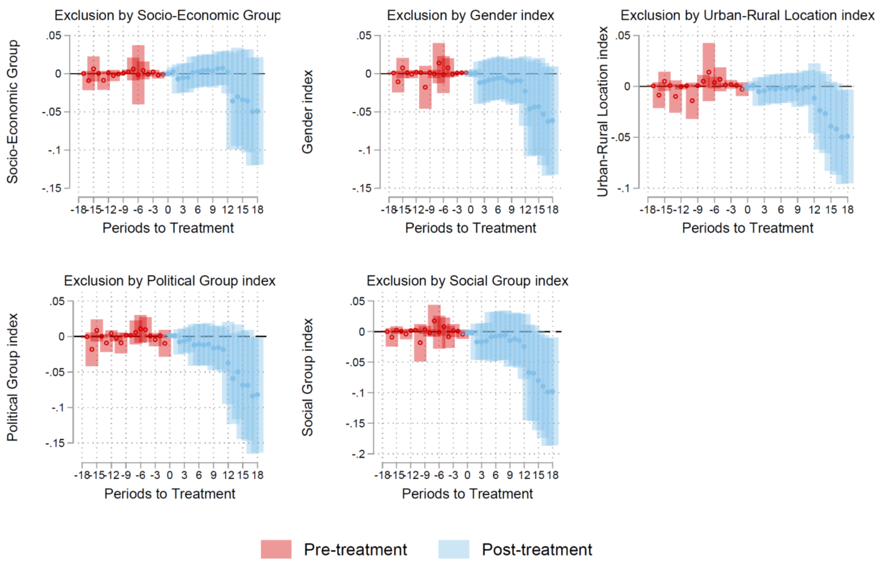

Figure 20: Exclusion dimensions in the aftermath of comprehensive agreements

Notes: The figures show the estimated impact of agreements on dimensions of exclusion in the V-Dem data in the 18 months before and after an agreement.

7. Discussion

Part of the reason why the conflict trap persists is that violence narrows the options for forward-looking decision making. This in turn is required for institutional or structural changes and for a sustainable reduction in violence. The conflict trap is also reflected in the level of power-sharing agreements, as more than half of all agreements are amended, or replaced by other agreements within a year.

This study shows that comprehensive power-sharing agreements nonetheless have a direct short and medium-term effect on violence. In some cases, the short-term reductions in armed violence seem to persist. We show that this effect is strongest for comprehensive agreements which are also associated with wider institutional developments and more dramatic reductions in violence.

Some countries escape the conflict trap. Power-sharing agreements should therefore be seen as both a short-term solution and a facilitator of broader long-run changes. We find that a mix of military power sharing and addressing vertical power relationships through passages treating governance, economic power-sharing transitional justice and human rights are particularly effective in an 18-month window. When we look at broader institutional changes, we find that gender, rural-urban dimensions, and the distribution of resources all shift in response. Further, comprehensive agreements seem to cover rights-based elements a lot more and access to justice is also most strongly associated with reductions of violence in the long run. This leaves justice as a key element that is present in all our results.

A caveat of our results is that the agreement of a power-sharing agreement could signal an international and internal context which is conducive to agreements and violence reductions. The attention by international actors and the help they provide after an agreement might contribute to reductions in violence. Also, the fact that the two conflict parties have built enough trust or resolved uncertainty regarding relative strengths could by themselves reduce violence and lead to the signature of an agreement. We cannot rule these explanations out entirely, but we have shown that our results are robust to controlling for a key action associated with international engagement (peacekeeping) and also do not change if we match countries according to their likelihood of coming to an agreement. In addition, the fact that we find sharp treatment effects are evidence against a more nuanced buildup of trust or reduction in uncertainty being the root cause of the violence reduction. Results should still be interpreted as showing the effect of coming to an agreement endogenously, i.e. with everything this entails in terms of international and local environment. Imposing agreements internationally will likely not have the same effect.

This is also where our findings echo the case study in the Pathways for Peace report by the World Bank/UN (2018), who conclude that countries that find pathways to sustainable peace have eventually tackled the messy and contested process of institutional reform. Often, the transition moment that led to sustainable peace is based on a shift away from security-led responses and toward broader approaches that mobilized a range of sectors in support of institutional reforms.

The quantitative results in this report can directly be interpreted within the influential framework proposed by Cheng et al. (2018), who stress that interactions and (mis)alignments between political settlements, elite bargaining and peace agreements may explain whether and how wars are terminated, and differing trajectories of post-war transition. According to Cheng et al. (2018), large-scale violence will only stabilize *"when the distribution of benefits in a society, supported by its institutions (e.g., political positions, business opportunities) is consistent with the distribution of power in society, and the economic and political outcomes of these institutions are sustainable over time"*. They stress that this includes both the horizontal relationships between different parts of the elites and the vertical relationships between elites and their constituencies. Our results can give some hints with regard to binding constraints. Military power-sharing seems to be an extremely powerful element for violence reduction which reflects the importance of tailoring agreements to de-facto military power. But our results also show that governance, human rights, and transitional justice are key elements. This suggests that the most stable and peaceful bargain results seem to be those that manage to combine a solution for the horizontal elite bargain with institutional changes that address the vertical dimension.

But there are other, complementary views to our findings which suggest that power-sharing agreements could be an entry point into changing the logic of the elite bargain itself by introducing elements of public goods, such as access to justice. Besley and Persson (2011) argue that the incentives of the incumbent government to invest in state capacity is key to understand how economic development, the distribution of resources and political violence evolve in the long run. Investment incentives increase when either institutions are cohesive, power is not contested, or the state is needed by the incumbent group to provide non-excludable public services. North et al (2007) propose that fragile countries represent a *limited access order* where elites use the state order to extract rents. Escaping this set-up is a necessary condition to escape the conflict trap. In this view, peace agreements need to complement the elite bargain with elements that allow for a change in the overall equilibrium in which the state can provide broader benefits. In this view, the goal of power-sharing agreements should be to improve social cohesion and frame the state as providing services, as compared to distributing rents. Put differently, the elite needs to be able to shift towards providing broader benefits beyond their narrow in-group. This makes agreements of power sharing more robust to shifts in de facto power.

How realistic is this? The differences between comprehensive and all power-sharing agreements we find are consistent with this view. Also, it has been shown that a sense of national identity can be affected by single events (Depetris-Chauvin et al 2020). If a peace agreement establishes a nucleus of unity this might be a starting point for broader legitimacy.

Rohner et al (2013) point to lack of trust between the different groups in conflict as one of the reasons of the conflict trap. Practitioners involved in mediation and peace talks stress the importance of building trust during negotiations (Freeman and Clark 2020). The reduction of exclusion, protection of rights and access to justice might be essential elements in re-building this trust towards other groups and, hence, provide an escape route out of the conflict trap.

Several caveats apply to this project. First, identification hinges on our controls for predicted future risk. If these fail to capture conflict dynamics systematically correlated with the adoption of a power-sharing agreement, the result may be a biased estimate of the effect of a power-sharing agreement. The biggest problem for our quantitative estimates occurs if other unrelated initiatives co-occur with the month of a power-sharing agreement. We see this as unlikely given the sharp, monthly variation we exploit. Our results should nonetheless be seen in this context – agreements are not concluded in isolation but impacted by the initiatives of a range of actors, including external actors, the provision of financial incentives (aid), and security guarantees (deployment of peacekeepers). Quantitative researchers could, given data on state visits, aid, peacekeeping or adopted UN resolutions, try to disentangle the effects of these different contextual elements.

Second, if armed groups anticipate that agreements will cement the distribution of political power, then it is possible that violence increases before an agreement is struck. However, our results never suggest clear pre-trends before agreements and are robust to matching on earlier risk data. This makes us confident that short-term tactical considerations in the context of negotiations do not drive our results. A clear way forward for research is a further disaggregation of our quantitative analysis. With an actor-based focus it would be possible to see whether actors that are excluded from a peace agreement are more likely to engage in violence than those that are included. Insights could also be gained from a geographic disaggregation in which the participation or exclusion of specific ethnic groups could be linked to spatially disaggregated violence data. Another possibility is the development of a fully dynamic model in which armed political violence is modelled jointly with the timing and content of peace agreements. Recent theoretical work has opened avenues in this direction (Meirowitz et al 2019).

A clear limitation for practitioners is that deals with external actors or actors which are shunned by the international community might yield results but may have a legal, strategic, and moral price which is too high to pay. Our quantitative results should be interpreted with this in mind: all agreements are trying to achieve something very difficult in a particular context. In this context their effect is even more remarkable.

REFERENCES

Abadie, A (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59 (2): 391-425.

Acemoglu, D., & Robinson, J.A. (2001). A theory of political transitions. *American Economic Review*, 91(4), 938-963.

Acemoglu, D. and J.A. Robinson (2006) De Facto Political Power and Institutional Persistence. *American Economic Review* , 96(2), 325-330.

Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2005). From education to democracy?. *American Economic Review*, 95(2), 44-49.

Armand, A., Marending, M., and Vysotskaya, G. (2023). Windows of peace: the effect of ceasefires on economic recovery (No. wp-2023-35). World Institute for Development Economic Research (UNU-WIDER).

Bell, C., S. Badanjak, J. Beujouan, R. Forster, T. Epple, A. Jamar, K. McNicholl, S. Molloy, K. Nash, J. Pospisil, R. Wilson, L. Wise (2021). PA-X Codebook, Version 5. Political Settlements Research Programme, University of Edinburgh, Edinburgh. www.peaceagreements.org

Bell, C. and Badanjak, S. (2019) 'Introducing PA-X: A new peace agreement database and dataset', *Journal of Peace Research*, 56 (3).

Berghof Foundation and the United Nations Department of Political and Peacebuilding Affairs (2020). Constitutions and Peace Processes: A Primer. Berlin: Berghof Foundation.

Besley, T. and Persson, T. (2011). Pillars of Prosperity: The Political Economics of Development Clusters, Princeton: Princeton University Press. https://doi.org/10.1515/9781400840526

Blattman, C. (2022). Why We Fight: The Roots of War and the Paths to Peace. Penguin.

Bormann, Nils-Christian, Lars-Erik Cederman, Scott Gates, Benjamin A. T.Graham, Simon Hug, Kaare W.S. and J. Wucherpfennig (2019) Power Sharing: Institutions, Behavior, and Peace. American Journal of Political Science, 63(1), pp. 84-100.

Borusyak, K., Jaravel, X., and Spiess, J. (2022). Revisiting Event Study Designs: Robust and Efficient Estimation. Available at SSRN 2826228.

Callaway, B. and P.H.C. Sant'Anna (2021) Difference-in-Differences with multiple time periods, Journal of Econometrics, 225(2).

Canidio, Andrea and Joan Esteban (2022) Optimal Political Institutions in the Shadow of Conflict. Unpublished Manuscript.

Cheng, C., Goodhand, J., and Meehan, P. (2018). Synthesis paper: Securing and sustaining elite bargains that reduce violent conflict.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal*, 21(1).

Collier, Paul, and Nicholas Sambanis (2002). Understanding Civil War: A New Agenda. *Journal of Conflict Resolution*, 46(1): 3–12.

Coppedge, M., J. Gerring, C.H. Knutsen, S.I. Lindberg, J. Teorell, N. Alizada, D. Altman, M. Bernhard, A. Cornell, M.S. Fish, L. Gastaldi, H. Gjerløw, A. Glynn, A. Hicken, G. Hindle, N. Ilchenko, J. Krusell, A. Luhrmann, S.F. Maerz, K.L. Marquardt, K. McMann, V. Mechkova, J. Medzihorsky, P. Paxton, D. Pemstein, J. Pernes, J. von Römer, B. Seim, R. Sigman, S.-E. Skaaning, J. Staton, A. Sundström, E. Tzelgov, Y. Wang, T. Wig, S. Wilson and D. Ziblatt. (2021)."V-Dem [Country–Year/Country–Date] Dataset v11.1" Varieties of Democracy Project. https://doi.org/10.23696/vdemds21

Croicu, M. and R. Sundberg, (2016), "UCDP GED Codebook version 5.0", Department of Peace and Conflict Research, Uppsala University.

Davies, S., T. Pettersson and M. Öberg (2022). Organized violence 1989-2021 and drone warfare. *Journal of Peace Research* 59(4).

De Chaisemartin, C., and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964-2996.

Depetris-Chauvin, E., R. Durante, and F. Campante. (2020). "Building Nations through Shared Experiences: Evidence from African Football." *American Economic Review*, 110 (5): 1572-1602.

Fearon, J. D. (1995). Rationalist explanations for war. International organization, 49(3), 379-414.

Fearon, J. D. (2004). Why Do Some Civil Wars Last So Much Longer than Others? Journal of Peace Research, 41(3), 275–301.

Fetzer, T., and Kyburz, S. (2022). Cohesive institutions and political violence. *Review of Economics and Statistics*, 1-46.

Fortna, V. (2003). Scraps of Paper? Agreements and the Durability of Peace. International Organization, 57(2), 337-372.

Francois, P., Rainer, I. and Trebbi, F., (2015). How is power shared in Africa?. *Econometrica*, 83(2), pp.465-503.

Freeman, M. and P. Clark (2020). Building Peace from the Middle. Peace Science Digest, Special Issue, 4.

Gates, S., Graham, B. A., Lupu, Y., Strand, H., and Strøm, K. W. (2016). Power sharing, protection, and peace. *The Journal of Politics*, 78(2), 512-526.

Hörner, J., Morelli, M., and Squintani, F. (2015). Mediation and peace. The Review of Economic Studies, 82(4), 1483-1501.

Krasner, S. (2002) Realist Views of International Law. *Proceedings of the Annual Meeting (American Society of International Law)*, 96 (13-16), pp. 265-268.

Licklider, R. (1995). The Consequences of Negotiated Settlements in Civil Wars, 1945–1993. *American Political Science Review*, 89(3), 681-690.

Meirowitz, A., M. Morelli, K. Ramsay and F. Squintani (2019) Dispute Resolution Institutions and Strategic Militarization, *Journal of Political Economy*, vol. 127(1), 378-418.

Mueller, H. and C. Rauh (2021). https://www.conflictforecast.org

Mueller, H. and C. Rauh (2022a). The Hard Problem of Prediction for Conflict Prevention. *Journal of the European Economic Association*, Journal of European Economic Association, vol. 20 (6), p.2440–2467.

Mueller, H. and C. Rauh (2022b). Using Past Violence and Current News to Predict Changes in Violence, *International Interactions*, 48:4, 579-596.

North, DC., Wallis, J.J., Webb, S.B. and Weingast, B R (2007). Limited Access Orders in the Developing World: A New Approach to the Problems of Development. Policy Research Working Paper; No. 4359. World Bank.

Nunn, N., and Qian, N. (2014). US food aid and civil conflict. American Economic Review, 104(6), 1630-1666.

Rios-Avila, F., A. Naqvi, and P.H.C. Sant'Anna (2021) drdid: Doubly robust difference-in-differences estimators in Stata, Working Paper, 2021.

Rohner, D., M. Thoenig and F. Zilibotti (2013) War Signals: A Theory of Trade, Trust, and Conflict. *The Review of Economic Studies*, Vol. 80, No. 3 (284): pp. 1114-1147.

Rohner, D., and Thoenig, M. (2021). The elusive peace dividend of development policy: From war traps to macro complementarities. *Annual Review of Economics*, 13, 111-131.

Sant'Anna, P. H., and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1), 101-122.

Srinivasan, S. (2021). When peace kills politics: international intervention and unending wars in the Sudans. Hurst Publishers.

Strøm, K., Gates, S., Graham, B., and Strand, H. (2017). Inclusion, Dispersion, and Constraint: Powersharing in the World's States, 1975–2010. *British Journal of Political Science*, 47(1), 165-185. doi:10.1017/S0007123415000174

Sundberg, R., and E. Melander (2013), "Introducing the UCDP Georeferenced Event Dataset", *Journal of Peace Research*, vol.50, no.4, 523-532

United Nations; World Bank. 2018. Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict. Washington, DC: World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/28337 License: CC BY 3.0 IGO.

Walter, Barbara (2009) Bargaining Failures and Civil War. Annual Review of Political Science, 12:243–61.

World Bank (2022) World Development Indicators." https://datatopics.worldbank.org/world-development-indicators/

APPENDIX

A Predicting power-sharing agreements

In this section we describe the prediction of power-sharing agreements through a cross-validation fit on the entire data. We use the same model as when predicting conflict outbreaks with the addition of two counts: the count of months since the last power sharing agreement and the count of months since a comprehensive agreement.

We use a random forest classifier to predict agreements three months ahead. The coding we use defines the y-variable as “an agreement within the next three months”, i.e. this includes all instances one, two or three months in the future.

We predict all agreements and comprehensive agreements. The procedure is the same in both cases: we first optimize the hyperparameters using a 11-fold cross validation method. We take many folds to make sure enough agreements are always in the training data. For both types of agreements we end up with a relatively flat tree and high minimum observations in the leafs. The maximum depth is 4, the minimum sample leaves are 200 and we use 700 trees. This suggests that overfitting is a problem that the cross-validation reveals.

Figure A1 displays the feature importance of the 15 most important predictors for all power-sharing agreements (top) and comprehensive agreements (bottom). The number of months since the last agreement and conflict dynamics are what contribute most to the prediction of power-sharing agreements. As when predicting conflict, the "armed conflict" topic is the most

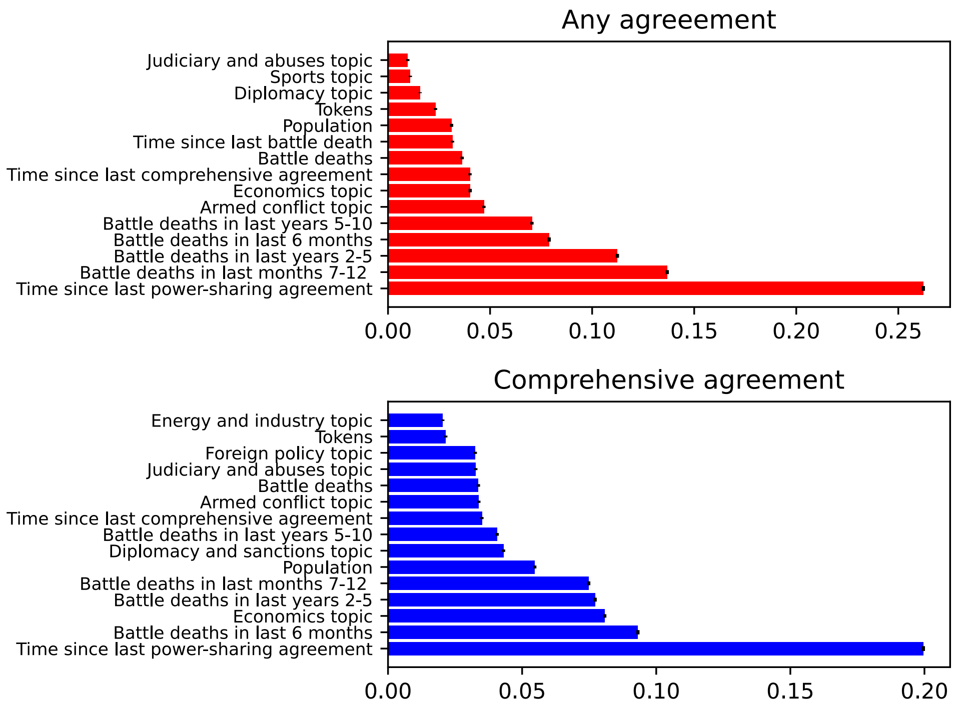

Figure A1: Feature importance of random forest predicting agreements

important text model followed by other models like "economics" and "diplomacy".

We then produce cross-validated fitted values so that we produce out-of-sample fitted scores for agreements and not only for conflict outbreaks. When predicting all agreements, we reach an ROC-AUC of 0.9 and an average precision score of 0.13. For comprehensive agreements the ROC-AUC is even better 0.92 but average precision is terrible – just 0.04. Low precisions despite high ROC-AUCs is a common feature of imbalanced classes with many zeros than ones.

Note, that precision is low even when predicting all agreements. This means, despite a very good ROC-AUC score indicating a strong forecast system, it is very hard to put any confidence on a positive prediction. It is simply very unlikely that any given country concludes a peace agreement in a specific month (quarter) and this makes the task of predicting agreements very difficult. We conclude that power-sharing agreements are hard to predict – or at least with our forecasting procedure.

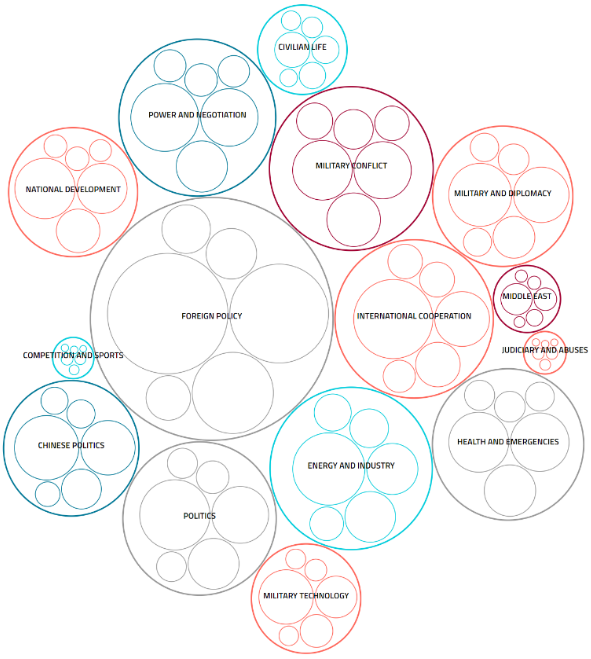

Figure A2: News topics in June 2014 for Mozambique

Notes: The snapshot was retrieved from conflictforecast.org on 24/02/2023.

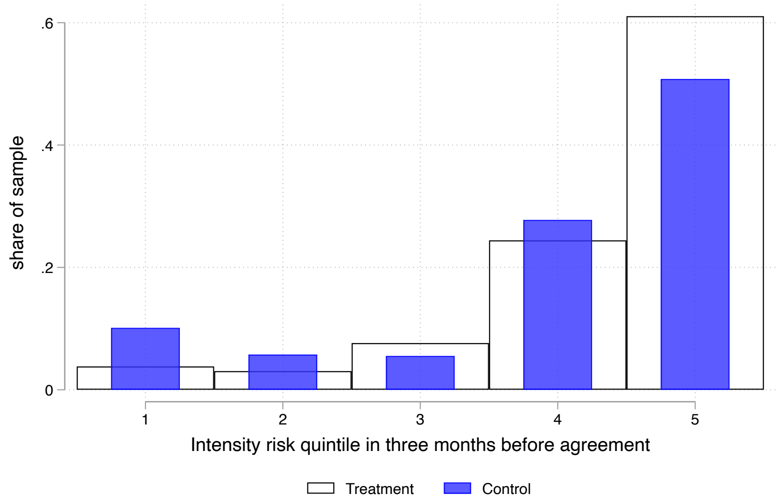

Figure A3: Distribution in terms of pre-treatment predicted intensity risk of all peace agreements in treatment and control group for difference-in-difference analysis

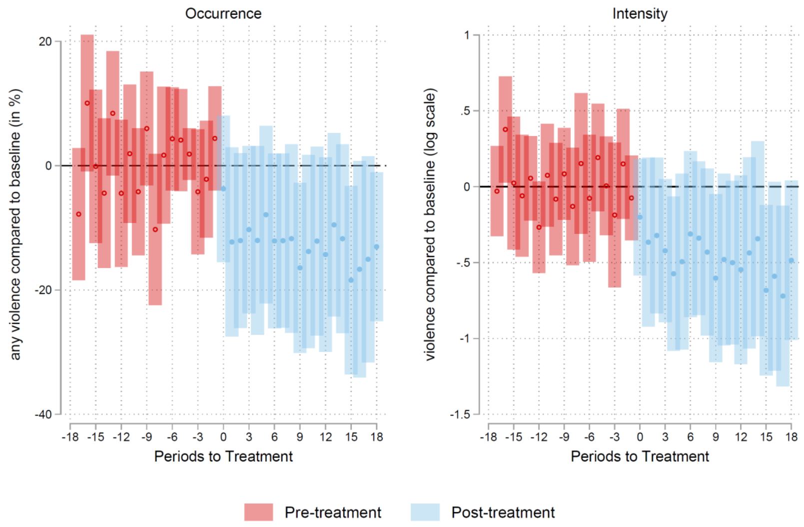

Figure A4: Violence occurrence and intensity after power sharing (18-month window)

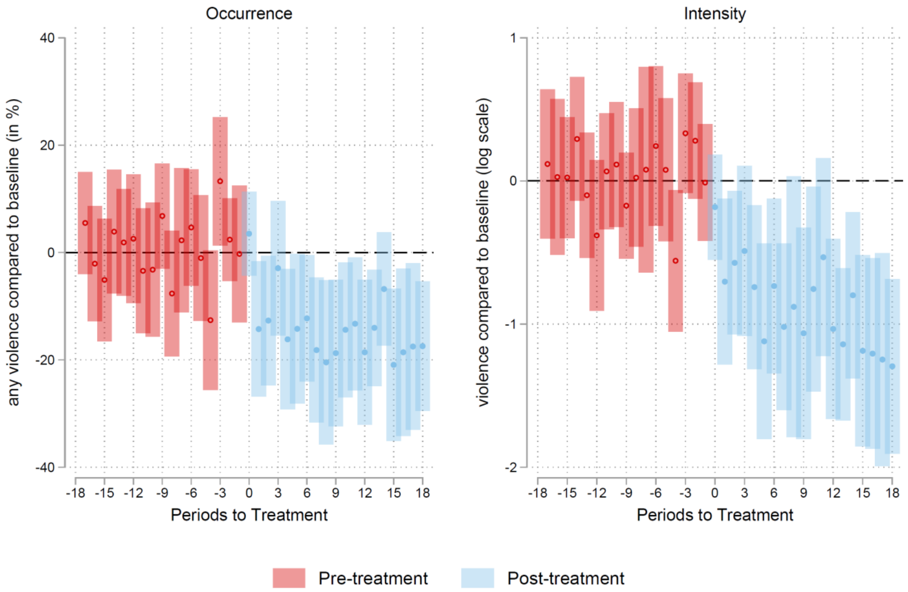

Figure A5: Violence occurrence and intensity after comprehensive agreements (18-month window)

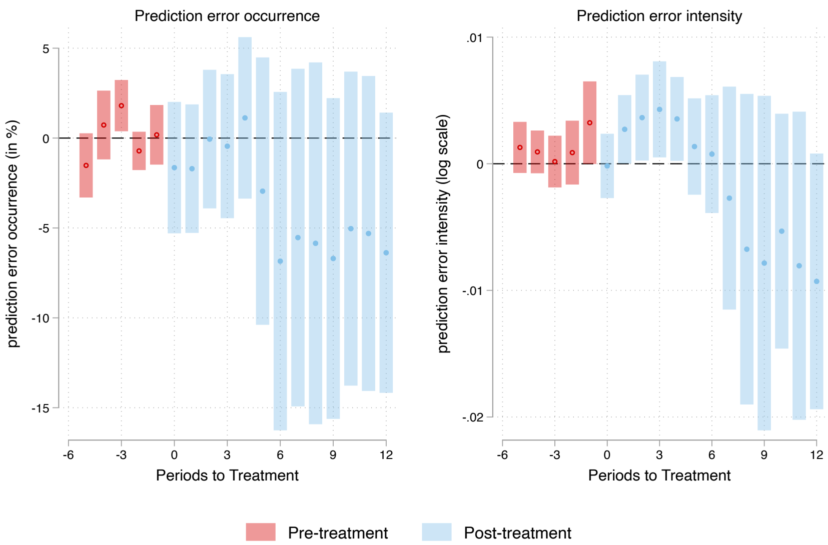

Figure A6: Forecast error for comprehensive power-sharing agreements

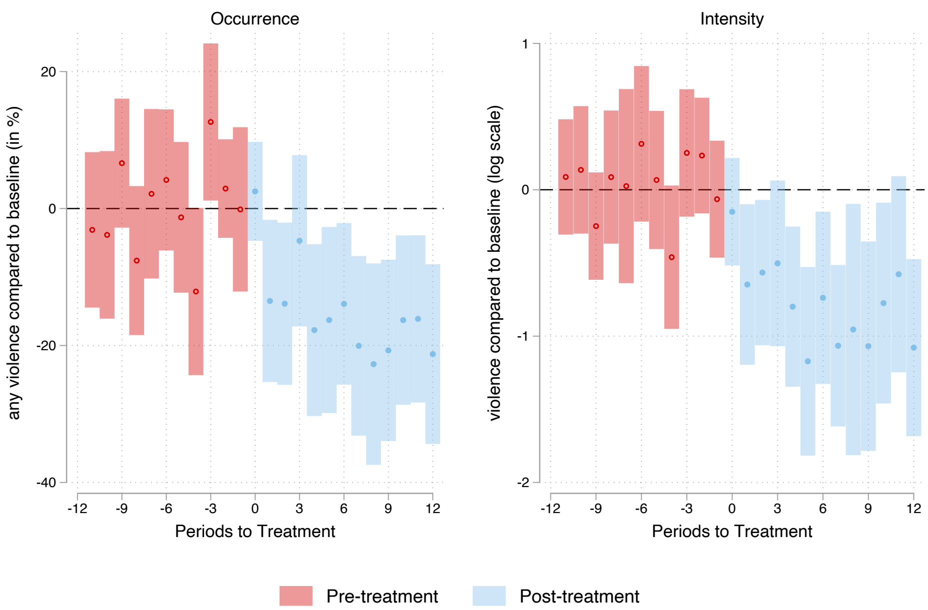

Figure A7: Violence occurrence and intensity after comprehensive power sharing matching on conflict occurrence likelihood

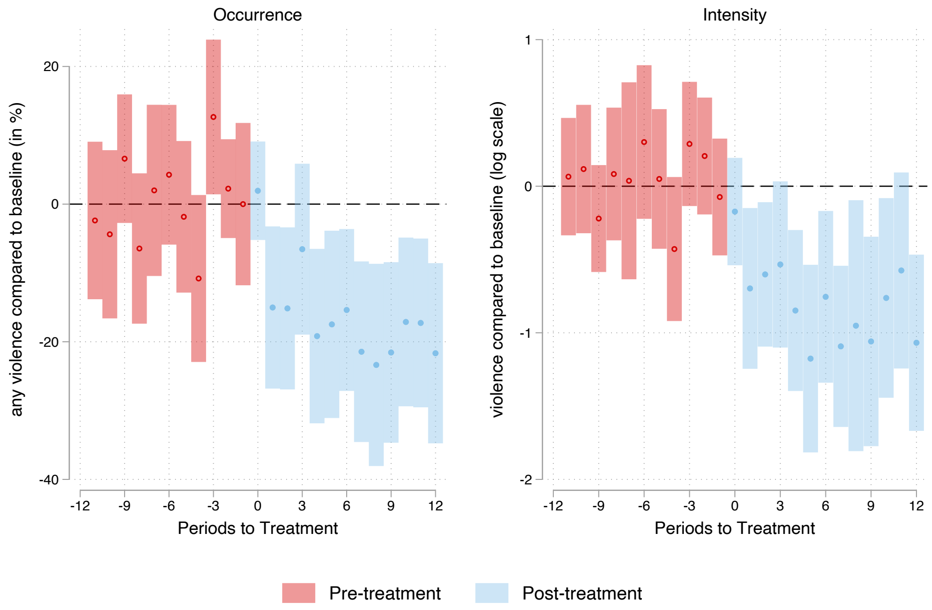

Figure A8: Violence occurrence and intensity after comprehensive agreement with random matching

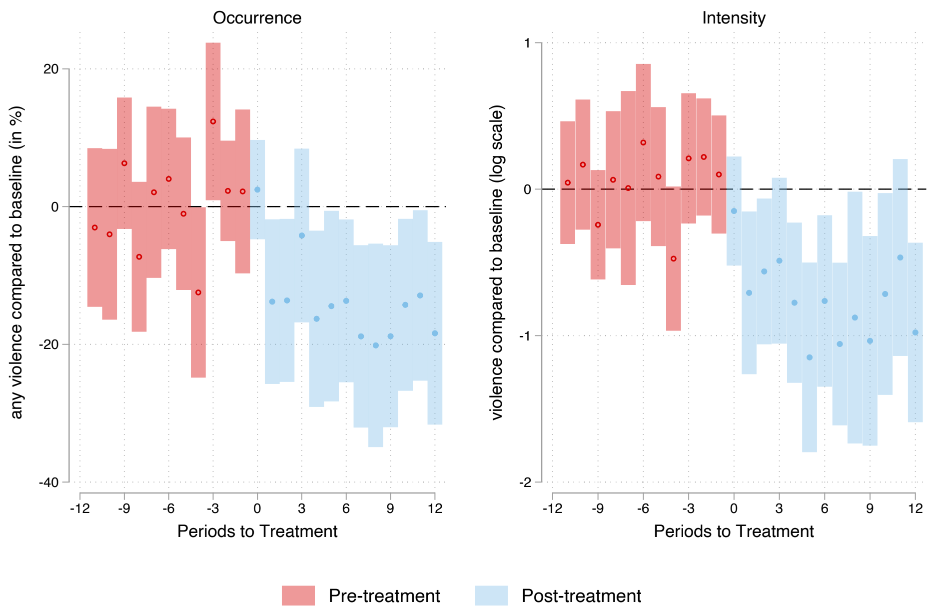

Figure A9: Violence occurrence and intensity after comprehensive agreement with matching on predicted likelihood of comprehensive agreement

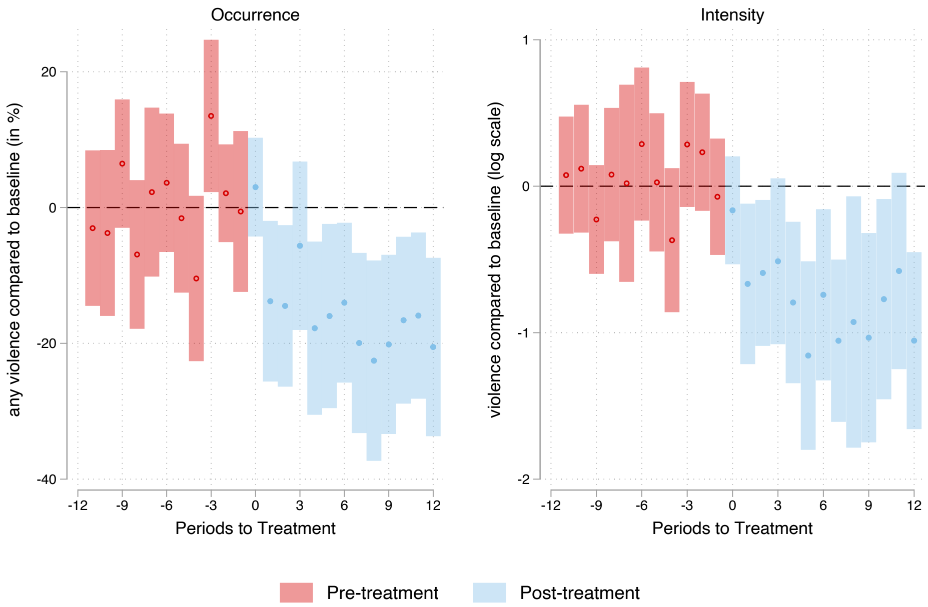

Figure A10: Violence occurrence and intensity after comprehensive agreement matching on 4-6 months before agreement

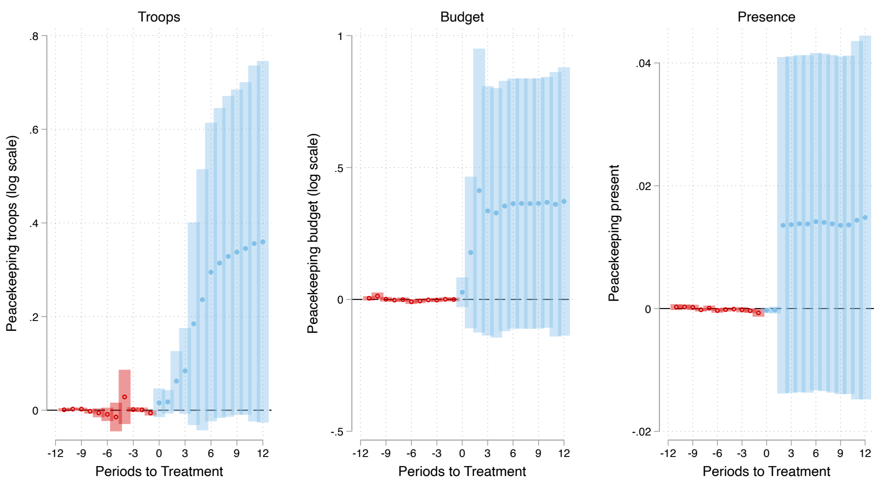

Figure A11: Peacekeeping activity after all power-sharing agreements

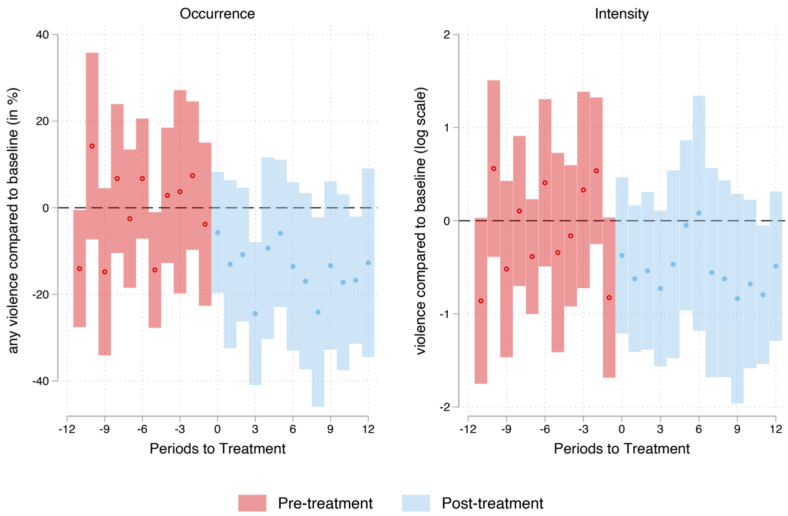

Figure A12: Violence occurrence and intensity when peacekeeping troops appear

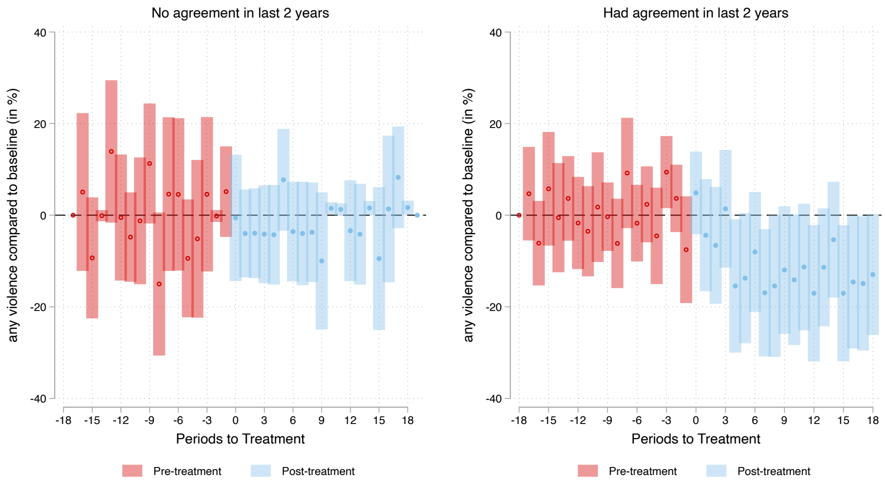

Figure A13: Violence occurrence after any power-sharing agreement for agreements without previous agreement and with agreement within past two years

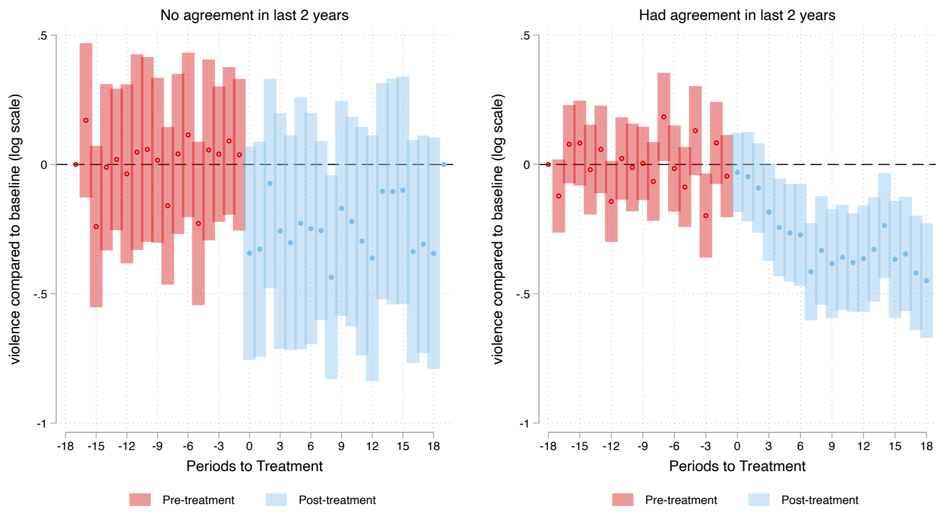

Figure A14: Violence intensity after any power-sharing agreement for agreements without previous agreement and with agreement within past two years

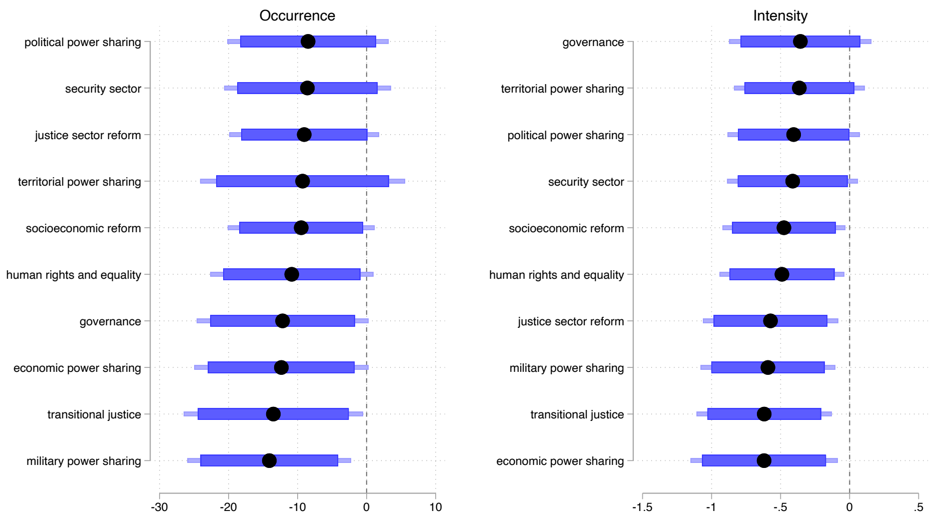

Figure A15: ATT for different components of agreements

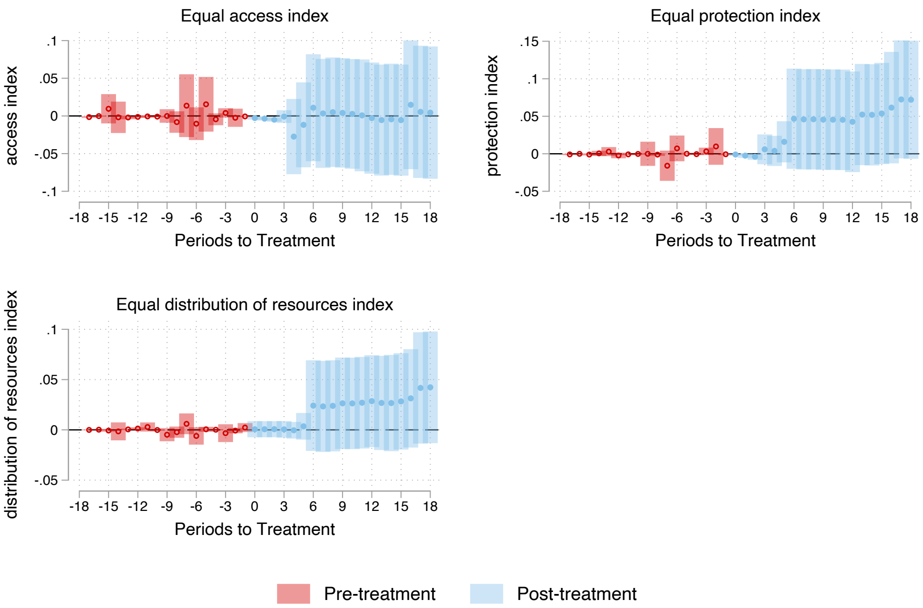

Figure A16: Equality dimensions in the aftermath of all power-sharing agreements

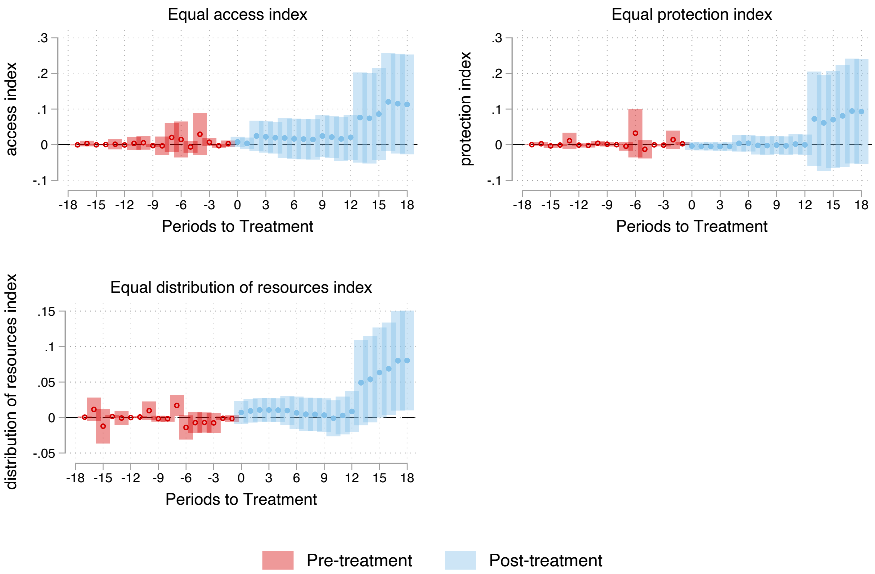

Figure A17: Equality dimensions in the aftermath of comprehensive power-sharing agreements

Table A1: Summary statistics

Variable	Obs	Mean	Std. dev.	Min	Max
Any power-sharing agreement	73,401	0.0060	0.0772	0	1
Comprehensive agreement	73,401	0.0010	0.0315	0	1
Ln(fatalities+1)	73,401	0.5568	1.4614	0	13.1503
Occurrence of violence	73,401	0.1601	0.3667	0	1
Ln(peacekeeping budget+1)	73,401	1.0491	3.9306	0	20.2203
Ln(peacekeeping troops+1)	73,401	0.3500	1.6182	0	10.5614
Conflict risk forecast	43,770	0.2764	0.3378	0.0016	0.9643
Intensity per capita forecast	43,770	0.0016	0.0081	0.0000	0.2622

Table A2: ATT of power-sharing agreements on conflict occurrence and intensity using different matching strategies

	(1)	(2)	(3)	(4)	(5)	(6)
	Random matching		Conflict likelihood		Predicted agreement	
	any	intensity	any	intensity	any	intensity
Panel A: All power-sharing agreements						
ATT power sharing	-13.085**	-0.474*	-12.366*	-0.465*	-11.242*	-0.419*
	(6.477)	(0.250)	(6.472)	(0.249)	(6.486)	(0.250)
Panel B: Comprehensive power-sharing agreements						
ATT power sharing	-16.287***	-0.932***	-15.094***	-0.917***	-13.730***	-0.884***
	(4.806)	(0.265)	(4.796)	(0.266)	(4.898)	(0.267)

Robust standard errors, clustered at the country level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Any violence is equal to 100 if there is any fatality according to UCDP in that month and zero otherwise. Violence intensity are log(fatalities +1). All regressions restrict the sample to a window of 18 months around the adoption of power-sharing agreements for treated countries and control for month fixed effects. The control group is a random sample of countries without power-sharing agreements in the first two columns, with the same distribution of predicted conflict likelihood in columns (3) and (4), and with the same distribution of predicted likelihoods of respective agreements in the last two columns.

Table A3: ATT of comprehensive agreements on V-Dem components

	(1)	(2)	(3)	(4)	(5)
Democracy index					
	Deliberative	Participatory	Electoral	Liberal	Egalitarian
ATT power sharing	0.054**	0.056**	0.057*	0.059***	0.058**
	(0.025)	(0.023)	(0.030)	(0.022)	(0.025)
Power distribution					
	SES	Social group	Gender	Sex orient.	Urban-rural
ATT power sharing	0.000	0.045	-0.009	0.015	0.017*
	(0.033)	(0.031)	(0.008)	(0.022)	(0.010)
Equality index					
	Access	Protection	Resources		
ATT power sharing	0.018	0.001	0.013		
	(0.021)	(0.009)	(0.011)		
Exclusion index					
	SES	Gender	Urban-rural	Political	Social group
ATT power sharing	0.000	-0.007	-0.001	-0.015	-0.013
	(0.008)	(0.011)	(0.006)	(0.011)	(0.015)
Institution index					
	Elections	Judiciary	Law enforce.	Public admin.	Justice
ATT power sharing	0.012	-0.034*	0.023	0.002	0.011
	(0.058)	(0.019)	(0.023)	(0.012)	(0.011)

Robust standard errors, clustered at the country-window level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The respective dependent variables are listed in the sub-headings and are standardized indices. All regressions restrict the sample to an 18-month window around the adoption of comprehensive power-sharing agreements for treated countries and control for month fixed effects. The randomized control group has the same distribution as the treatment group in terms of predicted conflict intensity before adoption.

Table A4: ATT of all power-sharing agreements on V-Dem components

	(1)	(2)	(3)	(4)	(5)
Democracy index					
	Deliberative	Participatory	Electoral	Liberal	Egalitarian
ATT power sharing	0.036*	0.023	0.020	0.030	0.014
	(0.021)	(0.018)	(0.020)	(0.020)	(0.013)
Power distribution					
	SES	Social group	Gender	Sex orient.	Urban-rural
ATT power sharing	-0.022	0.015	0.009	-0.005	-0.022
	(0.020)	(0.015)	(0.019)	(0.004)	(0.021)
Equality index					
	Access	Protection	Resources		
ATT power sharing	0.003	0.02	0.009		
	-0.017	-0.014	-0.009		
Exclusion index					
	SES	Gender	Urban-rural	Political	Social group
ATT power sharing	0.002	0.007	0.003	-0.014	-0.008
	(0.005)	(0.005)	(0.003)	(0.009)	(0.006)
Institution index					
	Elections	Judiciary	Law enforce.	Public admin.	Justice
ATT power sharing	-0.009	-0.001	0.017	-0.009	0.025**
	(0.030)	(0.008)	(0.019)	(0.014)	(0.012)

Robust standard errors, clustered at the country-window level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The respective dependent variables are listed in the sub-headings and are standardized indices. All regressions restrict the sample to an 18-month window around the adoption of all power-sharing agreements for treated countries and control for month fixed effects. The randomized control group has the same distribution as the treatment group in terms of predicted conflict intensity before adoption.

Table A5: ATT of comprehensive agreements on V-Dem components (whether agreement replaced or not)

	(1)	(2)	(3)	(4)	(5)
Democracy index					
	Deliberative	Participatory	Electoral	Liberal	Egalitarian
ATT power sharing	0.056**	0.059***	0.063**	0.052**	0.054**
	(0.024)	(0.021)	(0.028)	(0.020)	(0.022)
Power distribution					
	SES	Social group	Gender	Sex orient.	Urban-rural
ATT power sharing	0.010	0.064**	-0.008	0.015	0.022*
	(0.026)	(0.026)	(0.010)	(0.016)	(0.013)
Equality index					
	Access	Protection	Resources		
ATT power sharing	0.030	0.022	0.024**		
	(0.020)	(0.017)	(0.012)		
Exclusion index					
	SES	Gender	Urban-rural	Political	Social group
ATT power sharing	-0.008	-0.015	-0.003	-0.013	-0.024*
	(0.009)	(0.011)	(0.007)	(0.015)	(0.014)
Institution index					
	Elections	Judiciary	Law enforce.	Public admin.	Justice
ATT power sharing	0.055	-0.003	0.027	-0.007	0.030*
	(0.047)	(0.020)	(0.022)	(0.016)	(0.015)

Robust standard errors, clustered at the country-window level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The respective dependent variables are listed in the sub-headings and are standardized indices. All regressions restrict the sample to an 18-month window around the adoption of power-sharing agreements for treated countries and control for month fixed effects. The randomized control group has the same distribution as the treatment group in terms of predicted conflict intensity.